MS 302 In-class Problems

May 12, 2025

Table of Contents

Chapter 1 Section 1

Suppose that you produce structural metal bars by heat treating them using induction heating and that they must withstand 10,000 lbs.

You set as your goal a breaking point of 15,000 lbs.

Will the breaking point be exactly 15,000 lbs?

Almost every procedure we deal with has some form of variation built into it.

1. The failure point of structral metal bars
2. How long it takes to brush your teeth in the morning
3. The location that a basketball hits as you practice your free throw shot.

In order to understand and maybe even control the *uncertainty* and *variation* we use

Definition 1 (statistical inference)

The drawing of inferences about a population based on data taken from a sample of that population.*

Two *factors* that contribute to the strength of a bar in our process are temperature and time in the heat treatment.

Definition 2 (factors)

are properties or characteristics of a population.

*OED

To determine how much weight our bars can support we could put them into a machine that destructively tests them. But...?

Definition 3 (Population)

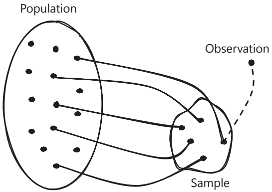
collection of all individuals in a particular scientific system or process.

Definition 4 (Sample)

collection of observations from the population.

Definition 5 (Observation)

individual or item of a particular type involved in the scientific system or process.



In addition to *Inferential statistics* there is *Descriptive statistics*.

Definition 6 (Descriptive statistics)

are used when seeking only to gain some summary of a set of data represented by a sample (single-number statistics that provide a sense of center, variability, or general nature of the distribution of the sample data).

An understanding of probability provides a basis to understand statistical inference.

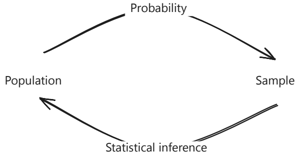
Ex The average height (in inches) of a male JSU student is $\mu_M = 68$ and the standard deviation is $\sigma_M = 3$ and for women $\mu_F = 63$ and $\sigma_F = 2.7$. Now suppose that someone believes that the average male JSU student height is $\mu_M = 63$. How can we show that this is not true without measuring almost every male JSU student? We can show that it is *probably* not true using statistical inference. Take a sample of heights

$$69.8, 68.0, 70.2, 70.6, 65.4, 72.0, 73.3, 63.1, 73.5, 69.1, 66.8, 64.8,$$

$$64.0, 66.4, 65.1, 70.8, 75.4, 71.5, 67.1, 66.0, 64.1, 64.9, 69.4, 68.2, 69.0$$

This $n = 25$ sample data set has a sample mean $\bar{x} = 68.4$ and sample standard deviation $s = 3.3$, but let's use $\sigma_M = 3$.

Population with know features + Probability: allow us to draw conclusions about characteristics of hypothetical data taken from the population.



Sample + Inferential Statistics: allow us to draw conclusions about the population.

Chapter 1 Section 2

Note to myself: Open mpg.csv

Definition 7 (Simple Random Sample)

Given a specified sample size, then every sample is as likely to be chosen as any other sample.

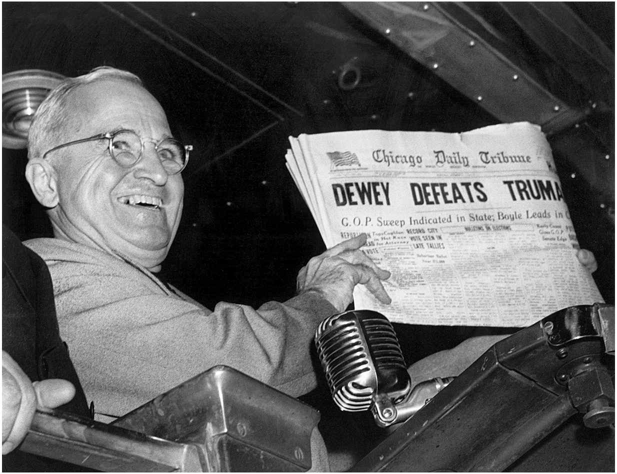
Ex Suppose that a city has two malls and you want to determine on average how many people visit each store per day. You could chose to use a sample of size of 10 and then

1. chose a mall at random
2. pick 10 stores at random.

While this is a *random sample* it is not a simple random sample. Why?
Because while every store has an equally likely chance of being chosen, there is no way to get a store from mall 1 and mall 2 in a sample at the same time. Therefore there are samples that could never be chosen.

Definition 8 (Biased Sample)

is a sample that does not accurately represent the population (it over/under-represents some segment of the population).



Chapter 1 Section 3
Measures of Location

Definition 9 (Sample mean)

If the observations in a sample are $x_1, x_2, \ldots, x_n$, *then the sample mean is*

$$\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Problem is that it is influenced by outliers. Note that as a single value gets larger so does the mean.

Ex Find the sample mean of

$$50000, 30000, 45000, 33000, 47000, 51000, 6744000$$

Definition 10 (Sample median)

If the observations in a sample are $x_1, x_2, \ldots, x_n$ *are arranged in* **increasing** *order, then the sample median is*

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}}, & \text{if } n \text{ odd} \\ \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right), & \text{if } n \text{ even} \end{cases}$$

Ex Find the sample median of

$$50000, 30000, 45000, 33000, 47000, 51000, 6744000$$

Chapter 1 Section 4
Measures of Variability

Both data sets

$$1, 2, 3, 7, 8, 9$$

and

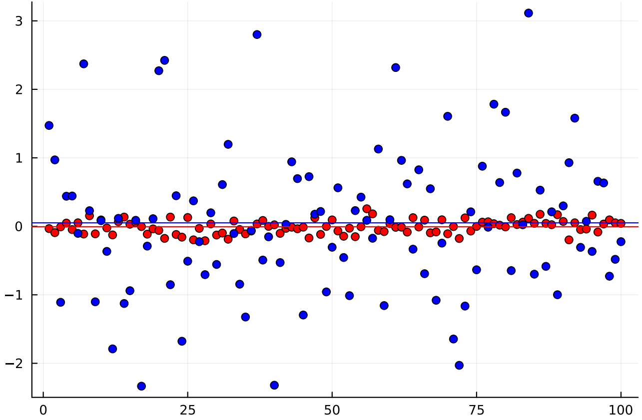
$$1, 1, 1, 9, 9, 9$$

have a mean of 5.
But the data sets are clearly different.

These two data sets of 100 points each have slightly different centers (means)

$$\mu_{red} = -0.006, \quad \mu_{blue} = 0.051$$

but there spreads are very different.



How should we measure this variation in the data?

Definition 11

If the observations in a sample are $x_1, x_2, \ldots, x_n$, then

sample range $x_{max} - x_{min}$

sample variance

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$$

sample standard deviation

$$s = \sqrt{s^2}$$

Ex Find the sample range, variance, and standard deviation for

$$1.7, 2.2, 3.9, 3.11, 14.7$$

Chapter 1 Section 5
Discrete and Continuous Distributions

Definition 12 (Discrete data)

can only take on certain values.

Example 5.1

1. *Number of students in my class, e.g.* $0, 1, 2, \ldots, 35$
2. *FM radio station frequencies can be from 88.1 MHz to 108.1 (or 107.9) MHz with a step size of 0.2 MHz, e.g. 106.9, 100.7, 103.5, 95.3, 95.1*

Definition 13 (Continuous data)

can take on any value in a range.

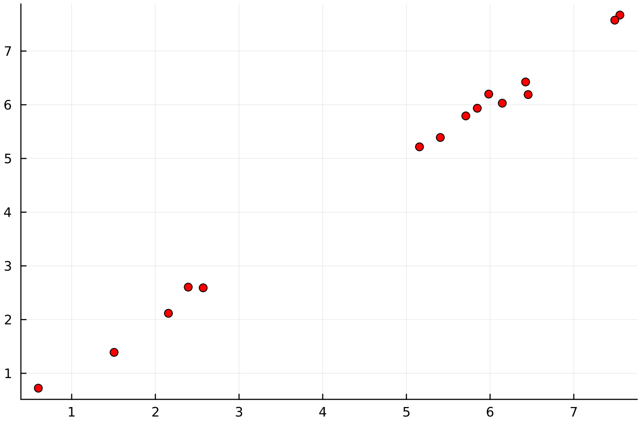
Example 5.2

Heights of adult humans can be anything from 20 *inches to* 110 *inches.*

Chapter 1 Section 6

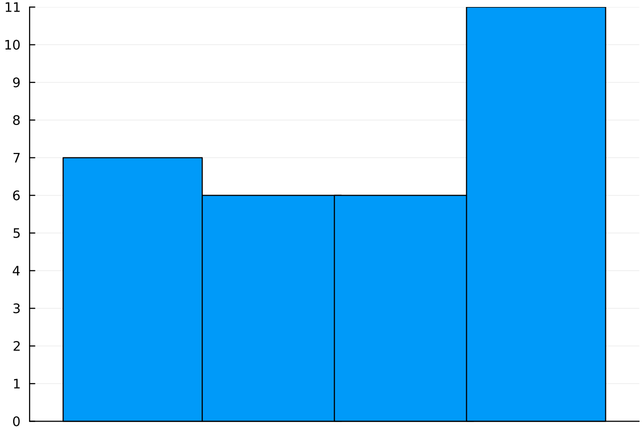
Scatter plot

x	y
0.601	0.723
1.507	1.390
2.156	2.118
2.393	2.605
2.572	2.592
5.156	5.216
5.405	5.391
5.710	5.793
5.847	5.935
5.984	6.199
6.146	6.030
6.425	6.424
6.455	6.190
7.490	7.575
7.551	7.670



Histogram

Sam data as before:
$20, 20, 20, 21, 22, 23, 27, 33, 33, 35, 38, 39, 39, 45, 46, 47, 47, 48, 48,$
$50, 52, 52, 53, 53, 54, 54, 56, 56, 56, 58$



Steam and leaf

Ordered data:
$20, 20, 20, 21, 22, 23, 27, 33, 33, 35, 38, 39, 39, 45, 46, 47, 47, 48, 48,$
$50, 52, 52, 53, 53, 54, 54, 56, 56, 56, 58$
The stem are the digits excluding the least significant digit and they lie on the left of the plot.
The leaves are what remains after removing the stem and are written to the right and are vertically aligned.

2	0001237
3	335899
4	567788
5	02233446668

Key $3|5 = 35$

Chapter 2 Section 1

Definition 14 (Experiment)

is any process that generates a set of data.

Definition 15 (Sample space)

is the set of all possible outcomes of a statistical experiment.

It is denoted by S.

Ex Flip a coin. What is the sample space as a set and a tree.

Ex Flip a coin twice. What is the sample space as a set and a tree.

Ex Flip a coin, if H roll a 4-sided die, else flip the coin again. What is the sample space as a set and a tree.

Ex Flip a coin until you get a H. What is the sample space as a set and a tree.

Ex Consider the set S to be all of the points that make up a circle with radius 5 that is centered at the origin.
List the sample space.

Chapter 2 Section 2

Definition 16 (Event)

is a subset of a sample space.

Definition 17 (Null (Empty) Set)

is the set with no elements.

Denoted by $\emptyset$.

Ex Flip a coin. List all events.

Ex Flip a coin twice.

1. $E_1 \equiv$ flip exactly one H.
2. $E_2 \equiv$ flip at least one H.
3. $E_3 \equiv$ the second flip is the same as the first.
4. $E_4 \equiv$ the coin lands on it's edge.

Ex Let $S = \{x : x \geq 0\}$ be the distance Forrest Gump ran before he tripped.
What is the event:

1. A that he trips after the 50^{th} mile and before or at the 90^{th} mile?
2. B that he never trips?

Definition 18 (Complement)

of the event A with respect to S is the set of all elements in S not in A.

Denoted as A'

Ex 2.15 Consider the sample space

$$S = \{\text{copper, sodium, nitrogen, potassium, uranium, oxygen, zinc}\}$$

and the events

$$A = \{\text{copper, sodium, zinc}\},$$
$$B = \{\text{sodium, nitrogen, potassium}\},$$
$$C = \{\text{oxygen}\}.$$

List the elements of the sets corresponding to the following events:

(a) A'

Definition 19 (Intersection of the events A and B)

is the set

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

Ex 2.15 Consider the sample space

$S = \{$copper, sodium, nitrogen, potassium, uranium, oxygen, zinc$\}$

and the events

$$A = \{\text{copper, sodium, zinc}\},$$
$$B = \{\text{sodium, nitrogen, potassium}\},$$
$$C = \{\text{oxygen}\}.$$

List the elements of the sets corresponding to the following events:

1. $A \cap B$

Definition 20 (Mutually disjoint)

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

Ex 2.15 Consider the sample space

$$S = \{\text{copper, sodium, nitrogen, potassium, uranium, oxygen, zinc}\}$$

and the events

$$A = \{\text{copper, sodium, zinc}\},$$
$$B = \{\text{sodium, nitrogen, potassium}\},$$
$$C = \{\text{oxygen}\}.$$

Which pairs of A, B, and C are disjoint?

Definition 21 (Union of the events A and B)

is the set

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

Ex 2.15 Consider the sample space

Consider the sample space

$$S = \{\text{copper, sodium, nitrogen, potassium, uranium, oxygen, zinc}\}$$

and the events

$$A = \{\text{copper, sodium, zinc}\},$$
$$B = \{\text{sodium, nitrogen, potassium}\},$$
$$C = \{\text{oxygen}\}.$$

List the elements of the sets corresponding to the following events:

(b) $A \cup C$

(c) $(A \cap B') \cup C'$

Ex 2.15

(d) $B' \cap C'$

(e) $A \cap B \cap C$

(f) $(A' \cup B') \cap (A' \cap C)$

Chapter 2 Section 3
Counting

If we flip a fair coin, what is the probability of getting a head?

$$P(H) = \frac{1}{2}$$

We will see that the way we calculate probabilities is that

$$P(E) = \frac{\text{how many ways to get the event } E}{\text{total number of outcomes}}$$

So what we need to figure out how to do is count.

If we flip a coin and then roll a 20-sided die, how many outcomes are possible?

Definition 22 (Multiplication rule)

If there are n_1 ways to perform an operation and each of these are followed by n_2 ways to complete a second operation, then there is a total of

$$n_1 \cdot n_2$$

ways to perform both operations.

Ex Flip a coin and then roll a 20-sided die.

Ex 2.24 Students at a private liberal arts college are classified as being freshmen, sophomores, juniors, or seniors, and also according to whether they are male or female. Find the total number of possible classifications for the students of that college.

Ex 2.25 A certain brand of shoes comes in 5 different styles, with each style available in 4 distinct colors. If the store wishes to display pairs of these shoes showing all of its various styles and colors, how many different pairs will the store have on display?

Definition 23 (Generalized multiplication rule)

If theree are k different tasks to be accomplished in sequence with n_1 ways to do the first, n_2 ways to do the second, …, n_k ways to do the k^{th}, then the total number of ways to accomplish these tasks is

$$n_1 \cdot n_2 \cdots n_k.$$

Ex How many ways can you pick a lower case letter, pick a digit, and then flip a coin to get H or T?

Ex 2.27 A developer of a new subdivision offers a prospective home buyer a choice of 4 designs, 3 different heating systems, a garage or carport, and a patio or screened porch. How many different plans are available to this buyer?

Ex Current Alabama license plates consist of a county code (there are 67 counties in Alabama - coded using the format 1 or 10 - but lets pretend we have a single character for each one, think hexadecimal - base 16 - but at least base 67) followed some number/letter combination depending on the year. See Vehicle registration plates of Alabama

1. In 1941 Alabama started using county codes in its license plates. The format was county code, followed by three digits. How many license plates can be issued.

2. Starting in January 2002 it consisted of the county code, a letter, three digits, a letter. How many license plates can be issued.

3. The current rule is it consists of the county code followed by 5 alphanumeric digits. How many license plates can be issued.

Suppose you have 3 books that you want to put on a shelf

Math, Computer Science, Physics

In how many ways can we do this?

Definition 24 (Permutation)

is an arrangement of a set of objects, either all or part of them.

Definition 25 (Factorial)

$$n! = n(n-1)\cdots 1; \quad n \in \mathbb{N}$$

and

$$0! = 1$$

Theorem 1

The number of permutations of n *objects is* $n!$.

Ex 2.32 (a) In how many ways can 6 people be lined up to get on a bus?

Ex 2.34

(a) How many distinct permutations can be made from the letters of the word COLUMNS?

(b) How many of these permutations start with the letter M?

Ex 2.35 A contractor wishes to build 9 houses, each different in design. In how many ways can he place these houses on a street if 6 lots are on one side of the street and 3 lots are on the opposite side?

Ex 2.37 In how many ways can 4 boys and 5 girls sit in a row if the boys and girls must alternate?

If there are 5 boys and 5 girls.

Sometimes you don't need all n objects.

Definition 26 (Permutation of n objects taken r at a time)

$$_nP_r = \frac{n!}{(n-r)!}$$

Ex 2.40 In how many ways can 5 starting positions on a basketball team be filled with 8 men who can play any of the positions?

When order does not matter

Definition 27 (Combination of n objects taken r at a time)

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

Ex 2.26 A California study concluded that following 7 simple health rules can extend a man's life by 11 years on the average and a woman's life by 7 years. These 7 rules are as follows: no smoking, get regular exercise, use alcohol only in moderation, get 7 to 8 hours of sleep, maintain proper weight, eat breakfast, and do not eat between meals. In how many ways can a person adopt 5 of these rules to follow

(a) if the person presently violates all 7 rules?

(b) if the person never drinks and always eats breakfast?

Ex 2.22 In a medical study, patients are classified in 8 ways according to whether they have blood type AB^+, AB^-, A^+, A^-, B^+, B^-, O^+, or O^-, and also according to whether their blood pressure is low, normal, or high. Find the number of ways in which a patient can be classified.

Ex How many four-digit numbers can be formed from the digits 8, 1, 2, 5, 6, and 9 where each digit is used once?

Ex A traditional poker hand has how many options?

Chapter 2 Section 4
Probability

Definition 28 (Probability)

is a fucntion $P : \mathcal{E} \to [0, 1]$ *such that for any event* A *in* $\mathcal{E}$ *such that the following are satisfied:*

1. $0 \leq P(A) \leq 1$
2. $P(\emptyset) = 0$
3. $P(S) = 1$

and if $A_1, A_2, \ldots$ *are mutually exclusive then*

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

Ex 2.49 Find the errors in each of the following statements:

1. The probabilities that an automobile salesperson will sell 0, 1, 2, or 3 cars on any given day in February are, respectively,

$$0.19, 0.38, 0.29, \text{ and } 0.15$$

2. The probability that it will rain tomorrow is 0.40, and the probability that it will not rain tomorrow is 0.52.

3. The probabilities that a printer will make 0, 1, 2, 3, or 4 or more mistakes in setting a document are, respectively,

$$0.19, 0.34, -0.25, 0.43, \text{ and } 0.29$$

4. On a single draw from a deck of playing cards, the probability of selecting a heart is $\frac{1}{4}$, the probability of selecting a black card is $\frac{1}{2}$, and the probability of selecting both a heart and a black card is $\frac{1}{8}$.

In this chapter all sample spaces S are finite.
The probability of an event A is the sum of the probability of every element in A.

Ex Flip a fair coin once.

1. Find the sample space.

2. State the set of all events.

3. What are
   3.1 $P(\emptyset) =$

   3.2 $P(\{H, T\}) = P(H, T) =$
4. Derive $P(H)$ and $P(T)$.

Example (pg 53) 2.25 A die is loaded in such a way that an even number is twice as likely to occur as an odd number. If E is the event that a number less than 4 occurs on a single toss of the die, find $P(E)$.

Theorem 2

Let an experiment have N different equally likely outcomes. If the event A consists of exactly n of these outcomes, then

$$P(A) = \frac{n}{N}$$

Ex Flip a coin twice ad let $A = \{HT, TH, TT\}$. Find $P(A)$.

Ex 2.51 A box contains 500 envelopes, of which 75 contain \$100 in cash, 150 contain \$25, and 275 contain \$10. An envelope may be purchased for \$25. What is the sample space for the different amounts of money? Assign probabilities to the sample points and then find the probability that the first envelope purchased contains less than \$100.

Ex 2.52 Statement Suppose that in a senior college class of 500 students it is found that 210 smoke, 258 drink alcoholic beverages, 216 eat between meals, 122 smoke and drink alcoholic beverages, 83 eat between meals and drink alcoholic beverages, 97 smoke and eat between meals, and 52 engage in all three of these bad health practices.

Ex 2.52 Questions If a member of this senior class is selected at random, find the probability that the student

(a) smokes but does not drink alcoholic beverages;

(b) eats between meals and drinks alcoholic beverages but does not smoke;

Example (pg 55) 2.28 In a poker hand consisting of 5 cards, find the probability of holding 2 aces and 3 jacks.