ECE286
Probability and Statistics

Michael Boyadjian

February 2, 2020

Contents

1 Introduction to Statistics and Data Analysis

1.1 Definitions

- **Probability:** Mathematical theory that models and analyses uncertainty based on 3 fundamental axioms.
- **Statistics:** Empirical; based on data and measurements; we infer characteristics of a phenomenon

1.2 Set Theory

A set is a set of objects (or elements) represented using capital letters A, B, C, S

- **Universal Set:** Set of all elements or objects of interest
- **Empty Set:** Set without any elements, denoted $\varnothing$
- **Compliment of a Set:** $A' = \{x|x \notin A\}$
- **Union of Two Sets:** $A \cup B = \{x|x \in A \text{ or } x \in B\}$
- **Intersection of Two Sets:** $A \cap B = \{x|x \in A \text{ and } x \in B\}$
- **Disjoint Sets:** Also called mutually exclusive; A and B are disjoint iff $A \cap B = \varnothing$
- **Subsets:** Denoted $A \subset B$; A is a subset of B iff $x \in A \implies x \in B$

1.3 Counting

There are three different methods of counting

1. **With Replacement, With Ordering**
    - Given n distinct objects pick one, note down its kind, and put it back
    - Repeat k times
    - Produces n^k possibilities
    - *ex.* Number of 8-character passwords $= 62^8$
2. **Without Replacement, With Ordering**
    - Same as previous, but no replacement
    - Produces $\frac{n!}{(n-k)!}$ possibilities
    - Describes number of **permutations**, denoted ${}_nP_k$
    - *ex.* Number of 8-character passwords without repeating characters $= \frac{62!}{54!}$

3. **Without Replacement, Without Ordering**

   - Ordering does not matter, two sequences with same objects count as the same
   - Produces $\frac{n!}{k!(n-k)!}$ possibilities
   - Describes number of **combinations**, denoted ${}_nC_k = \binom{n}{k}$
   - *ex.* Number of ways to pick 5 good chips and 3 bad chips out of 90 good and 10 defective $= \binom{90}{5}\binom{10}{3}$

2 Probability

Probability theory rests on the notion of a random experiment. We don't know what the outcome will be until we actually run the experiment. There are three kinds of experiments:

- Designed
- Observational
- Retrospective

An experiment has a procedure and a measurement and always produces one outcome

2.1 Sample Spaces

The sample space describes the set of all possible outcomes (S). There are three different kinds of sample spaces

- Finite $\rightarrow$ discrete
- Countably Infinite $\rightarrow$ discrete
- Continuous

2.2 Events

An event is the set of outcomes we are interested in. This must be a subset of the sample space. The **event class** is the set of all subsets of the sample space. Probabilities will be assigned to the events in ε

2.3 Relative Frequency

We want to assign probabilities consistent with the idea of relative frequency, which is defined as

$$\lim_{n \to \infty} \frac{n_A}{n} = p_A$$

This is only useful to model the likelihood of events occurring, so probabilities are assigned using probability axioms

2.4 Axioms of Probability

1. For any event A, $P(A) \geq 0$
2. $P(S) = 1$
3. For any two events A and B such that $P(A \cap B) = \varnothing$, then $P(A \cup B) = P(A) + P(B)$

2.5 Properties

1. $P(A') = 1 - P(A)$
2. $P(A) \leq 1$
3. $P(\varnothing) = 0$
4. For any events A and B: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2.6 Conditional Probability

Conditional probability describes the probability of some event B occurring given an event A has occurred. This is denoted as $P(B|A)$, which is read as *"the probability of B given A"*.

A conditional probability relative to a subspace A of S may also be calculated directly from the probabilities assigned to the elements of the original sample space S

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

2.6.1 Independent Events

Two events are said to be independent if and only if

$$P(B|A) = P(B) \qquad \text{or} \qquad P(A|B) = P(A)$$

A and B are otherwise dependent

2.6.2 Product Rule and Multiplicative Rule

If in an experiment, both events A and B can occur, then

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

We can use this to find another expression for the independence of events. Two events A and B are independent if and only if

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

This multiplicative rule can be extended to more than just two-event situations. If, in an experiment, the events A_1, A_2, ... A_k occur, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_k|A_1 \cap A_2 \cap \cdots A_{k-1})$$

If these events are independent, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2)P(A_3) \cdots P(A_k)$$

2.7 Bayes' Rule

Bayes' Rule tells us the following. If the events B_1 , B_2 , . . . , B_k constitute a partition of the sample space S such that $P(B_i) \neq 0$ for$i$ = 1,2,...,k, then for any event A in S such that $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}$$

3 Random Variables and Probability Distributions

3.1 Concept of a Random Variable

A random variable is a function that associates a real number with each element in the sample space. This allows us to do math on a random experiment and map outcomes to the real line. We can classify a sample space of random variables in two ways:

- **Discrete Sample Space:** Contains a finite number of possibilities or an unending sequence with as many elements as there are whole numbers
- **Continuous Sample Space:** Contains an infinite number of possibilities equal to the number of points on a line segment

3.2 Discrete Probability Distributions

A discrete random variable assumes each of its values with a certain probability. The set of ordered pairs $(x, f(x))$ is a **probability function**, **probability mass function**, or **probability distribution** of the discrete random variable X if, for each possible outcome x

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$
3. $P(X = x) = f(x)$

In cases where we want to determine the probability that the observed value X is less than or equal to some real x, we can use what is called the **cumulative distribution function**. $F(x)$ of a discrete random variable X with probability distribution $f(x)$ is

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \qquad \text{for } -\infty < x < \infty$$

3.3 Continuous Probability Distributions

A continuous random variable has a probability of 0 of assuming exactly any of its values, so we are unable to give its probability distribution in tabular form. In dealing with continuous distributions, note the following:

$$P(a < x \leq b) = P(a < x < b) + P(x = b) = P(a < x < b)$$

This tells us that we do not need to include an endpoint of the interval if X is continuous. We can define a probability density function for a continuous random variable X. The function $f(x)$ is a **probability density function** for the continuous random variable X, defined over the set of real numbers, if

1. $f(x) \geq 0$, for all $x \in R$

2. $\int\limits_{-\infty}^{\infty} f(x)dx = 1$

3. $P(a < x < b) = \int\limits_{a}^{b} f(x)dx$

Similar to discrete distributions, we can also define a probability distribution function to determine the probability that the observed value X is less than or equal to some real x. The **cumulative distribution function** $F(x)$ of a continuous random variable X with density $f(x)$ is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \qquad \text{for } -\infty < x < \infty$$