# Material Summary: Probability and Combinatorics

## Probability

**1.1 Some Definitions**

* The scientific method relies on **experiments**
  + Initial conditions outcome
    - Usually, we control the initial conditions to isolate the outcome
* **Random event**
  + A set of outcomes of an experiment
  + Each outcome happens with a certain **probability**
* **Random variable**
  + An expression whose value is the outcome of the experiment
  + Usually denoted with , , … (capital letters)
* **It is not possible to predict the next outcome of a   
  random event!**
  + But we can perform the same experiment **many times** (trials)
  + The patterns and laws become more apparent with more trials

**1.2 Frequency**

* Let's perform the same experiment many times
  + Under the same conditions
  + … such as throwing a dice
* A black background with a black square

  Description automatically generated with medium confidenceAssign a frequency to each number that the dice shows
  + – number of trials we got , – all trials
* As increases, "stabilizes" around some number
* We cannot perform infinitely many experiments
  + A black background with a black square

    Description automatically generated with medium confidenceBut we can "extend" the trials mathematically
  + We call this the probability of outcome **A**
    - **Statistical definition** of probability

**1.3 Examples**

* Rolling a dice
  + Possible outcomes:
  + A black background with a black square

    Description automatically generated with medium confidence**Fair dice** – all outcomes are equally likely
* Tossing a fair coin
  + A black background with a black square

    Description automatically generated with medium confidencePossible outcomes:
* A black background with a black square

  Description automatically generated with medium confidenceTossing an unfair coin
* Note that
  + The probability
    - It can also be expressed as percentage:
  + The sum of all probabilities is equal to

**1.4 Countable and Uncountable Outcomes**

* In some cases, the number of outcomes is finite
* But some random variables have **infinitely many** outcomes
* Example: intervals
  + What is the probability that a real number chosen at random, is in the interval ?
  + A black background with a black square

    Description automatically generated with medium confidenceAnswer: it depends only on the lengths of the intervals
    - The number of outcomes in infinite, but we are still able to   
      compute probabilities
  + A black background with a black square

    Description automatically generated with medium confidence**Probability density** – the probability of the result being in a tiny interval
    - – both ends of the interval

**1.5 Visualizing Random Variables**

* To visualize a random variable, we plot the value as a function  
  of the trial number
  + We can generate random values using **numpy**
  + A screenshot of a computer

    Description automatically generatedExample: throwing a dice

**1.6 Visualizing Random Variables**

* The function we got is not very informative
  + Better way: show the frequency of each output
    - For each possible value of the random variable, count how many times we got that value
  + A screenshot of a computer

    Description automatically generatedThis is called a **histogram**

## Combinatorics

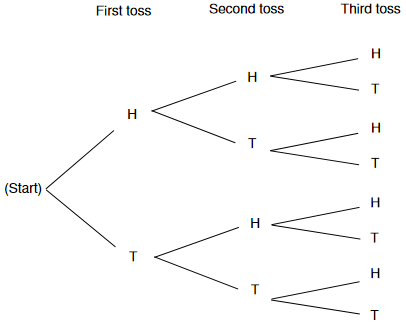
**2.1 Combinatorics**

* Combinatorics deals with **counting** objects and groups of objects
* Prerequisites
  + Finite (countable) number of outcomes
  + All outcomes have equal probability
* Examples: gambling games
  + Roulette – all segments are equally likely
  + Card games – all card backs are the same
* Counting rules
  + Rules for computing a **combinatorial probability**
  + Show how many "desired" outcomes exist
* Notation
  + All outcomes:
  + All experiment outcomes:
    - Usually, is fixed and depends on the experiment
* Types of **samples**
  + with repetition / without repetition
  + ordered / unordered
* Example: taking numbered balls out of a box
  + Take a ball, then return it to the box
  + Take a ball without returning it to the box (in this case )
  + Take balls in a specific order (e.g., if they are numbered or colored)
  + Take balls in no specific order

**2.2 Counting Rules**

* **Rule of sum**
  + choices for one action, choices for another action
  + The two actions **cannot be done at the same time**
  + There are ways to choose one of these actions
* Example
  + A woman will shop at **one** store in town today
    - North part of town – mall, furniture, jewellery (3 stores)
    - South part of town – clothing, shoes (2 stores)
  + In how many ways she could visit one shop?
  + Answer: 3 + 2 = 5 ways
* **Rule of product**
  + choices for one action, choices for another action
  + The two actions are performed **one after the other**
  + There are ways to do both actions
* Example
  + You have to decide what to wear
    - Shirts – red, blue, purple (3 colors)
    - Pants – black, white (2 colors)
  + In how many ways can you create one outfit (shirt and pants)?
  + Answer: 3.2 = 6 ways
    - For each choice of shirt, you can choose one color of pants
    - These are **independent**

**2.3 Example: Three Coin Tosses**

* Let's explore a graphic method of solving combinatorial problems called a **tree diagram**
  + Draw all intermediate results and the links between them
  + A "path" through the tree represents an outcome
  + Useful when the outcomes are relatively few
* What's the probability of getting 3 tails out of 3 coin tosses?
  + Answer:
* What's the probability that both of these   
  are true?
  + The first outcome is a head
  + The second outcome is a tail
  + Answer:

**A black background with a black square

Description automatically generated with medium confidence2.4 Example 2: Eating at a Restaurant**

* A restaurant offers
  + 5 choices of appetizer
  + 10 choices of main course
  + 4 choices of dessert
* You can choose one course, two **different** courses, or all three
* How many possible meals can you make?
* One course: either appetizer, main course, or dessert:
* Two courses: total
* Appetizer + main course:
* Main course + dessert:
* Appetizer + dessert:
* Three courses:
* Total: possible meals

**2.5 Permutations**

* A permutation (without repetition) of a set is any shuffling of all   
  elements in
  + The order matters
  + Notation:
* Example:
  + If , some permutations are
* Number of permutations of elements
  + choices for the first element
  + for the second one
    - Because the first one is already taken
  + for the third one
  + for the last one
  + Total:

**2.6 Variations**

* A variation is an ordered subset of elements from A
* Notation:
  + We read this as "Variations of elements, th class"
* Example:
  + If , some variations are
* Number of variations
  + **A black background with a black square

    Description automatically generated with medium confidence**Same technique as in permutations
  + choices for the first element
  + for the second one
  + for the last one

**2.7 Combinations**

* A combination is an **unordered subset** of k elements from A
* Notation:
* Example:
  + If , some combinations are
* Number of combinations of n elements
  + A black background with a black square

    Description automatically generated with medium confidenceUsing a similar (but more involved) logic, we can prove that
  + A black background with a black square

    Description automatically generated with medium confidenceThis is also known as **"n choose k"** (we choose elements from )

**2.8 Example Usages**

* Shuffle a deck of cards
  + The same as generating a random permutation of 52 (or 54) elements
* Crack a password
  + How many 3-letter passwords are there ( letters total)?
* Generate all anagrams of a given word
  + Anagram: a different word using the same letters
    - Example: emits items, mites, smite, times
  + Method:
    - Generate all permutations of the letters
    - For each permutation, find whether it’s a valid word (check with a dictionary)
    - Return all valid words
* Make a fruit salad
  + Generate combinations of fruits (the order doesn’t matter)
    - Possibly, combinations with repetition (if I love bananas, I’ll take a double serving)

## Probability Algebra

**3.1 Events**

* **Event** – a result from the experiment
* **Elementary event**
  + One particular outcome
  + Example: outcomes of two coin flips:
* Compound event
  + Consists of many elementary events
  + Example: getting an odd number from a dice
    - Consists of the elementary events
* **Event space** – the set of all possible events
* The algebra of events is the same as the algebra of sets
  + … and we already know these :)

**3.2 Algebra of Events**

* If event A happens with event B, A is a consequence of B:
* If and , then
* Complementary event: happens iff does not happen
* Impossible event: contains no elementary events:
* Product of events: happens iff A and B happen:
  + Incompatible events:
* Sum of events: happens if A, B or both happen:
  + If A and B are incompatible,
* Observe that
  + Logical relations are the same as set operations (and event operations)
    - AND: intersection
    - OR: union
    - NOT: complement

**3.3 Conditional Probability**

* Additional information about the experiment outcome can change   
  the probabilities
  + "a priori" "a posteriori"
* Example:
  + "Hidden dice": someone rolls a dice and doesn't tell us the result
  + Probabilities: for every number
    - These are also called "a priori" probabilities
  + Now we know the number is even
    - This changes all outcome probabilities:
      * These are called "a posteriori" probabilities
* Conditional probability
  + Probability of event given event
  + Math notation:
* More formally
  + A black background with a black square

    Description automatically generated with medium confidenceIf we know happened, the probability corresponds to the part of the set which is shared between and
* In our example
  + Event : number on a fair dice
  + A white circles with green and blue circles

    Description automatically generatedEvent : the number is even

**3.4 Event Independence**

* Sometimes, an event doesn't influence another event
  + They are called independent events
* If two events are independent, knowledge of one does not tell us anything about the other
* More formally,
  + If ,
  + The same can be applied to if
* Example
  + 99% of all people who died of cancer, have consumed pickles
  + 99,8% of all soldiers have eaten pickles
    - [*http://www.pleacher.com/mp/mhumor/pickles.html*](http://www.pleacher.com/mp/mhumor/pickles.html)
  + [*http://www.dhmo.org/facts.html*](http://www.dhmo.org/facts.html)

**3.5 Bayes' Theorem**

* The theorem tells us how to update the probabilities when we know some evidence
* A black background with a black square

  Description automatically generated with medium confidenceA black background with a black square

  Description automatically generated with medium confidenceA black background with a black square

  Description automatically generated with medium confidenceExample usage: spam detection

A black background with a black square

Description automatically generated with medium confidence

* + Consider each word ; compute number of emails which contain it
    - spam emails containing ; total emails containing :
    - "Spamminess" of word: frequency
    - "Spamminess" of email:

**3.6 Example: Family Paradox 1**

* A family has two children
  + One of them is a boy
  + What is the probability that both children are boys?
    - A child has a 0,5 chance of being a boy or a girl
* Intuitive answer: 0,25
  + **A diagram of a child

    Description automatically generatedA diagram of triangles with text

    Description automatically generated with medium confidence**But wait… let's exhaust all possibilities

**3.7 Example: Family Paradox 2**

* Mr. Smith is the father of two children
  + When we meet him on the street, he introduces one as his son
  + What's the probability that the other child is a boy?
* Assumption
  + **A diagram of a person walking with his name

    Description automatically generated with medium confidenceA diagram of a family tree

    Description automatically generated**He is equally likely to take any child to a walk

**3.8** Example: Monty Hall Problem

* In a game show, you have to choose between three doors
  + Behind one is a car, behind the other two – goats
* You pick a door
* The host reveals one of the two other doors – it's always a goat
* You have the option to keep your choice or switch doors
  + Which is the winning strategy?
* It turns out that the winning strategy  
  is to always switch
  + This gives you chanceof winning   
    the car
* **A diagram of a donkey door

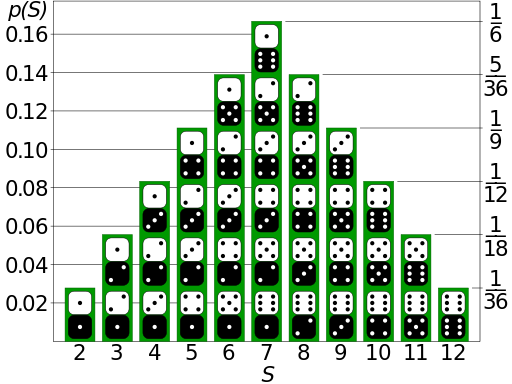
  Description automatically generated**More details: [Quora](https://www.quora.com/What-is-the-Monty-Hall-problem-and-what-is-its-solution)

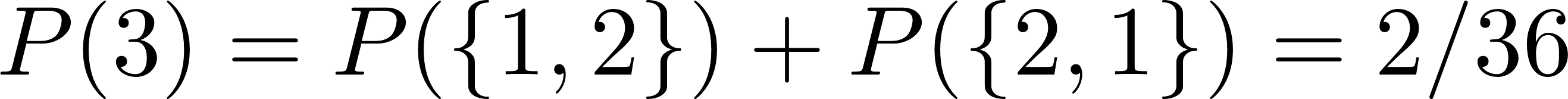
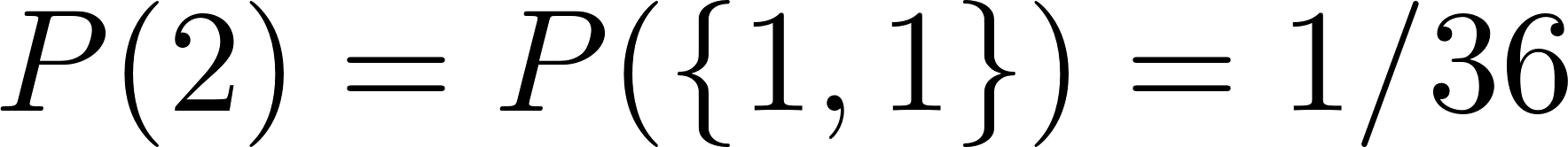
## Statistical Distributions

**4.1 Distributions**

* We saw that random variables can be treated as functions
  + But they look funky
    - Don't have derivatives at most points
    - Difficult to work with
* We can instead take functions of these functions
  + Like we counted each outcome
    - Instead of graphing the real function, we made a histogram of counts
    - This gives us a much better idea what the random variable looks like
* These functions of functions are called **distributions**
  + In our example, we looked at the **frequency distribution**

**4.2 Discrete Distribution**

* Probability distribution function
  + A table which maps each outcome of a random variable  
    to a probability:
  + Also called **probability mass function** (pmf)
* Example: two die rolls
  + Random variable: sum of numbers
  + Outcomes:
  + Probabilities

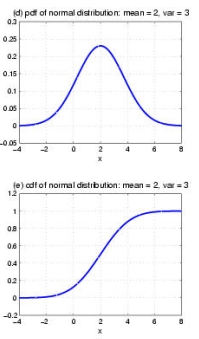


* **Cumulative distribution function**
  + A black background with a black square

    Description automatically generated with medium confidenceA table which maps each outcome of a random variable to the probability of its value being less than or equal to a given number
  + Also called **cumulative mass function** (cmf) or **cumulative density function** (cdf)
  + Every cmf is non-decreasing
    - Usually starts at 0
    - A graph of a graph of a graph

      Description automatically generated with medium confidenceAlways ends at 1

**4.3 Continuous Distribution**

* Cumulative density function (cdf)
  + ****A black background with a black square

    Description automatically generated with medium confidenceDefined in the same way as the cmf:
  + Probability density function
  + A black background with a black square

    Description automatically generated with medium confidenceDerivative of the cdf:
  + Meaning: the probability of the function taking values in an infinitely small interval around
  + The probability of observing any single value is exactly 0
    - The number of outcomes is

## Common Distributions

**5.1 Bernoulli and Uniform Distributions**

* Bernoulli distribution
  + The simplest distribution of a random variable
    - Value with probability
    - Value with probability
  + The two events are incompatible (mutually exclusive)
  + Example: coin flip (fair coin: )
  + … Not so interesting on its own
    - But takes part in other distributions
* Uniform distribution
  + All values in some range are equally likely
  + Example: number on a fair dice
    - Also generalizes to continuous variables

**5.2 Binomial Distribution**

* Bernoulli trials
  + Each trial has a "success" probability
  + Bernoulli distribution
* Discrete distribution
* Notation:
  + "X follows the binomial distribution with parameters and "
* A black background with a black square

  Description automatically generated with medium confidenceProbability mass function
* Cumulative function

A black background with a black square

Description automatically generated with medium confidence

**5.3 Normal Distribution**

* Origin: random errors in measurements
  + When we perform an experiment, there are many sources of error
* Example: throwing a dart at the origin of the -plane
  + We aim at the origin
  + Random errors prevent us from hitting it every time
  + Sources of error
    - Hand shaking, air currents, distribution of mass inside the arrow, different viewing angles… and many more, some of which we can’t even know
* Assumptions
  + The errors don’t depend on the orientation of the coordinate system
  + The errors in and directions are independent: one doesn't influence the other
  + Large errors are less likely than small errors
* We can derive the distribution of errors
  + Distances from the origin
* Normal (Gaussian) distribution
  + pdf:
  + – parameters
    - We'll see their real meaning next time
  + cdf: doesn't exist as a function, we can integrate numerically
* Complete derivation of the formula: [*here*](http://courses.ncssm.edu/math/Talks/PDFS/normal.pdf)
* **Standard normal distribution:** 
  + A graph of a function

    Description automatically generatedMainly for convenience

**5.3 Central Limit Theorem**

* The sum of many independent random variables tends to a normal distribution even if the original random variables are not normally distributed
  + In other words: The sampling distribution of the mean of any independent random variable will be normal or nearly normal if the sample is large enough
  + Large enough?
    - for most statisticians, but more is better
* Example: customers in a shop
  + Every customer has their own behavior, reasons, money, etc.
    - We can treat them as random variables with unknown distributions
  + The shop's earnings are approximately normally distributed
    - If there are many customers
  + We **don't even care** about the many sources of error: they will produce a **normal distribution** anyway