1 Relative Frequency

- $f_k(n) = \frac{N_k(n)}{n} \leftarrow$ **Relative Frequency**
  - k is the outcome
  - $N_k(n)$ is the number of times outcome k
- $\lim\limits_{n\to\infty} f_k(n) = p_k \leftarrow$ **Statistical Regularity**
  - p_k is the probability of event k occurring

1.1 Properties of Relative Frequencies

1. $f_k(n) = \frac{N_k(n)}{n}$
2. $0 \leq N_k(n) \leq n$
3. $0 \leq f_k(n) \leq 1 = \frac{0}{n} \leq \frac{N_k(n)}{n} \leq \frac{n}{n}$
4. $\sum_{k=1}^{k} f_k(n) = \sum_{k=1}^{k} \frac{N_k(n)}{n} = \frac{\sum_{k=1}^{k} N_k(n)}{n} = \frac{n}{n} = 1$
5. $\sum_{k=1}^{k} f_k(n) = 1$
6. If events A and B are disjoint and event C is "A or B", then $F_C = F_A(n) + F_B(n)$

2 Set Theory

- A *set* is a collection of objects, denoted by capital letters
- Denote the *universal set,* U; consisting of all possible objects of interest in a given setting/application
- For any set A, we say that *"x is an element of A"*, denoted $x \in A$ if object x of the universal set U is contained in A
- We say that *"x is not an element of A"*, denoted $x \notin A$ if object x of the universal set U is not contained in A
- We say that *"A is a subset of B"*, denoted $A \subset B$ if every element in A also belongs to B, $x \in A \rightarrow x \in B$
- The *empty set,* $\emptyset$ is defined as the set with no elements
  - The empty set is a subset of every set
- Sets *A and B are equal* if they contain the same elements. To show this:
  1. Enumerate the elements of each set
  2. Thm: $A = B \iff A \subset B$ AND $B \subset A$
- The *union of 2 sets A, B*, denoted $A \cup B$ is defined as the set of outcomes that are either in A, or in B, or both
- The *intersection fo 2 sets, A, B*, denoted $A \cap B$ is defined as the set of outcomes in A and B
- The 2 sets A, B are said to be *disjoint or mutually exclusive* if $A \cap B = \emptyset$
- The *complement of a set A*, denoted A^C is defined as the set of elements of U not in A
  - $A^C = \{x \in U | x \notin A\}$
- *Relative complement* or *difference*, denoted $A - B$, is the set of elements in A that are not in B
  - $A - B = A \cap B^C$
  - $A^C = U - A$

2.1 Properties of Set Operations

Set Operators are:

1. Commutative, Equation (1)
$$\begin{aligned} A \cup B &= B \cup A \\ A \cap B &= B \cap A \end{aligned} \tag{1}$$
2. Associative, Equation (2)
$$\begin{aligned} A \cup (B \cup C) &= (A \cup B) \cup C \\ A \cap (B \cap C) &= (A \cap B) \cap C \end{aligned} \tag{2}$$
3. Distributive, Equation (3)
$$\begin{aligned} A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\ A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \end{aligned} \tag{3}$$
4. Set Operations obey De Morgan's Laws, Equation (4)
$$\begin{aligned} (A \cup B)^C &= A^C \cap B^C \\ (A \cap B)^C &= A^C \cup B^C \end{aligned} \tag{4}$$

Additionally,

Defn 1 (Union of n Sets)**.** The *union of n sets* $\bigcup_{k=1}^{n} A_k = A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n$ is the set consisting of all elements such that $x \in A_k$ for some $1 \leq k \leq n$.

- All sets need to be empty to make $\bigcup_{k=1}^{n} A_k = \emptyset$

Defn 2 (Intersection of n Sets)**.** The *intersection of n sets* $\bigcap_{k=1}^{n} A_k = A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n$ is the set consisting of all elements such that $x \in a_k$ for all $1 \leq k \leq n$

- Just one set needs to be empty to make $\bigcap_{k=1}^{n} A_k = \emptyset$

3 Probability Theory

There are 3 main components to Probability Theory.

1. Set Theory
2. Axioms of Probability
3. Conditional Probability and Independence

3.1 Random Experiments

Defn 3 (Random Experiment)**.** A *random experiment* is an experiment whose outcome varies in an unpredictable fashion when performed under the same conditions.

Defn 4 (Sample Space)**.** A *sample space, S* of a random experiment is the set of all possible experiments.

Defn 5 (Outcome/Sample Point)**.** An *outcome*, or *sample point* of a random experiment is a result that cannot be decomposed into other results.

Defn 6 (Event)**.** An *event* corresponds to a subset of the sample space. We say an event occurs if and only if (iff) the outcome of the experiment is in the subset representing the event.

Defn 7 (Event Classes)**.** An *event class* $\mathcal{F}$ is the collection of the all the events' sets. $\mathcal{F}$ should be closed under unions, intersections, and complements.

- For S finite, or countably infinite, then we can let $\mathcal{F}$ be all subsets of S.
- For S uncountably infinite, instead we can let $\mathcal{F}$ consist of the subsets that can be obtained as countable unions and intersections of some sets of $\mathcal{F}$.

Defn 8 (Probability Law)**.** A *probability law* for a random experiment E, with sample space S, and an event class $\mathcal{F}$ is a rule that assigns to each event $A \in \mathcal{F}$ a number $P[A]$, called the probability of A that satisfies the axioms:

Axiom I: $0 \leq P[A]$
Axiom II: $P[S] = 1$
Axiom III: If $A \cap B = \emptyset$, then $P[A \cup B] = P[A] + P[B]$
Axiom III': If A_1, A_2, … is a sequence of events such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $P\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} P[A_k]$

3.2 Probability Law Corollaries

Axiom I: $0 \leq P[A]$
Axiom II: $P[S] = 1$
Axiom III: If $A \cap B = \emptyset$, then $P[A \cup B] = P[A] + P[B]$
Axiom III': If A_1, A_2, … is a sequence of events such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $P\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} P[A_k]$

Corollary 3.1. $P\left[A^C\right] = 1 - P[A]$

Corollary 3.2. $P[A] \leq 1$

Corollary 3.3. $P[\emptyset] = 0$

Corollary 3.4. If A_1, A_2, …, A_n are pairwise mutually exclusive ($A_1 \cap A_2 \cap \ldots \cap A_n = \emptyset$), then $P\left[\bigcup_{k=1}^{n}\right] = \sum_{k=1}^{n} P[A_k]$ for $n \geq 2$

Corollary 3.5. $P[A \cup B] = P[A] + P[B] - P[A \cap B]$

Corollary 3.6. $P[A \cup B] = \sum_{j=1}^{n} P[A_j] - \sum_{j<k} P[A_j \cap A_k] + \ldots + (-1)^{n+1} P[A_1 \cap \ldots \cap A_n]$

Corollary 3.7. If $A \subset B$, then $P[A] \leq P[B]$

3.3 Conditional Probability

Defn 9 (Conditional Probability)**.** The *conditional probability* of event A **GIVEN THAT** event B occurred is denoted $P[A|B]$ and is defined as

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \tag{5}$$

Theorem 1 (Theorem of Total Probability)**.** *Let* $B_1, B_2, \ldots, B_n$ *be mutually exclusive events whose union equals the sample space* S*, i.e.* $B_1, B_2, \ldots, B_n$ *is a* partition of S*.*

Defn 10 (Baye's Rule)**.** Let $B_1, B_2, \ldots, B_n$ be a partition of sample space S.

$$P[B_j|A] = \frac{P[A \cap B_j]}{P[A]} = \frac{P[A|B_j] * P[B_j]}{\sum_{k=1}^{n} P[A|B_k] * P[B_k]} \tag{6}$$

4 Single Discrete Random Variables

5 Single Continuous Random Variables

6 Multiple Random Variables

6.1 Joint Probability Mass Function

Defn 11 (Joint Probability Mass Function)**.** The *joint probability mass function (joint PMF)* of 2 discrete random variables X, Y is defined as:

$$p_{X,Y} = P[\{X = x\} \cap \{Y = y\}] \text{ for all } x, y \in S_{X,Y} \tag{7}$$

- This satisfies ALL propoerties of single random variable PMFs

6.1.1 Marginal Probability Mass Function

Defn 12 (Marginal Probability Mass Function)**.** Given a joint PMF of discrete random variables X, Y, the *Marginal Probability Mass Function (Marginal PMF)* of X is defined as:

$$p_X(x_i) = P[X = x_i] \text{ for } x_i \in S_X \tag{8}$$

and is calculated as:

$$p(x_i) = \sum_{y \in S_Y} p_{X,Y}(x_i, y) \tag{9}$$

6.2 Joint Cumulative Distribution Function

Defn 13 (Joint Cumulative Distribution Function)**.** The *Joint Cumulative Distribution Function (Joint CDF)* of X and Y is defined as the probability of the event $\{X \leq x\} \cap \{Y \leq y\}$

$$\begin{aligned} F_{X,Y}(x, y) &= P[\{X \leq x\} \cap \{Y \leq y\}] \text{ for all } (x, y) \in \mathbb{R}^2 \\ &= P[\{X \leq x\}, \{Y \leq y\}] \end{aligned} \tag{10}$$

(i) $F_{X,Y}(x, y)$ is non decreasing.

$$F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2) \text{ if } x_1 \leq x_2 \text{ and } y_1 \leq y_2 \tag{11}$$

(ii)

$$\begin{aligned} \lim_{y \to -\infty} F_{X,Y}(x, y) &= 0 \\ \lim_{x \to -\infty} F_{X,Y}(x, y) &= 0 \\ \lim_{(x,y) \to (\infty,\infty)} F_{X,Y}(x, y) &= 1 \end{aligned} \tag{12}$$

(iii) The Marginal CDFs can be obtained from the Joint CDF by removing restrictions for all but one variable.

$$\begin{aligned} F_X(x) &= P[\{X \leq x\}, \{Y \text{ is anything}\}] \\ &= P[\{X \leq x\}, \{-\infty \leq y \leq \infty\}] \\ &= \lim_{y \to \infty} F_{X,Y}(x, y) \\ F_Y(y) &= \lim_{x \to \infty} F_{X,Y}(x, y) \end{aligned} \tag{13}$$

(iv) The Joint CDF is continuous from ∞ to $-\infty$.

$$\begin{aligned} \lim_{x \to a^+} F_{X,Y}(x, y) &= F_{X,Y}(a, y) \\ \lim_{y \to b^+} F_{X,Y}(x, y) &= F_{X,Y}(x, b) \end{aligned} \tag{14}$$

(v) The probability of the "rectangle" $\{x_1 \le X \le x_2, y_1 \le Y \le y_2\}$

$$\begin{aligned} P[\{x_1 \le X \le x_2, y_1 \le Y \le y_2\}] &= P[\{X \le x_2, Y \le y_2\}] - P[\{X \le x_1, Y \le y_2\}] - \\ &\quad P[\{X \le x_2, Y \le y_1\}] + P[\{X \le x_1, Y \le y_1\}] \\ &= F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1) \end{aligned} \tag{15}$$

6.2.1 Marginal Cumulative Distribution Function

Defn 14 (Marginal Cumulative Distribution Function)**.** We obtain the *Marginal Cumulative Distribution Functions (Marginal CDFs)* by removing the constraint on one of the variables.

$$\begin{aligned} F_X(x) &= P[\{X \le x\}, \{Y \text{ is anything}\}] \\ &= P[\{X \le x\}, \{-\infty \le y \le \infty\}] \\ &= \lim_{y \to \infty} F_{X,Y}(x, y) \\ F_Y(y) &= \lim_{x \to \infty} F_{X,Y}(x, y) \end{aligned} \tag{16}$$

6.3 Joint Probability Density Function