Elementary Mathematics – Statistics

1. Probability **2. Combined Events** **3. Statistical Measures**

1 Probability

Probability theory turns randomness into structure.

Introduction to Probability

Probability measures how likely an event is to happen. A probability of 1 means the event is certain, 0 means it is impossibility, and 0.35 means there is a 35% chance of it happening.

Probability of
- $P(E) = \dfrac{\text{Number of outcomes that match event } E}{\text{Total number of possible outcomes}}$

Event E

Example 1:

A card is drawn from a standard deck of 52 playing cards.

a) Find the probability that the card is a spade.

b) Find the probability that the card a black ace.

The card drawn is a 9 of clubs. A second card is then drawn without replacement.

c) Find the probability that the second card beats the first (larger number, ace included).

d) Find the probability that the second card is of the same suit.

a) – Total No. of cards $= 52$
- No. of spades $= 13$
- $P(\text{draw a spade}) = \dfrac{13}{52} = \dfrac{1}{4}$

b) – No. of black aces $= 2$
- $P(\text{draw a black ace}) = \dfrac{2}{52} = \dfrac{1}{26}$

c) – No. of cards left $= 51$
- No. of cards larger than 9 left $= 5 \times 4 = 20$
- $P(\text{draw a card larger than 9 next}) = \dfrac{20}{51}$

d) – No. of clubs left $= 12$
- $P(\text{draw another club}) = \dfrac{12}{51}$

Events as Sets

A convenient way of expressing events is by using set notation. If we define S as the sample space (the set of all possible outcomes), then any event E can be expressed as a subset of S.

- $P(E) = \dfrac{\text{Number of elements in event } E}{\text{Number of elements in sample space } S} = \dfrac{n(E)}{n(S)}$

Example 1:

A fair 6-sided die is rolled.

a) List the sample space S in roster notation along with its count.

b) Find the probability of rolling an odd number.

c) Find the probability of rolling a perfect square.

a) – $S = \{1, 2, 3, 4, 5, 6\}$ – $n(S) = 6$

b) – Let E_O be the event of rolling an odd number.
- $E_O = \{1, 3, 5\}$ – $P(E_O) = \dfrac{n(E_O)}{n(S)} = \dfrac{3}{6} = \dfrac{1}{2}$

c) – Let E_P be the event of rolling a perfect square.
- $E_P = \{1, 4\}$ – $P(E_P) = \dfrac{n(E_P)}{n(S)} = \dfrac{2}{6} = \dfrac{1}{3}$

Example 2:

A letter is chosen randomly from the alphabet.

a) Find the probability the letter is a vowel.

b) Find the probability the letter is contained in "MATH HAWKER".

a) – Let E_V be the event of choosing a vowel.
- $P(E_V) = \dfrac{n(E_V)}{n(S)} = \dfrac{5}{26}$

b) – Let E_M be the event of choosing a relevant letter.
- $E_M = \{M, A, T, H, W, K, E, R\}$
- $P(E_M) = \dfrac{n(E_M)}{n(S)} = \dfrac{8}{26} = \dfrac{4}{13}$

Complement of Event

Being a set, an event can have a complement made up of all outcomes where the event does not happen. In probability, working with complements is even simpler with the following identity.

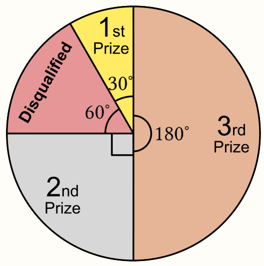
- $P(E') = 1 - P(E)$

Example 1:

For a game event, the wheel below is spun quickly while a contestant threw a dart to determine their prize. Assume the location the dart lands is random.

a) Find the probability of winning the second prize.

b) Find the probability of not getting disqualified.



a) – Let E_S be the event of winning the second prize.
- $P(E_S) = \dfrac{\text{Area of second prize}}{\text{Area of circle}} = \dfrac{\text{Angle of second prize}}{\text{Angle of circle}} = \dfrac{90^\circ}{360^\circ} = \dfrac{1}{4}$

b) – Let E_D be the event of getting disqualified.
- $P(E_D') = 1 - P(E_D) = 1 - \dfrac{60^\circ}{360^\circ} = \dfrac{5}{6}$

Example 2:

A basket contains x green apples, $2x$ red apples, and $x + 2$ yellow apples. An apple is picked at random.

a) Write an expression, in terms of x, for the total number of apples in the basket.

b) Write an expression, in terms of x, for the probability that the apple picked is yellow.

c) Given that the probability in part b) is $\frac{1}{3}$, find the value of x.

a) – $S = x + (2x) + (x + 2) = 4x + 2$

b) – $P(\text{picking a yellow apple}) = \dfrac{x + 2}{4x + 2}$

c) – $\frac{x+2}{4x+2} = \frac{1}{3}$
$3(x+2) = 4x+2$
$x = 4$

2 Combined Events

Calculating probabilities of multiple related events can be challenging. These tools can guide calculations based on the situation.

Probability of Combined Events

If event E is the combination of events A and B, $A \cap B$ refers to the event where both A and B happen, while $A \cup B$ refers to the event where either A or B happen.

- $P(A \cap B) = P(A \text{ and } B)$
- $P(A \cup B) = P(A \text{ or } B)$

In general, it can be challenging to calculate probabilities of combined events, but there are a few conditions and tools that can help. Let's start with independence.

Independence of Events

Two events are independent if the outcome of one does not affect the outcome of the other. For independent events A and B, the probability that both A and B occur is the product of their individual probabilities.

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

Examples

- When flipping two coins, the result of the second flip is independent of the result of the first flip.
- My breakfirst choice is independent from your exam results.

This is known as the multiplicative law of probability or the multiplication rule.

Example 1:

Two fair coins are tossed.

a) Find the probability that both are heads.

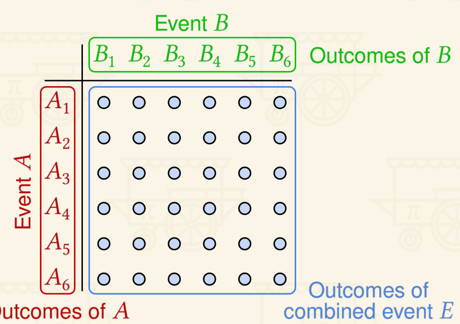
a) Find the probability that neither are heads.

a) – Let E be the event of flipping 2 heads.
Let E_1 be the event of flipping 1 head.
– $P(E) = P(E_1) \times P(E_1) = 0.5 \times 0.5 = 0.25$

b) – $P(E') = 1 - P(E) = 1 - 0.25 = 0.75$

Possibility Diagram

A possibility diagram arranges all possible outcomes of two independent events in a structured grid, allowing for easy visualization and counting of combinations.



The exact format and contents of the table depends on the types of events modeled.

Example 1:

Two fair 6-sided dice are rolled one after the other.

a) Draw a possibility diagram of all possible outcomes of the two rolls.

b) Find the probability of both dice rolling the same number.

c) Find the probability that the sum of the numbers is greater than 10.

d) Find the probability that the second number is larger than the first.

a)

Second Die \ First Die	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

b) – $P(\text{rolling the same number}) = \frac{6}{36} = \frac{1}{6}$

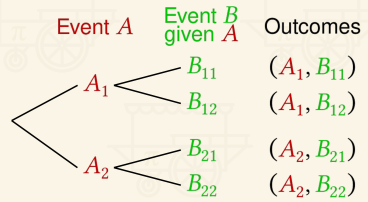
c) – $P(\text{sum being larger than } 10) = \frac{15}{36} = \frac{5}{12}$

d) – $P(\text{second number larger than the first}) = \frac{3}{36} = \frac{1}{12}$

Outcome Tree Diagram

A tree diagram shows outcomes step by step, with branches representing each possible result. It is useful when events are not independent or when there are more than two events involved.

Given an event E that is determined by the outcomes of events A and B,



To find the probability of an event E, count the number of outcomes that fulfill the event, then divide that by the total number of outcomes.

$$P(E) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}$$

Example 1:

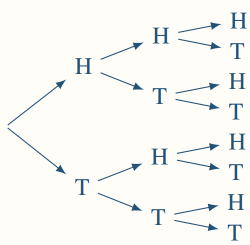
Three fair coins are tossed one after the other.

a) Draw a tree diagram of all possible outcomes of the three tosses.

b) Find the probability that there is 1 head and 2 tails.

c) Find the probability that there is at least 2 heads.

a)



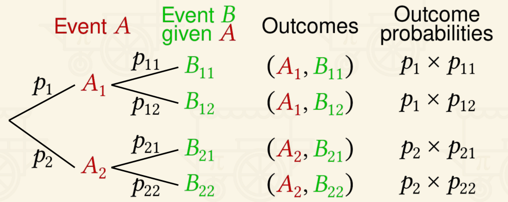
b) – $P(\text{flipping 1 head and 2 tails}) = \frac{3}{8}$

c) – $P(\text{flipping at least 2 heads}) = \frac{4}{8} = \frac{1}{2}$

Probability Tree Diagram

A probability tree diagram replaces repeated branches with weighted probabilities, making it less cluttered and easier to compute the likelihood of specific outcomes.

Given an event E that is determined by the outcomes of events A and B,



To find the probability of an event E, add the probabilities of all the outcomes that match the event. The probability of an outcome is found by multiplying the probabilities along the path.

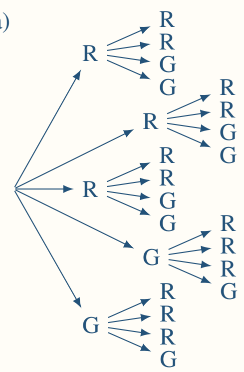
$$P(E) = \sum \text{(Favorable outcome probabilities)}$$

Example 1:

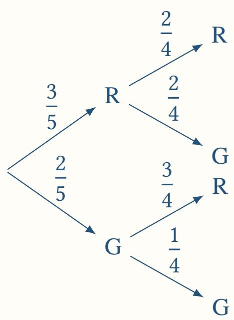
A bag contains 3 red balls and 2 green balls. Two balls are drawn from the bag one after the other, without replacement.

a) Draw an outcome tree diagram for the outcomes of the two draws.
b) Draw a probability tree diagram for outcomes of the two draws.
c) Find the probability that 1 red and 1 green ball are drawn.
d) Find the probability that both balls drawn are of the same color.

a)



b)



c) – $P(\text{1 red and 1 green ball}) = \left(\frac{3}{5} \times \frac{2}{4}\right) + \left(\frac{2}{5} \times \frac{3}{4}\right) = \frac{3}{5}$

d) – $P(\text{same color}) = \left(\frac{3}{5} \times \frac{2}{4}\right) + \left(\frac{2}{5} \times \frac{1}{4}\right) = \frac{2}{5}$

Additive Law of Probability

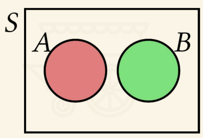
When two events can occur at the same time, simply adding their probabilities would double-count the outcomes they share. The addition law corrects this by subtracting the overlap.

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

Mutual Exclusivity of Events

Two events are mutually exclusive if they cannot happen at the same time. For mutually exclusive events A and B, the probability that either A or B occurs is simply the sum of their individual probabilities.

- $P(A \cap B) = 0$
- $P(A \cup B) = P(A) + P(B)$



Examples

- When flipping a coin, getting a head and getting a tail are mutually exclusive.
- When enrolling in a university, choosing NUS and NTU are mutually exclusive.

Example 1:

A card is drawn from a standard deck of 52 playing cards.

a) Find the probability that the card is a red king.
b) Find the probability that the card is either a spade or an ace.
c) Find the probability that the card is an ace or a king.

a) – $P(\text{red king}) = \frac{2}{52} = \frac{1}{26}$

b) – $P(\text{spade or ace}) = P(\text{spade}) + P(\text{ace}) - P(\text{ace of spade})$

$$= \frac{12}{52} + \frac{4}{52} - \frac{1}{52}$$
$$= \frac{15}{52}$$

c) – $P(\text{ace or king}) = P(\text{ace}) + P(\text{king})$ (mutually exclusive)

$$= \frac{4}{52} + \frac{4}{52}$$
$$= \frac{2}{13}$$

Example 2:

In a school of 100 students, 35 students are in the Art Club, 30 students are in the Music Club, 12 students are in both clubs.

a) Find the probability that a student is in the Art or Music Club.
b) Find the probability that a student is in only one club.
c) If 20 students are in the Drama Club and no student in the Drama Club is in the Art Club, find the probability that a student is in the Art or Drama Club.

a) – $P(\text{art or music}) = P(\text{art}) + P(\text{music}) - P(\text{both})$

$$= \frac{35}{100} + \frac{30}{100} - \frac{12}{100}$$
$$= \frac{53}{100}$$

b) – $P(\text{only one club}) = P(\text{only art}) + P(\text{only music})$

$$= \frac{35-12}{100} + \frac{30-12}{100}$$
$$= \frac{41}{100}$$

c) – $P(\text{art or drama}) = P(\text{art}) + P(\text{drama})$

$$= \frac{35}{100} + \frac{20}{100}$$
$$= \frac{55}{100}$$

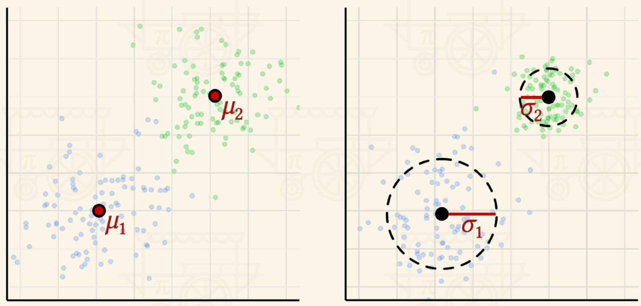
3 Statistical Measures

Besides probability theory, statistics is also about analyzing data. Statistical measures help bring focus to messy data.

Introduction to Statistical Measures

It's hard to interpret or compare large datasets directly. Statistical measures help by summarizing the data with a few key values. Measures of central tendency show the typical value, while measures of spread show how much the data varies around it.

Measure of Central Tendency   **Measure of Spread**



We'll cover 3 sets of measures of central tendency and spread.

Parametric Measures

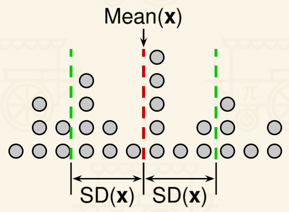
These measures are precise but sensitive to outliers and best used with symmetrical, numerical data.

Mean (central tendency)

- $\text{Mean}(\mathbf{x}) = \dfrac{\text{Sum of data points}}{\text{Number of data points}} = \dfrac{\sum_i^n x_i}{n}$
- Also represented by the symbol $\overline{x}$ (read "x bar").

Standard Deviation (spread)

- $\text{SD}(\mathbf{x}) = \sqrt{\dfrac{\sum_i^n (x_i - \overline{x})^2}{n}} = \sqrt{\dfrac{\sum_i^n x_i^2}{n} - \overline{x}^2}$



BTW: The name "parametric" comes from the use of parameters in a statistical family. The mean and standard deviation are the parameters of the normal (a.k.a. Gaussian) distribution, which is not covered in this syllabus.

Example 1:

The following dataset was collected from measuring the heights of students in a class. Find the mean height and the standard deviation.

height (m)	1.36	1.65	1.37	1.38	1.59	1.26	1.39	1.26

– Let x be the height of a student.

– $\overline{x} = \dfrac{1.36 + 1.65 + 1.37 + 1.38 + 1.59 + 1.26 + 1.39 + 1.26}{8} = 1.41$

x	1.36	1.65	1.37	1.38	1.59	1.26	1.39	1.26
x^2	1.85	2.72	1.88	1.90	2.53	1.59	1.93	1.59

– $\sum x^2 = 1.85 + 2.72 + 1.88 + 1.90 + 2.53 + 1.59 + 1.93 + 1.59 = 16.0$

– $\overline{x}^2 = 1.41^2 = 1.99$

– $\text{SD}(x) = \sqrt{\dfrac{16.0}{8} - 1.99} = 0.10$

Example 2:

A store clerk weighs 5 apples and recorded the following results.

$$48, 52, x, 50, y$$

He recorded the mean weight of the apples to be 51 grams with a standard deviation of 1.87.

a) Find the values of x and y.

b) It was later discovered that the weighing machine was miscalibrated, under-measuring each apple by 5 grams. What are the correct mean and standard deviation of the apple weights?

a) – $\dfrac{48 + 52 + x + 50 + y}{5} = 51$

$y = 105 - x$ — (1)

– $\sqrt{\dfrac{48^2 + 52^2 + x^2 + 50^2 + y^2}{5} - 51^2} = 1.87$

$48^2 + 52^2 + x^2 + 50^2 + y^2 = 13022.5$

$x^2 + y^2 = 5514.5$ — (2)

– Sub (1) into (2):

$x^2 + (105 - x)^2 = 5514.5$

$2x^2 - 210x + 5510.5 = 0$

$x = 53.5 \quad \text{or} \quad x = 51.5$

– Choosing $x = 53.5$, and sub into (1):

$y = 105 - 53.5 = 51.5$

b) – Correct Mean $= 51 + 5 = 56$

– Adding or subtracting the same number to all values doesn't change their spread, so the standard deviation stays the same.

⇒ Correct Standard Deviation $= 1.87$

Parametric Measures for Grouped Data

When data is grouped into intervals, a small adjustment to the formulas allows us to find the mean and standard deviation.

Mean (central tendency)

- $\text{Mean}(\mathbf{x}) = \dfrac{\sum_i^n f_i x_i}{\sum_i^n f_i}$

Standard Deviation (spread)

- $\text{SD}(\mathbf{x}) = \sqrt{\dfrac{\sum_i^n f_i (x_i - \overline{x})^2}{\sum_i^n f_i}} = \sqrt{\dfrac{\sum_i^n f_i x_i^2}{\sum_i^n f_i} - \overline{x}^2}$

where x is the bin value and f is the frequency of the value. For interval bins, x is the middle value of the interval.

Example 1:

A tour company offers several packages, each with its own cost and number of customers as shown below.

City	Bangkok	Tokyo	Munich	New York
Cost ($)	1200	2500	3800	4800
Customers	12	33	7	16

a) Find the mean spending of the customers.

b) Find the standard deviation of the customer spending.

x	1,200	2,500	3,800	4,800
f	12	33	7	16
fx	14,400	82,500	26,600	76,800
fx^2	17,280,000	206,250,000	101,080,000	368,640,000

a) – $\sum f = 12 + 33 + 7 + 16 = 68$

– $\sum fx = 14400 + 82500 + 26600 + 76800 = 200300$

– $\overline{x} = \dfrac{200300}{68} = 2946$

b) – $\overline{x}^2 = 2946^2 = 8678916$

– $\sum fx^2 = 17280000 + 206250000 + 101080000 + 368640000$
$= 693250000$

– $\text{SD}(x) = \sqrt{\dfrac{693250000}{68} - 8678916} = 1231$

Example 2:

For a project, a teacher told students to time their journeys to school. The following table shows the results collected.

Time (min)	$0 < x \le 20$	$20 < x \le 40$	$40 < x \le 60$	$x > 60$
Students	3	17	14	4

a) Find the mean travel time of the students.

b) Find the standard deviation of the travel time of the students.

Interval	$0 < x \le 20$	$20 < x \le 40$	$40 < x \le 60$	$60 < x \le 80$
x	10	30	50	70
f	3	17	14	4
fx	30	510	700	280
fx^2	300	15,300	35,000	19,600

a) – $\sum f = 3 + 17 + 14 + 4 = 38$

– $\sum fx = 30 + 510 + 700 + 280 = 1520$

– $\overline{x} = \dfrac{1520}{38} = 40$

b) – $\overline{x}^2 = 40^2 = 1600$

– $\sum fx^2 = 300 + 15300 + 35000 + 19600 = 70200$

– $\text{SD}(x) = \sqrt{\dfrac{70200}{38} - 1600} = 15.7$

Quartile-based Measures

These are robust tools that remain reliable even when data is skewed or contains outliers.

Quartiles

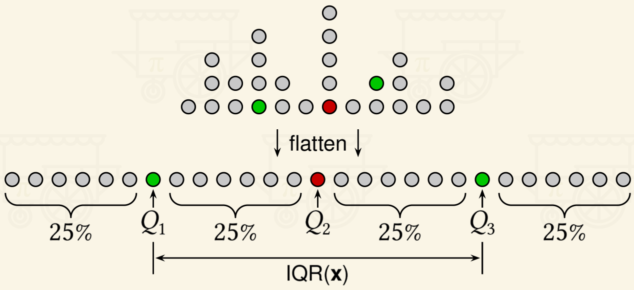
- Quartiles divide a sorted dataset into four equal parts.
- For an even number of points, a quartile is the average of the two middle number in that section.
- Lower quartile (Q_1) is the value one-quarter of the way through the data.
- Median (Q_2) is the value halfway through the data.
- Upper quartile (Q_3) is the value three-quarters of the way through the data.

Median (central tendency)

- $\text{Median}(\mathbf{x}) = Q_2$

Interquartile Range (spread)

- $\text{IQR}(\mathbf{x}) = Q_3 - Q_1$



Example 1:

An ice cream vendor records his sales over 13 days.

18, 22, 19, 25, 20, 40, 17, 24, 21, 23, 19, 26, 22

Find the quartiles and interquartile range.

- Sorted list: 17, 18, 19, 19, 20, 21, 22, 22, 23, 24, 25, 26, 40
- $Q_1 = \dfrac{19+19}{2} = 19$
- $Q_2 = 22$
- $Q_3 = \dfrac{24+25}{2} = 24.5$
- $IQR = Q_3 - Q_1 = 24.5 - 19 = 5.5$

Descriptive Measures

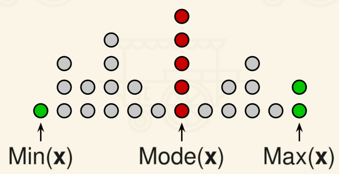
Simple and intuitive, these measures give a quick sense of the most common value and the overall spread.

Mode (central tendency)

- The mode is the value that appears most often in a dataset.

Range (spread)

- $\text{Range}(\mathbf{x}) = \text{Max}(\mathbf{x}) - \text{Min}(\mathbf{x})$



Example 1:

A restaurant offers the following meal courses. They recorded the their sales using the following table.

Courses	3-course	4-course	5-course	6-course
Cost ($)	40	55	70	85
Customers	32	19	11	5

a) Find the mode and range of the customer spending.

b) Find the median and interquartile range of the customer spending.

a) – $\text{Mode}(x) = 40$

– $\text{Range}(x) = 85 - 40 = 45$

b) – Total customers $= 32 + 19 + 11 + 5 = 67$

– Position of $Q_2 = \dfrac{67+1}{2} = 34$

$Q2 = 55$

– Position of $Q_1 = \dfrac{33+1}{2} = 12$

$Q1 = 40$

– Position of $Q_3 = 34 + 12 = 46$

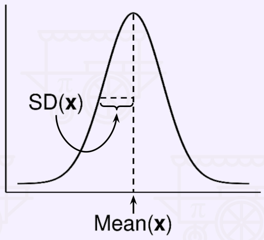
$Q3 = 70$

Why So Many Measures?

Understanding the full power of these statistical measures goes beyond the syllabus, but building some intuition can help us see why we use different types of measures in the first place.

Parametric Measures

- These are the default tools in statistics. The mean and standard deviation describe the normal distribution, which is a common model for real-world data, allowing us to do more advanced analysis and make useful predictions.



Quartile-based Measures

- Unlike parametric measures, quartiles don't rely on strong assumptions about the data distribution and are less affected by outliers or extreme values. While they may not support as many fancy calculations, they give more reliable insights when data is messy, skewed, or riddled with outliers.

$$\mathbf{x} = (1 \quad 5 \quad 2 \quad 4 \quad 1 \quad 7 \quad 9999) \quad \text{(9999: outlier)}$$

$\text{Mean}(\mathbf{x}) = 1431.3$ $\quad$ $\text{Median}(\mathbf{x}) = 4$

Descriptive Measures

- The mode is great for quick insights, especially with categorical data. For example, it's useful for answering questions like "What's the best-selling item in the store?".
- The range is the most intuitive measure of spread. Saying "the temperature ranges from 22°C to 32°C" is easy to understand, while saying "the standard deviation of temperature is 3.2" feels abstract without context.

REACH US AT:

MATHHAWKER.COM.SG – ALL NOTES

@MATHHAWKER
@MATHHAWKER
– LATEST UPDATES

PATREON.COM/MATHHAWKER – SUPPORT US :)