Probability and Statistics Notes

Chapter One

Jesse Crawford

Department of Mathematics
Tarleton State University

Outline

A Probability Problem

Example

In a large city, 70% of the population approves of the mayor. In a random sample of size 5, what is the probability that 3 people will approve of the mayor?

Big Picture of Probability:

- Have a statistical model with *known* parameters.
- Use the model to calculate probabilities of outcomes.
- $$\text{Information about model} \rightarrow \text{Information about outcomes}$$

A Statistics Problem

Example

A random sample of size 5 is taken from a city with a large population, and 4 people in the sample approve of the mayor. Estimate the percentage of people in the city who approve of the mayor.

Big Picture of Statistics:

- Have a statistical model with some *unknown* parameters.
- Make observations.
- Use observations to *estimate* unknown parameters.
- $$\text{Information about model} \leftarrow \text{Information about outcomes}$$

Outline

Sample Space

Definition

The set of all possible outcomes for a random experiment is called the *sample space* S. The terms *outcome space* or *observation space* are also used.

Example

- A six sided die is rolled.
- Find the sample space.
- $S = \{1, 2, 3, 4, 5, 6\}$.

Example

- A coin is flipped twice, and the sequence of heads and tails is recorded.
- Find the sample space.
- $S = \{HH, HT, TH, TT\}$.

Example

- A coin is tossed until heads is observed, and the number of flips is recorded.
- Find the sample space.
- $S = \{1, 2, 3, \ldots\}$

Example

- Suppose everyone who works for a certain company weighs between 100 and 200 pounds, inclusive. An employee of this company is selected at random and his or her weight is noted.
- Find the sample space.
- $S = \{x \mid 100 \leq x \leq 200\}$

Frequency, Relative Frequency, and Histograms

Definition

Suppose an experiment is repeated n times. The number of times an outcome occurs is called its *frequency* f. The proportion of times it occurs, $\frac{f}{n}$, is called its *relative frequency*.

Example

A die is rolled 8 times with the following results:

$$5, 4, 5, 6, 2, 3, 2, 6$$

Find the frequency and relative frequency of each outcome in the sample space.

Probability

Definition (Informal)

- Suppose x is an outcome in the sample space.
- $$\text{Probability of } x = \lim_{n\to\infty} \text{Relative Frequency of } x.$$

Example

In 1000 die rolls, the following frequencies were observed:

Outcome	Frequency
1	162
2	165
3	180
4	178
5	159
6	156

Outline

Sets and Elements

Definition

- A *set* A is a collection of mathematical objects, called the *elements* of A.
- $x \in A$ means that x is an element of A.
- $x \notin A$ means that x is *not* an element of A.

Example

Suppose $A = \{1, 2, 3\}$.

- $1 \in A$
- $2 \in A$
- $3 \in A$
- $4 \notin A$
- $2.75 \notin A$

Example

Suppose $B = \{5, 10, 15, 20, \ldots\}$

- $40 \in B$
- $48 \notin B$

Example

Suppose $C = \{x \mid 1 < x \leq 4\}$

- $2 \in C$
- $4 \in C$
- $0 \notin C$
- $1 \notin C$

Union

Definition

The *union* of A and B is the set of elements contained in A, or B, or both.

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

Example

Suppose $A = \{1, 3, 4, 7\}$ and $B = \{1, 4, 8, 9\}$. Find $A \cup B$.
$A \cup B = \{1, 3, 4, 7, 8, 9\}$.

Example

Define intervals $A = [0, 5)$ and $B = (2, 7)$. Find $A \cup B$.
$A \cup B = [0, 7)$.

Intersection

Definition

The *intersection* of A and B is the set of elements contained in both A and B.

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

Example

Suppose $A = \{1, 3, 4, 7\}$ and $B = \{1, 4, 8, 9\}$. Find $A \cap B$.
$A \cap B = \{1, 4\}$.

Example

Define intervals $A = [0, 5)$ and $B = (2, 7)$. Find $A \cap B$.
$A \cap B = (2, 5)$.

Complements

Definition

The *universal set* U is the set of all elements under consideration. In probability theory, the universal set is the sample space S.

Definition

The complement of A is the set of elements in U that are not in A.
$A' = \{x \mid x \in U \text{ and } x \notin A\}$

Example

Suppose $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $A = \{1, 3, 4, 7\}$, and $B = \{1, 4, 8, 9\}$. Find A' and B'.

- $A' = \{2, 5, 6, 8, 9, 10\}$
- $B' = \{2, 3, 5, 6, 7, 10\}$

Example

Suppose $U = \mathbb{R}$, $A = [0, 5)$ and $B = (2, 7)$. Find A' and B'.

- $A' = (-\infty, 0) \cup [5, \infty)$
- $B' = (-\infty, 2] \cup [7, \infty)$

Subsets

Definition

$A \subseteq B$ means that every element of A is an element of B.

Example

Suppose $A = \{2, 3, 7\}$ and $B = \{1, 2, 3, 6, 7, 9\}$. Is either set a subset of the other?
$A \subseteq B$.

Example

Suppose $A = [0, 5]$ and $B = (1, 3)$. Is either set a subset of the other?
$B \subseteq A$.

Example

What about for $A = \{1, 3, 4, 7\}$, and $B = \{1, 4, 8, 9\}$?
Neither set is a subset of the other.

Example

What about for $A = [0, 5)$ and $B = (2, 7)$?
Neither set is contained in the other.

Example

What about the universal set U and any set A?
$A \subseteq U$

The Empty Set

Definition

The empty set $\emptyset$ is the set with no elements.
$\emptyset = \{\}$

Properties of the Empty Set

Let A be any set.

- $A \cup \emptyset = A$
- $A \cap \emptyset = \emptyset$
- $\emptyset' = U$
- $\emptyset \subseteq A$

Disjoint Sets

Definition

A and B are *disjoint* if $A \cap B = \emptyset$.

Example

- Are $A = \{4, 6, 9\}$ and $B = \{2, 3, 6, 8\}$ disjoint?
- No, $A \cap B$ is nonempty.
- Are $C = (3, 6)$ and $D = [7, 10]$ disjoint?
- Yes, $C \cap D = \emptyset$.

Mutually Exclusive Sets

Definition

Let $A_1, A_2, \ldots, A_k$ be sets such that

$$\text{if } i \neq j, \text{ then } A_i \cap A_j = \emptyset.$$

Then the sets $A_1, A_2, \ldots, A_k$ are *mutually exclusive* or *pairwise disjoint*.

Example

Determine whether the sets are mutually exclusive.

- $A_1 = \{1, 2\}$, $A_2 = \{3, 4\}$, $A_3 = \{4, 5\}$, and $A_4 = \{6, 7\}$.
- No, because $A_2 \cap A_3$ is nonempty.
- $B_1 = \{6, 8\}$, $B_2 = \{1, 4\}$, and $B_3 = \{5, 11\}$.
- Yes, because $B_1 \cap B_2$, $B_1 \cap B_3$, and $B_2 \cap B_3$ are all empty.

Example

Determine whether the sets are mutually exclusive.

- $A_1 = (0, 2)$, $A_2 = (5, 8)$, $A_3 = [10, 12]$, and $A_4 = [15, 16]$.
- Yes.
- $B_1 = [3, 7]$, $B_2 = (6, 8)$, and $B_3 = (10, 12)$.
- No, because $B_1 \cap B_2$ is nonempty.

Outline

Events

The universal set is the sample space S.

Definition

A subset $A \subseteq S$ is called an *event*.

Example

Rolling a die, our sample space is

$$S = \{1, 2, 3, 4, 5, 6\}.$$

Then $A = \{2, 4, 6\}$ is the event that an even number is rolled.

Probability Measures

Definition

A *probability measure* is a function P that assigns to each event $A \subseteq S$ a number $P(A)$, called the *probability* of A, such that the following properties are satisfied:

- $P(A) \geq 0$ for all events A
- $P(S) = 1$
- If $A_1, A_2, \ldots, A_k$ are mutually exclusive, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_k) = P(A_1) + P(A_2) + \cdots + P(A_k).$$

- If $A_1, A_2, \ldots$ are mutually exclusive, then

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots .$$

Example

A basketball player shoots five free throws, and the number of successful shots is noted.

- What is the sample space?
- Let $A_1 = \{0, 1\}$, $A_2 = \{2, 3\}$, and $A_3 = \{4, 5\}$, and suppose $P(A_1) = 0.3$ and $P(A_2) = 0.5$.
- Find $P(A_3)$.
- 0.2

Example

Suppose $S = \{1, 2, 3, 4\}$, and assume $P(\{x\}) = \frac{x}{10}$, for every $x \in S$.

- Find $P(\{2, 4\})$.
- Find $P(\{1, 2\})$.

Additional Properties

Proposition

For any event A,

$$P(A) = 1 - P(A').$$

Example

Suppose $S = \{1, 2, 3, \ldots, 10\}$, and assume $P(\{x\}) = \frac{x^2}{385}$, for every $x \in S$. Find $P(\{3, 4, \ldots, 10\})$.

Proposition

- $P(\emptyset) = 0$
- If $A \subseteq B$, then $P(A) \leq P(B)$.
- For any event A, $P(A) \leq 1$.

The Inclusion Exclusion Formula

Proposition

For any events A and B,

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

Example

There is a 30% chance it will rain tomorrow and a 40% chance the temperature will exceed 70 degrees. There is a 10% chance it will rain and the temperature will exceed 70 degrees. Find the probability that it will rain or the temperature will exceed 70 degrees.

Inclusion Exclusion Formula for Three Sets

Proposition

For any events A, B, and C,

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

Equally Likely Outcomes

Proposition

If all outcomes in the sample space are equally likely, then

$$P(A) = \frac{N(A)}{N(S)},$$

where $N(A)$ is the number of outcomes in A, for any event A.

Example

What's the probability of rolling an even number on a fair die?

Example

A card is selected at random from a standard 52 card deck. Let

$$A = \{x \,|\, x \text{ is a heart}\}$$
$$B = \{x \,|\, x \text{ is a jack}\}$$
$$C = \{x \,|\, x \text{ is a jack, queen, or king}\}$$

Calculate

- $P(A)$
- $P(B)$
- $P(A \cup B)$
- $P(A \cap C)$

Example

A fair coin is tossed three times, and we assume that the outcomes in the sample space are equally likely. Find the probability that

- all tosses are heads
- at least 2 tosses are heads
- at least 1 toss is heads
- the first toss is heads and the last is tails

Outline

The Multiplication Principle

The Multiplication Principle (Two Steps)

Assume an experiment satisfies the following:

- it consists of 2 steps,
- there are n_1 possible outcomes for the first step, and
- there are n_2 possible outcomes for the second step, *regardless of the outcome on the first step*.

Then there are $n_1 n_2$ possible outcomes for the experiment.

Example

An ice cream sundae is constructed by choosing a flavor of ice cream, vanilla, chocolate, or strawberry, and a topping, fudge or caramel. How many possible sundaes are there?

Example

Suppose you have two urns. One urn contains a silver dollar and a quarter, while the other urn contains a dime, a nickel, and a penny. You select an urn at random and then select a coin at random from that urn. How many possible outcomes are there?

The Multiplication Principle (Multiple Steps)

Assume an experiment satisfies the following:

- it consists of m steps, and
- there are n_i possible outcomes for the ith step, regardless of the outcomes of the previous steps.

Then there are $n_1 n_2 \cdots n_m$ possible outcomes for the experiment.

Example

Suppose you have 10 shirts, 5 pairs of pants, 3 pairs of shoes, and 4 hats. An outfit consists of a shirt, a pair of pants, a pair of shoes, and a hat. How many possible outfits are there?

Example

How many ways can the letters in the word “math” be rearranged?

Permutations

There are $n!$ ways to arrange n objects in order. The different arrangements are called *permutations*.

Example

A traveling salesman has to visit 5 cities. An itinerary for his trip consists of a list of these cities in the order that he will visit them. How many possible itineraries are there?

Example

How many ways can a chairperson, a secretary, and a treasurer be selected from an 8 person committee?

${}_n\mathrm{P}_r$

- An ordered arrangement of r objects from a larger set of n objects is called a *permutation of n objects taken r at a time*.
- There are ${}_n\mathrm{P}_r = \frac{n!}{(n-r)!}$ such permutations.

Example

How many 6-letter passwords can be formed if repetitions are not allowed?

Example

Seven people are selected at random, without replacement from a group of 30, and they are noted in the order that they are selected. How many possible outcomes are there?

Ordered Samples without Replacement

When an *ordered* sample of size r is selected from a population of size n *without* replacement, the number of possible samples is ${}_n\mathrm{P}_r$.

Counting when Order Matters with Replacement/Repetitions

Example

How many 6-letter passwords can be formed if repetitions are allowed?

Example

Seven people are selected at random, with replacement from a group of 30, and they are noted in the order that they are selected. How many possible outcomes are there?

Combinations

Example

How many 5-card poker hands are there?

A set containing n elements has $\binom{n}{r}$ (unordered) subsets of size r, where

$$\binom{n}{r} = {}_nC_r = \frac{n!}{r!(n-r)!}.$$

- These numbers are called *binomial coefficients*, and
- the function ${}_nC_r$ is called the *combination* function.

Example

Seven people are selected at random, without replacement from a group of 30. How many possible samples are there?

Unordered Samples without Replacement

When an *unordered* sample of size r is selected from a population of size n *without* replacement, the number of possible samples is $\binom{n}{r}$.

Example

How many subcommittees containing 3 people can be formed in a committee containing 8 people?

Example

An urn contains 10 red balls and 8 green balls. If 7 balls are selected at random without replacement, what is the probability that 4 of them are red?

Example

In a 5-card poker hand, what is the probability of selecting 2 kings and 2 queens?

Distinguishable Permutations

Example

How many 10-letter passwords consist of 7 A's and 3 D's?

Example

How many 10-letter passwords consist of 5 A's, 3 B's, and 2 C's?

Example

In how many distinguishable ways can 4 red balls, 3 green balls, and 7 yellow balls be arranged, assuming that balls of the same color are indistinguishable?

Example

Find the coefficient of x^4y^2 in the expansion of $(x+y)^6$.

Example

Find the coefficient of $x^2y^5z^7$ in the expansion of $(x+y+z)^{14}$.

Example

If you buy 10 fruits consisting of apples, oranges, and bananas, how many possible outcomes are there?

Example

In how many ways can 15 identical loaves of bread be distributed amongst 7 houses?

Example

A Hyundai dealer sells the following models of cars: Accent, Elantra, Sonata, and Santa Fe. The available colors are silver, black, and blue, and there are 9 available options (leather interior, tinted windows, etc.).

- How many different cars can be made, assuming 4 options are selected?
- How many different cars can be made, assuming 0 to 9 options are selected?
- How many different cars can be made, assuming 5 to 7 options are selected?

Yahtzee and Poker

Example

If 5 six-sided dice are rolled, what is the probability of getting

- five of a kind?
- four of a kind?
- three of a kind?
- a full house?
- one pair?
- two pair?

Example

Find the probability of getting one pair in a five card poker hand.

Outline

Example

At a university with 1000 students, 185 students have taken an upper level math course, 162 students have taken an upper level physics course, and 114 have taken both. What is the probability that a randomly selected student

- has taken an upper level math course?
- has taken an upper level physics course?
- has taken an upper level math course, given that he or she has taken an upper level physics course?
- has taken an upper level physics course, given that he or she has taken an upper level math course?

Definition

Suppose A and B are events, and $P(B) > 0$. Then the *conditional probability* of *A given B* is

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

Example

If two fair dice are rolled, what is the probability that one of the dice is a 4, given that the sum of the dice is 10?

Equally Likely Outcomes

If all outcomes in the sample space are equally likely,

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

Example

In a five-card poker hand, find the probability of drawing at least 3 kings, given that at least 2 kings are drawn.

Example

The following table describes the genders and colors of 10122 mice.

	Male	Female	Total
White	2	110	112
Black	145	9865	10010
Total	147	9975	10122

Compute the probability that a randomly selected mouse is

- white, given that it's female.
- black, given that it's male.
- male, given that it's black.
- Is a male mouse likely to be black?
- Is a black mouse likely to be male?

Example

Suppose A and B are events such that $P(A) = 0.5$, $P(A \cap B') = 0.2$, and $P[(A \cup B)'] = 0.1$. Find $P(A \mid B)$.

Multiplication Rule

For any events A and B,

$$P(A \cap B) = P(A)P(B \mid A), \text{ and}$$
$$P(A \cap B) = P(B)P(A \mid B),$$

assuming these conditional probabilities are defined.

Example

Suppose an urn contains 7 red balls and 10 green balls. If two balls are drawn successively without replacement, what is the probability that

- the first ball is red and the second is green?
- the first ball is green and the second is red?
- both balls are red?
- both balls are green?

Example

Suppose an urn contains 7 red balls and 10 green balls. If two balls are drawn successively without replacement, what is the probability that the second ball drawn is red?

Example

At a certain company, it is known that 90% of workers who have completed a training program will meet their quotas, while only 70% of other workers will do so. Assuming 80% of the company's workers have completed the program, what percentage of the company's workers will meet their quotas?

Example

Balls are drawn sequentially without replacement from the above urn until the 4th red ball is drawn. What is the probability that this happens on the 8th draw?

Example

It is known that there are 4 defective batteries in a box of 20. Batteries are randomly selected from the box and tested, one by one, until all four defective batteries ar found.

- What is the probability that 12 batteries will be tested in this process?
- Find the probability that at least 7 batteries will need to be checked to find all 4 defective batteries.

Example

A bowl contains 9 red chips and 1 blue chip. 10 people successively draw chips from the bowl without replacement, until the blue chip is drawn. The person who draws the blue chip wins a prize.
Would you prefer to draw first, second,. . ., or last?

Outline

Definitions

Definition (Informal)

Events A and B are *statistically independent* if

- knowing that A happened gives us no information about whether B will or has happened
- and vice versa

Example

A fair die is rolled twice in a row. Let

- $A = \{\text{the first roll is a 3}\}$
- $B = \{\text{the second roll is a 5}\}$

Definition (Formal)

Events A and B are *statistically independent* if

- $P(A \cap B) = P(A)P(B)$
- Also called “stochastically independent”
- Usually just called “independent”

Example

A fair die is rolled twice in a row. Let

- $A = \{\text{the first roll is a } 3\}$
- $B = \{\text{the second roll is a } 5\}$
- Are these events independent?

Example

Suppose

- $A = \{\text{the first roll is a 5}\}$
- $B = \{\text{the sum of the two rolls is 11}\}$
- Are these events independent?

Example

Suppose $P(A) = 0.4$ and $P(B) = 0.5$. Find $P(A \cup B)$ if A and B are

- independent
- disjoint

A' and B'

Proposition

If A and B are indepenent, then

- A and B' are independent
- A' and B are independent
- A' and B' are independent

Example

Joe parks illegally on Monday and Tuesday. He has a 30% chance of getting a ticket on Monday and a 40% chance of getting a ticket on Tuesday. If receiving a ticket on Monday is independent of receiving a ticket on Tuesday, what is the probability that Joe does not receive a ticket on either day?

Definition

Events A, B, and C are *mutually independent* if the following conditions hold:

- They are *pairwise independent*:
  - $P(A \cap B) = P(A)P(B)$
  - $P(A \cap C) = P(A)P(C)$
  - $P(B \cap C) = P(B)P(C)$
- $P(A \cap B \cap C) = P(A)P(B)P(C)$

Example

An urn contains four balls numbered 1 to 4. One ball is drawn at random, and

- $A = \{1, 2\}$
- $B = \{1, 3\}$
- $C = \{1, 4\}$
- Are the events A, B, and C pairwise independent? Are they mutually independent?

Machine Components

Example

The failure probabilities for three machine components are 0.1, 0.3, and 0.5. Assuming the machine components are statistically independent, calculate the probability that

- all of the components fail
- at least one of the components fail
- exactly one of the components fails

Basketball Free Throws

Example

When Alice shoots a free throw, she has an 80% chance of making it, and her free throws are independent. In a sequence of 10 free throws, what is the probability that

- she makes exactly 7 of them?
- she makes at least 9 of them?

If Alice shoots indefinitely, what is the probability that her 3rd miss will occur on the 12th throw?

Urn Problems

Example

An urn contains 6 red balls and 4 green balls.

- If you draw five balls from the urn with replacement, what is the probability of drawing 3 red balls and 2 green balls?
- Rework this problem assuming the balls are not replaced.
- Rework this problem, assuming the balls are not replaced and assuming there are 60 red balls and 40 green balls in the urn.

Pistol Duels

Example

- Pistol duel
- Take turns shooting
- Both shooters have 75% accuracy
- Probability of winning if you shoot first?

Example

Pat's accuracy is $\frac{2}{3}$ and Quin's accuracy is $\frac{1}{2}$. Find the probability that Pat wins if he shoots first.

Outline

Example

Suppose urn 1 contains 3 red balls and 3 green balls, urn 2 contains 1 red ball and 2 green balls, and urn 3 contains 2 red balls and 3 green balls. You select one of these urns at random so that urns 1, 2, and 3 have probabilities $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{1}{6}$ of being selected, respectively. Then, you select a ball at random from the chosen urn. If the ball is green, what is the probability that you selected urn 3?

Example

Previous example about a company's workers

- 80% of the workers completed a training program
- 90% of workers completing the training program meet their quotas
- 70% of other workers meet their quotas
- Assuming a worker met his quota, find the probability that he completed the training program

Example

About 1 in 10,000 people have a certain disease. There is a test for this disease that makes two types of errors:

- for people without the disease, it yields a false positive 5% of the time
- for people with the disease, it yields a false negative 2% of the time

If the test identifies someone as having the disease, what is the probability that he or she actually has the disease?

Example

Seeds from supplier A have an 85% germination rate and those from supplier B have a 75% germination rate. A seed packaging company purchases 40% of their seeds from supplier A and 60% from supplier B. Given that a seed germinates, what is the probability that it was purchased from supplier A?

Bayes's Theorem

Theorem

Suppose the events $B_1, \ldots, B_m$ *form a partition of the sample space* S*, and assume* $P(B_i) > 0$*, for* $i = 1, 2, \ldots, m$*. For any event* A *such that* $P(A) > 0$*,*

$$P(B_k \mid A) = \frac{P(B_k)P(A \mid B_k)}{\sum_{i=1}^{m} P(B_i)P(A \mid B_i)}.$$

Review for Exam One

- Section 1.1: Basic Concepts
  - Sample space
  - Frequency, relative frequency, histograms
- Review of Set Theory
  - $\in, \cup, \cap, A', \subseteq, \emptyset$
  - Disjoint sets; mutually exclusive sets
- Section 1.2: Properties of Probability
  - Sample space; events
  - Probability axioms
  - Other properties of probability
  - Simple proofs and calculations
  - Venn diagrams
  - Equally likely outcomes

- Section 1.3: Methods of Enumeration
  - Multiplication principle
  - Permutations
  - ${}_nP_r$
  - $\binom{n}{r}$
  - Distinguishable permutations
  - Coding scheme. Example: 000|00||0|00
  - Combining different methods
- Section 1.4: Conditional Probability
  - $$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$
  - Equally likely outcomes:
    $$P(A \mid B) = \frac{N(A \cap B)}{N(B)}$$
  - Trees
- Section 1.5: Independent Events
  - Definition of independence: $P(A \cap B) = P(A)P(B)$
- Section 1.6: Bayes's Theorem