STAT 311: Introduction to Probability

STAT 311: Introduction to Probability
- Set Theory
- Chapter 1
- Chapter 2
- Chapter 3
- Chapter 4
- Chapter 5

Set Theory

$\mathbb{R}$: Collection of real numbers

$\mathbb{Q}$: Collection of rational numbers, that is consist of p/q where p and q are integers

$\mathbb{Z}$: Collection of integers: ..., -2, -1, 0, 1, 2, ...

$\mathbb{N}$: Collection of positive integer (does not include 0)

- **Definition 1.2**

  Let U be a set and let A, B and E be the subsets of U

  a) The **complement** of E, $E^c = \{x : x \notin E\}$

  b) The **intersection** of A and B is $A \cap B = \{x : x \in A \text{ and } x \in B\}$

  c) The **union** of A and B is $A \cup B = \{x : x \in A \text{ or } x \in B\}$

- **Proposition 1.1: De Morgan's Laws**

  Let A and B be subsets of U, Then

  a) $(A \cup B)^c = A^c \cap B^c$ b) $(A \cap B)^c = A^c \cup B^c$

- **Proposition 1.2: Distributive Laws**

  Let A, B and C be subsets of U, Then

  a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

- **Proposition 1.3: Associative and Commutative Laws**

  Let A, B, and C be subsets of U, Then

  a) $A \cap B = B \cap A$ b) $A \cup B = B \cup A$

  c) $A \cap (B \cap C) = (A \cap B) \cap C$ d) $A \cup (B \cup C) = (A \cup B) \cup C$

- **Disjoint sets**

Definition 1.4: Two sets, A and B are said to be **disjoint** if $A \cap B = \emptyset$, that is, they have no elements in common. Sets $A_1, A_2, ..$ are said to be **pairwise disjoint** if $A_n \cap A_m = \emptyset$ when $m \neq n$

- **Cartesian Product**

  Definition 1.5: Let A and B be two sets. The **Cartesian product** of A and B, denoted by $A \times B$ is the set of all ordered pairs (a,b), where $a \in A$ and $b \in B$.

  Example:

  Let $A = \{a, b\}$ and $B = \{1, 2\}$, then $A \times B = \{(a, 1), (a, 2), (b, 1), (b, 2)\}$

- **Venn Diagram**

Chapter 1

Event: An event is a collection or a set of outcomes.

Sample space: The sample space is the set of all outcomes. Denote the sample space using the letter S.

Empty set: The empty set, denoted by $\emptyset$, is the event that contains no outcomes whatsoever.

If all of the outcomes in the event A are also contained in event B, we say that A is a subset of B, denoted by $A \subset B$

If event A and its complement A^c are disjoint, and their union is all of the sample space S. We could write $A \cup A^c = S$

Number of events: $2^n \quad (n = number\ of\ outcomes)$

- **Theorem 1.22 DeMorgan's first law**

  For a finite or infinite collection of events, $A_1, A_2, ..$

$$\left(\bigcup_j A_j\right)^c = \bigcap_j A_j^c$$

- **Theorem 1.23 DeMorgan's second law**

  For a finite or infinite collection of events, $A_1, A_2, ..$

$$\left(\bigcap_j A_j\right)^c = \bigcup_j A_j^c$$

- **Distributive Laws**

  Let A, B and C be subsets of S, Then

  a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

  b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

- **Associative and Commutative Laws**

  Let A, B, and C be subsets of U, Then

  a) $A \cap B = B \cap A$

  b) $A \cup B = B \cup A$

  c) $A \cap (B \cap C) = (A \cap B) \cap C$

d) $A \cup (B \cup C) = (A \cup B) \cup C$

Chapter 2

- **Proof!**

- **Frequentist point of view**

  The probability of an event is the long-run proportion of times that the event occurs in independent repetitions of the random experiment.

$$\lim_{n \to \infty} \frac{N(E)}{n} \approx P(E)$$

- **Bayesian**: Conditional probability (See chapter 5)

- **Definition 2.2.** A pair of events A, B is **disjoint** (also called **mutually exclusive**) if they have no outcome in common, i.e., if their intersection is empty, $A \cap B = \emptyset$. A collection of events is disjoint if every pair of events is disjoint.

- **Axioms 2.4. Three Probability Axioms**

  1. **Non-negativity**: For each event A, $0 \leq P(A) \leq 1$
  2. **Certainty**: For the sample space, S, $P(S) = 1$
  3. **Additivity**: If $A_1, \ldots$ is a collection of disjoint events, then

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$$

- **Type of Probabilities**

  1. **Equally Likely Events**: If we have a situation (a "random process") in which there are n equally likely outcomes, and the event A consists of exactly m of these outcomes, we say that the probability of A is m/n, which is $P(A) = \frac{m}{n}$. (**finitely many equally likely outcomes**)
  2. **Empirical**: This perspective defines probability via a thought **experiment**.
  3. **Subjective**: Subjective probability is an **individual person's measure of belief** that an event will occur.

- **Theorem 2.5.** The probability of the empty set $\emptyset$ is always 0.

- **Theorem 2.6.** If $A_1, A_2, .., A_n$ is a collection of finitely many *disjoint* events, then the probability of the union of the events equals the sum of the probabilities of the events:

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j)$$

- **Definition 2.14.** If a collection of nonempty events is disjoint, and their union is the entire sample space, then the collection is called a **partition**. If $\bigcup_j B_j = S$ and the B_j's are disjoint events, then the collection of B_j's is called a **partition**.

- **Remark 2.16.** The probabilities of events in a partition always sum to 1.

- **Theorem 2.19. Complementation rule** The complement A^c of event A has probability $P(A^c) = 1 - P(A)$.

- **Theorem 2.20. Domination principle** If $A \subset B$ then $P(A) \leq P(B)$.

- **Theorem 2.22.** For any events A and B,

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

- **Theorem 2.23.** For any three events A, B, C,

$$\begin{aligned} P(A \cup B \cup C) = {} & P(A) + P(B) + P(C) \\ & - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ & + P(A \cap B \cap C) \end{aligned}$$

- **Theorem 2.25. Inclusion-Exclusion Rule** For any finite sequence of events $A_1, A_2, \dots, A_n$,

$$\begin{aligned} P\left(\bigcup_{j=1}^{n} A_j\right) = {} & \sum_{i=1}^{n} P(A_j) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \\ & - \sum_{i<j<k<l} P(A_i \cap A_j \cap A_k \cap A_l) \\ & \pm \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n) \end{aligned}$$

Chapter 3

- **Definition 3.1. Independence**

  Event A and B are called **independent** if

$$P(A \cap B) = P(A)P(B)$$
$$P(A|B) = P(A)P(A|B) = P(A)P(B)$$

- **Definition 3.2. Dependence**

  Events A and B are called dependent if they are not independent. In other words, A and B are dependent if

$$P(A \cap B) \neq P(A)P(B)$$

- **Remark 3.6. Independent vs Disjoint**

  If A and B both have positive probabilities, they cannot be both independent and disjoint.

  When $P(A) = 0$, then A is independent from any event B, because, $P(A \cap B) = 0 = P(A)P(B)$

- **Theorem 3.9. Subsets are dependent.** If $A \subset B$ and neither $P(A) = 0$ nor $P(B) = 1$, then A, B are dependent.
- **Theorem3.10. Complements are dependent.** If neither $P(A) = 0$ nor $P()A = 1$, then A, A^c are dependent.
- **Definition 3.12. Independent of three events**

  A collection of three events A, B, C is called (*mutually*) *independent* if all four of the following are satisfied:

$$P(A \cap B) = P(A)P(B)$$
$$P(A \cap C) = P(A)P(C)$$
$$P(B \cap C) = P(B)P(C)$$
$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

- **Theorem 3.20. Independent among complements**

  When events are independent, their complements are too, i.e., if A, B are independent, then A^c, B are independent, A, B^c are independent, and A^c, B^c are independent too.

- **Theorem 3.24. Probability of Good Occurring before Bad**

$$P(Good\ before\ bad) = \frac{p}{p+q}$$

Chapter 4

- **Definition 4.2.** The conditional probability of event $\boldsymbol{A}$, given event $\boldsymbol{B}$ is written as $P(A \mid B)$.

  In other words, conditional probability is written as

  $$P(\textit{event under consideration} \mid \textit{a given event})$$

  In general, if event $\boldsymbol{B}$ has nonzero probability (i.e., $P(B) > 0$), then the conditional probability $P(A \mid B)$ of $\boldsymbol{A}$ given $\boldsymbol{B}$ is defined as

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

- **Theorem 4.4.** If $P(B) > 0$, then A and B are independent if and only if $P(A) = P(A \mid B)$, i.e., when $B's$ occurrence does not affect the probability of $\boldsymbol{A}$ occurring

- **Theorem 4.10. Distributive Laws** For any event $A_1, A_2, \ldots,$

  $$\left(\bigcup_j A_j\right) \cap B = \bigcup_j (A_j \cap B)$$

  and

  $$\left(\bigcap_j A_j\right) \cup B = \bigcap_j (A_j \cup B)$$

  The unions and intersections over j can be finite, written as $\cup_{j=1}^{n} A_j$, or can be finite, $\cup_{j=1}^{\infty} A_j$.

- Conditional Probability Satisfy the Probability Axioms

  Consider an event B with $P(B) > 0$,

  1. For any event A

  $$0 \leq P(A \mid B) \leq 1;$$

  2. For the sample space S,

  $$P(S \mid B) = 1;$$

  3. For any disjoint events $A_1, A_2, \ldots,$

  $$P\left(\bigcup_{j=1}^{\infty} A_j \,\middle|\, B\right) = \sum_{j=1}^{\infty} P(A \mid B)$$

- Calculate intersections using the general multiplication rule

Chapter 5

- **Theorem 5.1. Bayes' Theorem**

  For any two events A and B with nonzero probabilities,

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

$$\big(P(B)P(A \mid B) = P(A)P(B \mid A)\big)$$

- **Theorem 5.5. Bayes' Theorem (Decomposition of Sample Space into Two Parts)**

  If $P(B) \neq 0$ and $0 < P(A) < 1$, then,

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + P(A^c)P(B \mid A^c)}$$

- **Remark 5.10.** Let $A_1, A_2, \ldots$ be a partition of the sample space (i.e., $\cup_{j=1}^{\infty} A_j = S$ and the A_j's are disjoint). Then we have the decomposition of B into infinitely many parts, $A_1 \cap B, A_2 \cap B$, etc. So $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots$ is a union of disjoint events. Thus

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \cdots = \sum_{j=1}^{\infty} P(A_j \cap B) = \sum_{j=1}^{\infty} P(A_j)P(B \mid A_j)$$

- **Theorem 5.5. Bayes' Theorem (Decomposition of Sample Space into Infinitely Many Parts)**

  If $P(B) > 0$, and if $A_1, A_2, \ldots$ form a partition of S, with all $P(A_j) > 0$, then

$$P(A_k \mid B) = \frac{P(A_k)P(B \mid A_k)}{\sum_{j=1}^{\infty} P(A_j)P(B \mid A_j)}$$

- **Remark 5.16. Multiplication with Conditional Probabilities**

  For any events $A_1, A_2, \ldots, A_n$ with $A_1 \cap A_2 \cap \cdots \cap A_n \neq \emptyset$,

$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_n) &= P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \\ &\quad \times P(A_4 \mid A_1 \cap A_2 \cap A_3) \\ &\quad \cdots \times P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1})) \end{aligned}$$

- **Pólya's Urn**

  In the basic Pólya urn model, the urn contains x white and y black balls; one ball is drawn randomly from the urn and its color observed; it is then returned in the urn, and an additional c balls of **the same color** is added to the urn. Then the selection process is repeated.

- **Electrical Circuits and Probability**

  In a **series** circuit, the current will only flow if all of the lights work. That is, if any of the light bulbs are burned out, none of them will turn in. Christmas tree lights used to be like this. This is equivalent of a '**and**'.

  In a **parallel** circuit , the current will continue to flow as long as any of the lights work. Christmas tree lights now are in parallel circuits so if one goes bad, the rest of the string will still light up. This is equivalent of an '**or**'.