

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
- $\lim_{n \rightarrow \infty} f_k(n) = p_k \leftarrow$ **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: $A = B \iff A \subset B \text{ AND } B \subset A$
- 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 of 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 (2.1)

$$\begin{aligned} A \cup B &= B \cup A \\ A \cap B &= B \cap A \end{aligned} \tag{2.1}$$

2. Associative, Equation (2.2)

$$\begin{aligned} A \cup (B \cup C) &= (A \cup B) \cup C \\ A \cap (B \cap C) &= (A \cap B) \cap C \end{aligned} \tag{2.2}$$

3. Distributive, Equation (2.3)

$$\begin{aligned} A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\ A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \end{aligned} \tag{2.3}$$

4. Set Operations obey De Morgan's Laws, Equation (2.4)

$$\begin{aligned} (A \cup B)^C &= A^C \cap B^C \\ (A \cap B)^C &= A^C \cup B^C \end{aligned} \tag{2.4}$$

Additionally,

Defn 1 (Union of n Sets). The *union of n sets* $\bigcup_{k=1}^n A_k = A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n$ 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* $\bigcap_{k=1}^n A_k = A_1 \cap A_2 \cap A_3 \cap \dots \cap A_n$ 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, \dots is a sequence of events such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $P[\bigcup_{k=1}^{\infty} A_k] = \sum_{k=1}^{\infty} P[A_k]$

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, \dots is a sequence of events such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $P[\bigcup_{k=1}^{\infty} A_k] = \sum_{k=1}^{\infty} P[A_k]$

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

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

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

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

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

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

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

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \quad (3.1)$$

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

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

$$P[B_j|A] = \frac{P[A \cap B_j]}{P[A]} = \frac{P[A|B_j] * P[B_j]}{\sum_{k=1}^n P[A|B_k] * P[B_k]} \quad (3.2)$$

3.4 Event Independence

Defn 11 (Independent). Two events A and B are *independent* if

$$P[A \cap B] = P[A] * P[B], P[A] \neq 0, P[B] \neq 0 \quad (3.3)$$

- If $A \cap B = \emptyset$, the A and B are **dependent**.
- If checking for independence between more than 2 events, you must check each pair, each triple, etc. until you check the independence of each event against each other. For 3 events, A, B, C :
 - Check $P[A \cap B \cap C] = P[A] * P[B] * P[C]$
 - Also need to check:
 1. $P[A \cap B] = P[A] * P[B]$
 2. $P[B \cap C] = P[B] * P[C]$
 3. $P[A \cap C] = P[A] * P[C]$

4 Counting

4.1 Ordered Sampling with Replacement

Defn 12 (Permutations). The number of distinct outcomes of an experiment, where the elements being sampled are replaced between each sampling.

$$\begin{array}{c} \text{If } k = n \\ \frac{n}{\text{First}} * \frac{n-1}{\text{Second}} * \frac{n-2}{\text{Third}} * \dots * \frac{n-k-1}{\text{kth Item}} = n! \end{array} \quad (4.1)$$

4.2 Ordered Sampling without Replacement

Defn 13. Choose k elements in succession without replacement from a population of n distinct objects, where $k \leq n$

$$\frac{n}{\text{First}} * \frac{n-1}{\text{Second}} * \frac{n-2}{\text{Third}} * \dots * \frac{n-k-1}{\text{kth Item}} \quad (4.2)$$

4.3 Unordered Sampling with Replacement

4.4 Unordered Sampling without Replacement

Defn 14. The number of ways to choose k items out of n items. Said n choose k :

$$\binom{n}{k} = \frac{n * (n-1) * (n-2) * \dots * (n-k+1)}{k!} = \frac{n!}{k! (n-k)!} \quad (4.3)$$

$$\binom{n}{k} = \binom{n}{n-k} \quad (4.4)$$

5 Single Discrete Random Variables

Defn 15 (Random Variable). A *random variable* X is a function that assigns a real number $X(\zeta)$ to each outcome ζ in the sample space of the random experiment.

6 Single Continuous Random Variables

7 Multiple Random Variables

7.1 Joint Probability Mass Function

Defn 16 (Joint Probability Mass Function). The *joint probability mass function (joint PMF)* of 2 discrete random variables X, Y is defined as:

$$p_{X,Y} = P[\{X = x\} \cap \{Y = y\}] \text{ for all } x, y \in S_{X,Y} \quad (7.1)$$

- This satisfies ALL propoerties of single random variable PMFs

7.1.1 Marginal Probability Mass Function

Defn 17 (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:

$$p_X(x_i) = P[X = x_i] \text{ for } x_i \in S_X \quad (7.2)$$

and is calculated as:

$$p(x_i) = \sum_{y \in S_Y} p_{X,Y}(x_i, y) \quad (7.3)$$

7.2 Joint Cumulative Distribution Function

Defn 18 (Joint Cumulative Distribution Function). The *Joint Cumulative Distribution Function (Joint CDF)* of X and Y is defined as the probability of the event $\{X \leq x\} \cap \{Y \leq y\}$

$$\begin{aligned} F_{X,Y}(x, y) &= P[\{X \leq x\} \cap \{Y \leq y\}] \text{ for all } (x, y) \in \mathbb{R}^2 \\ &= P[\{X \leq x\}, \{Y \leq y\}] \end{aligned} \quad (7.4)$$

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

$$F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2) \text{ if } x_1 \leq x_2 \text{ and } y_1 \leq y_2 \quad (7.5)$$

- (ii)

$$\begin{aligned} \lim_{y \rightarrow -\infty} F_{X,Y}(x, y) &= 0 \\ \lim_{x \rightarrow -\infty} F_{X,Y}(x, y) &= 0 \\ \lim_{(x,y) \rightarrow (\infty, \infty)} F_{X,Y}(x, y) &= 1 \end{aligned} \quad (7.6)$$

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

$$\begin{aligned} F_X(x) &= P[\{X \leq x\}, \{Y \text{ is anything}\}] \\ &= P[\{X \leq x\}, \{-\infty \leq y \leq \infty\}] \\ &= \lim_{y \rightarrow \infty} F_{X,Y}(x, y) \\ F_Y(y) &= \lim_{x \rightarrow \infty} F_{X,Y}(x, y) \end{aligned} \quad (7.7)$$

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

$$\begin{aligned} \lim_{x \rightarrow a^+} F_{X,Y}(x, y) &= F_{X,Y}(a, y) \\ \lim_{y \rightarrow b^+} F_{X,Y}(x, y) &= F_{X,Y}(x, b) \end{aligned} \quad (7.8)$$

- (v) The probability of the “rectangle” $\{x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2\}$

$$\begin{aligned} P[\{x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2\}] &= P[\{X \leq x_2, Y \leq y_2\}] - P[\{X \leq x_1, Y \leq y_2\}] - \\ &\quad P[\{X \leq x_2, Y \leq y_1\}] + P[\{X \leq x_1, Y \leq y_1\}] \\ &= F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1) \end{aligned} \quad (7.9)$$

7.2.1 Marginal Cumulative Distribution Function

Defn 19 (Marginal Cumulative Distribution Function). We obtain the *Marginal Cumulative Distribution Functions* (Marginal CDFs) by removing the constraint on one of the variables.

$$\begin{aligned} F_X(x) &= P[\{X \leq x\}, \{Y \text{ is anything}\}] \\ &= P[\{X \leq x\}, \{-\infty \leq y \leq \infty\}] \\ &= \lim_{y \rightarrow \infty} F_{X,Y}(x, y) \\ F_Y(y) &= \lim_{x \rightarrow \infty} F_{X,Y}(x, y) \end{aligned} \tag{7.10}$$

7.3 Joint Probability Density Function

Defn 20 (Joint Probability Density Function). We say that X, Y are jointly continuous if the probabilities of events involving X and Y can be expressed as an integral of a *Joint Probability Density Function* (Joint PDF).

i.e. There exists some nonnegative function $f_{X,Y}(x, y)$, which we call the joint PDF, that is defined on the real plane such that for every event B which is a subset of the xy plane

$$P[(X, Y) \text{ in } B] = \iint_B f_{X,Y}(x, y) dx dy \tag{7.11}$$

Remark 20.1. The probability mass of an event is found by integrating the PDF over the region in the xy plane corresponding to your event.

7.3.1 Properties

$$\iint_B f_{X,Y}(x, y) = 1 \tag{7.12}$$

$$x \geq 0, y \geq 0 \forall x \forall y \tag{7.13}$$

$$\tag{7.14}$$

7.3.2 Facts about Joint PDFs

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) = 1 \tag{7.15}$$

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(s, t) dt ds \tag{7.16}$$

$$f_{X,Y} = \frac{\partial^2 f_{X,Y}(x, y)}{\partial x \partial y} \tag{7.17}$$

$$\tag{7.18}$$

7.3.3 Marginal PDF

Defn 21 (Marginal Probability Density Function). The *Marginal Probability Density Functions* (Marginal PDFs) $f_X(x)$ and $f_Y(y)$ are obtained by taking the derivative of the marginal CDFs.

$$\begin{aligned} f_X(x) &= \frac{d}{dx} F_X(x) \\ &= \frac{d}{dx} \int_{-\infty}^x \left[\int_{-\infty}^{\infty} f_{X,Y}(s, t) dt ds \right] \\ &= \frac{d}{dx} \int_{-\infty}^x \int_{-\infty}^{\infty} f_{X,Y}(s, t) dt ds \end{aligned} \tag{7.19}$$

Simplified with Second Fundamental Theorem of Calculus

$$\begin{aligned} &= \int_{-\infty}^{\infty} f_{X,Y}(x, t) dt \\ f_X &= \int_{-\infty}^{\infty} f_{X,Y}(x, t) dt \end{aligned}$$

7.4 Independence of Multiple Random Variables

Defn 22 (Independent Random Variables). X and Y are independent random variables if **ANY** event A_1 defined in terms of S is independent of **ANY** event A_2 defined in terms of Y .

$$P[X \in A_1, Y \in A_2] = P[X \in A_1] * P[Y \in A_2] \quad (7.20)$$

There are 3 ways to phrase this:

1. For discrete random variables X and Y , X and Y are independent if and only if:

$$p_{X,Y}(x, y) = p_X(x) * p_Y(y) \quad (7.21)$$

2. For discrete random variables X and Y , X and Y are independent if and only if:

$$F_{X,Y}(x, y) = F_X(x) * F_Y(y) \quad (7.22)$$

3. For discrete random variables X and Y , X and Y are independent if and only if:

$$f_{X,Y}(x, y) = f_X(x) * f_Y(y) \quad (7.23)$$

You can prove Independence of Multiple Random Variables, Equation (7.21).

Independence of Discrete Random Variables with PMF. □

8 Random Vectors

Random Vectors are usually denoted:

$$\vec{X} = \langle X_1, X_2, X_3, \dots, X_n \rangle \quad (8.1)$$

8.1 Joint CDF of a Random Vector

$$\begin{aligned} F_{\vec{X}}(\vec{x}) &= F_{X_1, X_2, X_3, \dots, X_n}(x_1, x_2, x_3, \dots, x_n) \\ &= P[X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3, \dots, X_n \leq x_n] \end{aligned} \quad (8.2)$$

8.2 Joint PDF of a Random Vector

$$f_{\vec{X}}(\vec{x}) = \frac{\partial^n F_{\vec{X}}(\vec{x})}{\partial x_1 \partial x_2 \partial x_3 \dots \partial x_n} \quad (8.3)$$

8.2.1 Marginal PDF of a Random Vector

Integrate out the terms that you're not interested in.

$$f_{\vec{X}} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\vec{X}}(\vec{x}) \partial x_2 \partial x_3 \dots \partial x_n \quad (8.4)$$

For instance, say we want the marginal PDF of some function with respect to X_1 , X_3 , and X_4 .

$$f_{X_1, X_3, X_4}(x_1, x_3, x_4) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\vec{X}}(\vec{x}) \partial x_2 \partial x_5 \partial x_6 \dots \partial x_n \quad (8.5)$$

A Reference Material

A.1 Trigonometry

A.1.1 Trigonometric Formulas

$$\sin(\alpha) + \sin(\beta) = 2 \sin\left(\frac{\alpha + \beta}{2}\right) \cos\left(\frac{\alpha - \beta}{2}\right) \quad (A.1)$$

A.2 Calculus

A.2.1 Fundamental Theorems of Calculus

Defn 23 (First Fundamental Theorem of Calculus). The *first fundamental theorem of calculus* states that, if f is continuous on the closed interval $[a, b]$ and F is the indefinite integral of f on $[a, b]$, then

$$\int_a^b f(x) dx = F(b) - F(a) \quad (\text{A.2})$$

Defn 24 (Second Fundamental Theorem of Calculus). The *second fundamental theorem of calculus* holds for f a continuous function on an open interval I and a any point in I , and states that if F is defined by

$$F(x) = \int_a^x f(t) dt,$$

then

$$\begin{aligned} \frac{d}{dx} \int_a^x f(t) dt &= f(x) \\ F'(x) &= f(x) \end{aligned} \quad (\text{A.3})$$