MATH 460: Probability & Statistics — January_2019

Chapter 1: Basic Probability

Class 1 Notes — **Dr. Basilio**

Mon Jan_7 ∪ Tues Jan_8

* * *

Guiding Question(s)

(1) What is the mathematical language of collections of things?

(2) What does it mean for events to be "certain" or "uncertain"?

(3) How can we systematically study uncertainty?

(4) How can we apply this to our everyday lives?

Sets

To begin our study of certainty, we first need to discuss the basic language it is written in–mathematical set theory.

Definition 1: Set Theory

A **set** is a collection of things, or **elements**. Elements of a set can be anything, such as numbers or people. We denote a set with curly brackets { and }. So if we want to talk about the set of even numbers from 1 to 10, we can write it like this: $\{2, 10, 8, 6, 4\}$–note that we separate the elements by a comma and that the order doesn't matter.
Here are some important terms and notations:

- $\emptyset$ denotes the empty set or the set that contains no elements. We also write this $\{\}$.
- Capital letters denote sets: A, B, C etc. We also use Greek letters like "Omega" Ω.
- S denotes the universe of all possible elements in consideration. We call it the **sample space**.
- We denote $A \subset B$ to say that A **is a subset of** B. This means that every element of A is also an element of set B.
- A' (or A^c) denotes the elements of the universe S that are NOT in A. We call this the **complement of A**.
- $A \cup B$ is the set consisting of all elements in the set A combined with all the elements in set B. We call this A **union** B.
- $A \cap B$ is the set consisting of all elements that are in both A and B. We call this A **intersect** B.
- We write $A - B$ to mean the set containing elements that are in A and not in B. Notice that $A' = S - A$.
- We say that two sets are **disjoint** if they have no elements in common. In other words, A and B are disjoint if and only if $A \cap B = \emptyset$.

Activity 1: Venn Diagrams

What is a Venn Diagram? If you know what it is, use them to illustrate the all of the definitions from Definition 1.

Activity 2: Set Theory

Consider the sets $A = \{1, 2, 3, 4, 5\}$ and $B = \{2, 4, 6, 8, 10\}$ where $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Compute each of the following sets:

(a) $A \cup B$

(b) $A \cap B$

(c) A'

(d) B'

(e) $B - A$

(f) $A - B$

(g) $(A \cup B)'$

(h) $A' \cup B'$

(i) $(B - A') \cap (A \cap B)'$

Activity 3: Set Theory

Consider the sets $A = \{4, 5, 6, 7, 9, 13, 16\}$ and $B = \{3, 6, 9, 12, 15\}$ where the sample space S consists of all positive integers less than or equal to 16. Find the following:

(a) $A \cup B$

(b) $A \cap B$

(c) A'

(d) $(A \cap B)'$

Random Experiments

An **experiment** is any activity or situation in which two or more outcomes will result. A **random experiment** is an experiement where there is uncertainty in which outcome will occur.

Example 1: Random Experiments

(a) If we toss a die, the result of the experiment is that it will come up with one of the numbers in the set $\{1, 2, 3, 4, 5, 6\}$. Is it random?

(b) If we toss a coin twice, there are four results possible, as indicated by $\{HH, HT, TH, TT\}$, i.e., both heads, heads on first and tails on second, etc. Is it random?

Sample Space

A set S that consists of all possible **outcomes** of a random experiment is called a **sample space**, and each outcome is called a **sample point**. Often there will be more than one sample space that can describe outcomes of an experiment, but there is usually only one that will provide the most information.

Example 2: Sample Space

(a) If we toss a die, one sample space, or set of all possible outcomes, is given by $S = \{1, 2, 3, 4, 5, 6\}$ while another is $S = \{odd, even\}$. It is clear, however, that the latter would not be adequate to determine, for example, whether an outcome is divisible by 3.

(b) If we flip a coin the sample space is $S = \{H, T\}$ or $S = \{1, 0\}$ where 1 represents a H and 0 represents T.

(c) If we draw a card the sample space has 52 items $\{AH, 10K, 4D, \ldots\}$. Notice we use dots to denote more elements are there but are not completely listed. This can be either for finite or infinite sets.

(d) Buying a car: One possible sample space is $S = \{$(male, hybrid), (female, hybrid), (male, traditional), (female, traditional)$\}$

Activity 4: Sample Space/Outcomes

List all possible outcomes for the following.

(a) Flipping a coin 4 times.

(b) Rolling 2 six-sided dice at the same time.

Events

An **event** is a subset A of the sample space S, i.e., it is a set of possible outcomes. If the outcome of an experiment is an element of A, we say that the event A has occurred. An event consisting of a single point of S is often called a **simple** or **elementary event**.

Example 3: Events

(a) If we toss a coin, the result of the experiment is that it will either come up "tails," symbolized by T (or 0), or "heads," symbolized by H (or 1), i.e., one of the elements of the set $\{H, T\}$ (or $\{0, 1\}$).

(b) If we toss a die, and the result of the experiment is that a 5 comes up, then 5 is a simple event.

Definition 2: Events and Sets

We can use the set operations on events in S to obtain other events in S. The empty-set, $\emptyset$, is called an impossible event because it cannot occur. If A and B are events in S, then

(a) $A \cup B$ is the event that "either A or B or both" occur.

(b) $A \cap B$ is the event that "both A and B" occur.

(c) A' is the event "not in A"

(d) $A - B$ is the even "in A but not in B"

(e) If the sets corresponding to events A and B are disjoint, i.e., $A \cap B = \emptyset$, we often say that the events are **mutually exclusive**. This means that they cannot both occur.

Activity 5: Events

Let S be the sample space of flipping a coin twice. Let A be the event "at least one head occurs" and B be the event "the second toss results in a tail." Express S, A and B using the H and T notation and find:

(a) $A \cup B$

(b) $A \cap B$

(c) A'

(d) $A - B$

Definition 3: Concept of Probability

In any random experiment there is always uncertainty as to whether a particular event will or will not occur. As a measure of the **chance**, or **probability**, with which we can expect the event to occur, it is convenient to assign a number between 0 and 1. If we are sure or certain that the event will occur, we say that its probability is 100%, or 1, but if we are sure that the event will not occur, we say that its probability is zero. If, for example, the probability is $\frac{1}{4}$, we would say that there is a 25% chance it will occur and a 75% chance that it will not occur.
There are two important procedures by means of which we can estimate the probability of an event.

- **CLASSICAL APPROACH:** If an event can occur in h different ways out of a total number of n possible ways, all of which are equally likely, then the probability of the event is $\frac{h}{n}$. That is,

$$\boxed{P(E) = \frac{\#\text{ outcomes in } E}{\#\text{ outcomes in } S}}$$

- **FREQUENCY APPROACH:** If after n repetitions of an experiment, where n is very large, an event is observed to occur in h of these, then the probability of the event is $\frac{h}{n}$. This is also called the **empirical probability** of the event. That is,

$$\boxed{P(E) = \frac{\#\text{ observations}}{\#\text{ of total repetitions of experiment}}}$$

Odds are similar to, but different than probability and they are defined to be the ratio of wins to losses. Odds are customarily expressed using a colon ":" or "to". For example, 2 to 1 odds means that there are 2 ways to win and one way to lose, so the sample space has 3 total outcomes. The probability would be $2/3$ for comparison.

$$\boxed{\text{Odds: } W : W' \text{ or } W \text{ to } W'}$$

Example 4: Probability

(a) CLASSICAL APPROACH:
Suppose we want to know the probability that a head will turn up in a single toss of a coin. Since there are two equally likely ways in which the coin can come up, namely, heads and tails (assuming it does not roll away or stand on its edge) and of these two ways a head can arise in only one way, we reason that the required probability is $\frac{1}{2}$. In arriving at this, we assume that the coin is fair, i.e., not loaded in any way.

(b) FREQUENCY APPROACH:
If we toss a coin 1000 times and find that it comes up heads 532 times, we estimate the probability of a head coming up to be $\frac{532}{1000}$, or 0.532.

Activity 6: Classical Probability

A ball is drawn at random from a box containing 6 red balls, 4 white balls, and 5 blue balls. Determine the probability that it is (a) red, (b) white, (c) blue, (d) not red, (e) red or white.

Activity 7: Frequency Approach

Pair up into a group of 2 (or 3) students. One student will flip a coin and the other will record the results.

(a) Flip a coin ten times. What is the empirical probability from your experiment?

(b) Flip a coin 20 times. What is the empirical probability from your experiment?

Axioms of Probability

Both the classical and frequency approaches have serious drawbacks, the first because the words "equally likely" are vague and the second because the "large number" involved is vague. Also, it would be very impractical to have to conduct long experiments to determine probabilities. Because of these difficulties, mathematicians have been led to an axiomatic approach to probability.

Definition 4: Concept of Probability

Suppose we have a sample space S.
To each event A, we associate a real number $P(A)$. Then P is called a **probability function**, and $P(A)$ the **probability of the event** A, if the following axioms are satisfied:

- **Axiom 1** For every event A,
$$P(A) \geq 0 \tag{1}$$
- **Axiom 2** For the sample space S,
$$P(S) = 1 \tag{2}$$
- **Axiom 3** If two events A and B are mutually exclusive,
$$P(A \cup B) = P(A) + P(B) \tag{3}$$

Some observations:

- For any event A, the probability is between 0 and 1: $0 \leq P(A) \leq 1$. Of course, $P(A) = 0$ are for impossible events and $P(A) = 1$ are for events that are certain to happen.
- Because the sample space S consists of all possible outcomes for a chance experiment, in the long run an outcome that is in S must occur 100% of the time. This means that $P(S) = 1$.
- Axiom 3 can be applied to any number of mutually exclusive events. If $A_1, A_2, A_3, \ldots$ are all mutually exclusive events (that is $A_i \cap A_j = \emptyset$ for any two indices i and j) then
$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots \tag{4}$$
For example, if we roll a fair dice and A_1 is the event of rolling a 1, A_2 is the event of rolling a 2, etc then there are 6 mutually exclusive events and the sample space is $S = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_6$. Then
$$1 = P(S) = P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_6) = P(A_1) + P(A_2) + P(A_3) + \cdots + P(A_6). \tag{5}$$

Theorem 1: Theorems of Probability

For any event A,

(a) $P(A) + P(A') = 1$
This one makes sense since in words it says: any event either occurs or doesn't occur! Since A and A' are mutually exclusive (or disjoint), $P(A \cup A') = P(A) + P(A')$ by Axiom 3. Since $S = A \cup A'$ and $P(S) = 1$ (Axiom 2), $1 = P(S) = P(A \cup A') = P(A) + P(A')$.

(b) $P(A') = 1 - P(A)$
This follows from Property (a) by subtracting $P(A)$ from both sides and noting that $A' = S - A$.

(c) $P(\emptyset) = 0$.
Observe that $1 = P(S) = P(S \cup \emptyset) = P(S) + P(\emptyset) = 1 + P(\emptyset)$. Subtracting 1 from both sides gives the result.

(d) If A and B are two events (not necessarily disjoint), then
$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)} \tag{6}$$

Activity 8: Probability

Suppose A and B are two disjoint events in a sample space S and that $P(A) = .22$; $P(B) = .33$. Calculate the following probabilities.

(a) $P(A \cup B)$

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

(c) $P(A' \cup B)$

(d) $P(A' \cap B)$

(e) $P(A - B)$

Activity 9: Probability

Determine the probability for the following events.

(a) Roll a 7 or 11 from a pair of fair 6 sided dice.

(b) A non-defective television found next if out of 2,300 televisions already examined, 15 were defective.

(c) At least 1 tails appears in 5 tosses of a fair coin.

(d) The probability of drawing an ace or a club in a standard deck.

Conditional Probability

Definition 5: Conditional Probability

Let A and B be two events such that $P(A) > 0$. Denote by $P(B|A)$ the **probability of B given that A has occurred**. Since A is known to have occurred, it becomes the new sample space replacing the original S. From this we are led to the definition

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{7}$$

or, equivalently,

$$P(B \cap A) = P(A) \cdot P(B|A) \tag{8}$$

In words, (11) says that the probability that both A and B occur is equal to the probability that A occurs times the probability that B occurs given that A has occurred.

Example 5: Conditional Probability

Two cards are drawn from a well-shuffled ordinary deck of 52 cards. Find the probability that they are both aces if the first card is (a) replaced, (b) not replaced.

Let A be the event “ace on first draw” and B be the even “ace on second draw.” Then we are looking for $P(A \cap B)$.
(a) Since there are 4 aces in 52 cards, $P(A) = 4/52$. If the card is replaced (and the deck is re-shuffled) then $P(B|A) = P(B) = 4/52$ as well. Thus, using the formula

$$P(A \cap B) = P(A)P(B|A) = P(A)P(B) = \frac{4}{52} \cdot \frac{4}{52} = \frac{1}{169}.$$

(b) When the ace is not replaced, then there are 3 aces left and $P(B|A) = 3/51$ and

$$P(A \cap B) = P(A)P(B|A) = P(A)P(B) = \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}.$$

Activity 10: Conditional Probability

(a) Five marbles are picked at random out of a jar containing 10 red marbles, 15 white marbles, 20 blue marbles, 25 orange marbles, and 30 purple marbles. What is the probability of picking one of each color, assuming you pick a marble one at a time?

(b) Drawing a king and a queen for first and second draws from a well-shuffled 52 card deck.