1 Introduction

What does the term "probable" mean to you ?

How likely is it that...
- you will win the lottery this week?
- in a room of 23 people, at least two people share the same birthday?
- an insured driver will make a claim?
- the next outcome in a time series will be higher than the previous outcome?
- the evidence presented at a jury trial supports guilty verdict if the accused was actually innocent?

'When it is not in our power to determine what is true,
we must ascertain that which is most probable'
René Descartes

In order to learn from data, to make predictions
- we need to take into account randomness
- assign a probabilistic model to real-life phenomena
- we need to measure how probable events are likely to be

Some basic definitions ....

An **Experiment** -> anything that gives rise to a defined set of possible outcomes
A **Sample point** -> one basic single outcome, s.
A **Sample Space** -> set of all possible outcomes (i.e. sample points) $S = \{$ all s possible in an experiment$\}$.

Example:
a) Experiment = Toss a coin observe upper face. $S = \{H, T\}$
b) Experiment = Toss a coin twice and observe upper faces: $S = \{HH, HT, TH, TT\}$

Task 1: Write down the sample space for the following experiments:
a) Throw a die observe number on uppermost face
b) Throw two die observe 2 numbers thrown
c) Pick a card from a deck of 52 cards observe card
d) Pick a card from a deck of 52 cards observe card suit
e) Toss a coin until a head appears
f) Lifetime of battery.

Solution:

You will notice that sometimes specification of the sample space can be subjective to the experimenter !!

Types of sample space:
Finite sample spaces,
countable = one-to-one match with whole numbers,
discrete = finite or infinitely countable
Continuous sample spaces.

Task 2: Pick out examples of these types of sample spaces from the sample spaces you specified in the previous task.

Solution:

An **Event** -> any subset of sample points in the sample space, $A \subseteq S$.
Usually capital letters at the beginning of the alphabet are used to represent events, e.g. $A, B, C, .., E, ..A_1, A_2$.
Let $\mathcal{A}$ represent the collection of *all* events in a sample space, $\mathcal{A} = \{\text{all } A_i | A_i \subseteq S\}$.

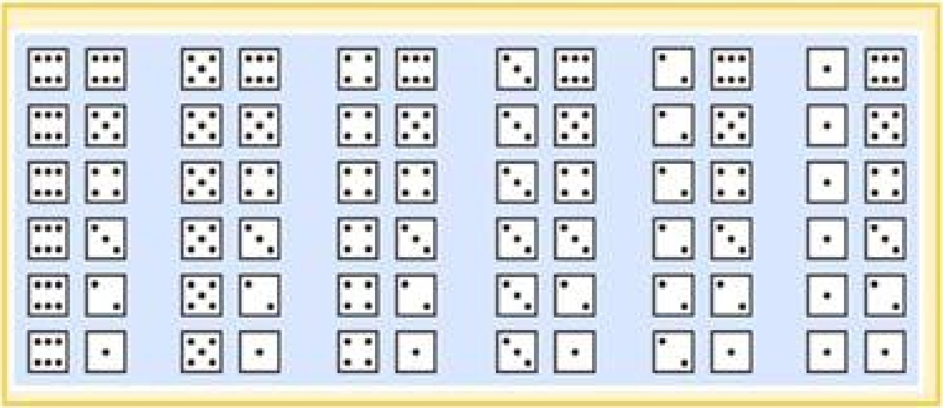
Example:
a) Experiment: Toss a coin twice observe upper faces,
  i) A = exactly one head occurs,
  ii) B = at least one head occurs.
b) Experiment: Toss two dice, Event E = sum of numbers equals 4.

Solution:
a) i) $A = \{HT, TH\}$, ii) $B = \{HT, TH, HH\}$
b) $E = \{(1, 3), (3, 1), (2, 2)\}$

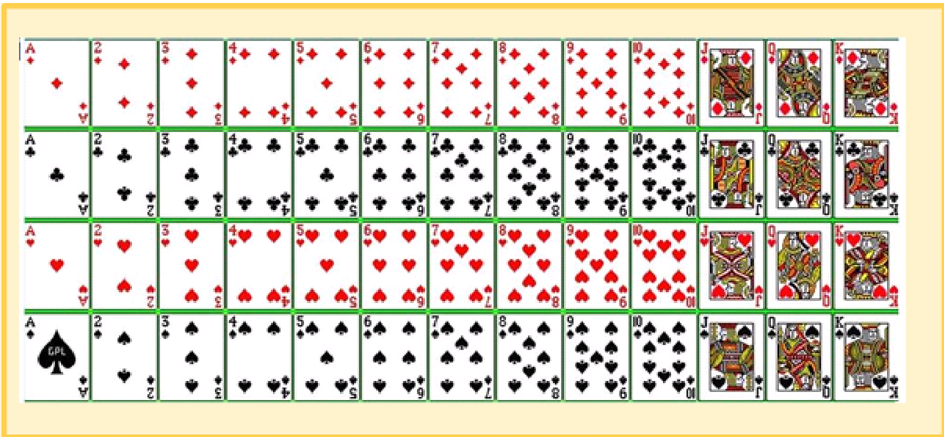
Task 3: List the sample points in the following events:
a) Throw a die, E = the number four or higher comes up
b) Throw two die, E = the sum of the two numbers is 7
c) Pick a card from a deck of 52 cards, E = a picture card is selected
d) Pick a card from a deck of 52 cards, E = a spade is selected
e) Toss a coin until a head appears, E = head appears within 5 tosses
f) E = lifetime of a battery is greater than 400 hours.

Solution:

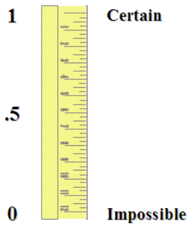
2 dice :



face:



2 Finite Sample Spaces and probability functions

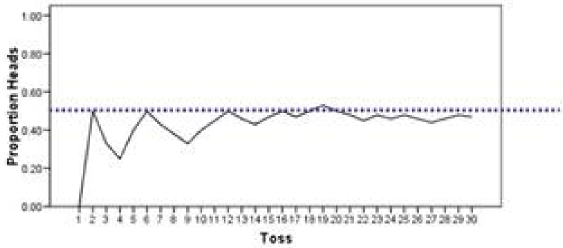


The **Probability** of an event A, represented by $P(A)$ or $Pr(A)$, represents the likelihood of the event occurring on a scale from 0, impossible, to 1, certain, i.e. $P(A) \in [0, 1]$

The probability of an event is its *long run relative frequency.*

A **Probability Distribution** outlines the probability of all events in the sample space, $P : \mathcal{A} \to [0, 1]$.

Task 4: Take a coin from your pocket. Starting from toss 1 and continuing tosses, calculate the proportion of heads you tossed, graph these relative frequencies against the number of tosses. What do you expect to happen as the number of tosses increases?



Experiment: Toss a coin. Let Event A= observe a head, $P(A) = 0.5$.
$S = \{H, T\}$.
$\mathcal{A} = \{H, T, HorT\}$??
The probability distribution:

Outcome:	H	T
Prob:	0.5	0.5

Postulates of probability:

1. The probability of an event is a nonnegative real number, that is, $P(A) \geq 0$ for any subset $A \subseteq S$.
2. $P(S) = 1$, since some outcome in the sample space must occur with certainty.
3. If $A_1, A_2, A_3, \cdots$, is a finite or infinite sequence of mutually exclusive events of S, then $\sum_{i=1}^{n} A_i = 1$

Probability of sample points: For a sample space with a finite # of outcomes/sample points, $S = \{s_1, ..., s_n\}$, define $p_i = P(s_i)$ as the probability function, where $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$.
If all outcomes, $\{s_1, ..., s_n\}$, have equal probabilities $p_1 = p_2 = \cdots = p_n$, then $p_i = \frac{1}{n}$.

Example:Write down the probabilities of these sample points:
a) Experiment: Toss a coin observe upper face, $P(H) =?$
b) Experiment: Throw a die observe upper face, $P(4) =?$
Solution: a)$P(H) = \frac{1}{2}$, b) $P(4) = \frac{1}{6}$

Probability of events:

$$P(A) = \frac{\#\{s|s \in A\}}{\#\{s|s \in S\}} = \frac{\#(A)}{\#(S)}$$

Example:Write down the probabilities of these events:

a) Experiment: Toss a coin twice observe upper faces,

i) A = exactly one head occurs,

ii) B = at least one head occurs.

b) Experiment: Toss two dice, Event E = sum of numbers equals 4.

Solution:

a) $S = \{HH, HT, TH, TT\}, \#S = 4.$

$A = \{HT, TH\}, P(A) = \frac{2}{4},$

$B = \{HT, TH, HH\}, P(B) = \frac{3}{4}$

b) $S = \{(i, j), i, j \in 1, 2, 3, 4, 5, 6\}, \#S = 36.$

$E = \{(1, 3), (3, 1), (2, 2)\}, P(E) = \frac{3}{36}$

Task 5: For the events from experiments with finite sample spaces that you described in Task 2, write down the probabilities for those events.

Solution:

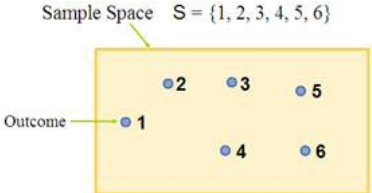
** If continuous, can't talk about $P(outcome)$, need to consider $P(set)$

Example: $S = [0, 1], 0 < a < b < 1$. $P([a, b]) = b - a, P(a) = P(b) = 0$. Need to group outcomes, not sum up individual points since they all have P = 0.

Sample Space: Venn Diagram representation

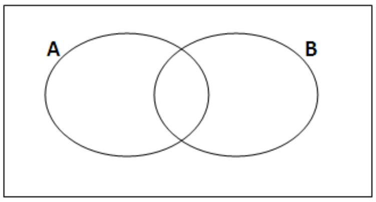
Finite sample spaces can be visualized by using a venn diagram.

For example:

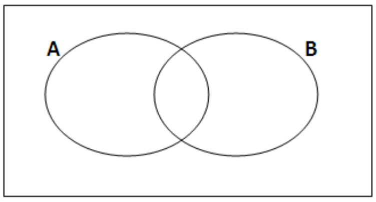


Definitions of operators:

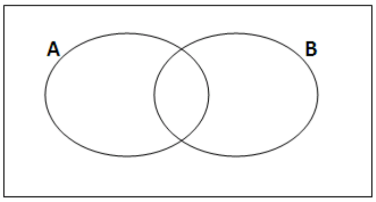
Task 6: Use venn diagrams to describe the following operators:
Union of Sets: $A \cup B = \{s \in S : s \in A \text{ or } s \in B\}$.



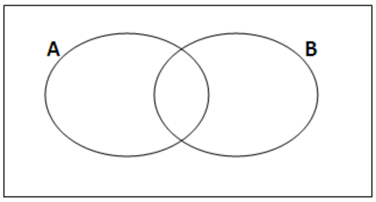
Intersection: $A \cap B = \{s \in S : s \in A \text{ and } s \in B\}$.



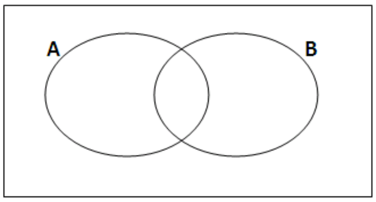
Complement: $A^c = \{s \in S : s \notin A\}$ also can be represented by $A^c = A' = \overline{A}$.



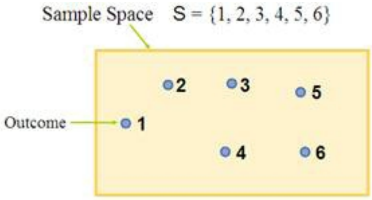
Set Difference: $A \backslash B = \{s \in S : s \in A \text{ and } s \notin B\}$
* $A \backslash B$ "A *less* B" is not to be confused with $A|B$ "A *given* B" ..which will be seen later.
$A \backslash B$ is usually written as $A \cap B^c$.



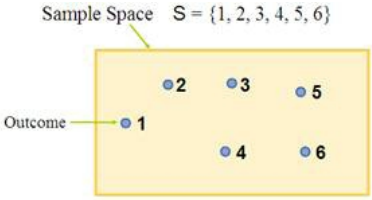
Symmetric Difference:
$A \triangle B = \{s \in S : (s \in A \text{ and } s \notin B) \text{ or } (s \in B \text{ and } s \notin A)\} = (A \cap B^c) \cup (B \cap A^c)$



Mutually Exclusive / Disjoint Events:



Non Mutually Exclusive / Disjoint Events:



2.1 Some properties:

- $A \cup B = B \cup A$
- $(A \cup B) \cup C = A \cup (B \cup C)$
- $A \cap B = B \cap A$
- $(A \cap B) \cap C = A \cap (B \cap C)$
- mixed operations $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
- $(A \cup B)^c = A^c \cap B^c$
- $(A \cap B)^c = A^c \cup B^c$ since
  $s \in (A \cap B)^c = s \notin (A \cap B)$
  $s \notin A$ or $s \notin B = s \in A^c$ or $s \in B^c$
  $s \in A^c \cup B^c)$

Task 7: Use Venn diagrams to help verify the set properties outlined above.

Solution:

3 Learning to count...

Motivating example:

What is the probability you will win the lotto by playing 1 line of 6 numbers chosen randomly from 1 to 45 inclusive ?

So evaluating probabilities, - how probable an event is likely to happen, is all about counting outcomes... So we must learn how to count ....
Aim:

- Solve counting problems using the Multiplication Principle
- Solve counting problems using permutations
- Solve counting problems using combinations
- Solve counting problems using permutations with repetition
- Solve counting problems with restrictions
- Compute probabilities involving permutations and combinations

3.1 Solve counting problems using the Multiplication Principle

Example: Menu.
The fixed-price dinner at a restaurant provides the following choices...

Appetizer: Soup or Salad
Main Course: Baked chicken,
Broiled beef patty,
Baby beef liver,
or Roast beef
Dessert: Ice-cream or Cheese cake

How many different three course meals can be ordered?
Solution: Appetizer: 2 choices, Main Course: 4 choices, Dessert: 2 choices.
i.e. 2 choices and 4 choices and 2 choices
Then $2 \times 4 \times 2 = 16$ different ways to order a three course meal.

The multiplication principle :
If a task consists of a sequence of choices in which there are
p selections for the first choice,
q selections for the second choice,
r selections for the third choice,
and so on,
then the task of making these selections can be done in $p \times q \times r$.. different ways.

Task 8: Airports
The IATA, the International Airline Transportation Association assigns three-letter codes to represent airport locations. For example, the airport code for Dublin Airport is DUB. How many different airport codes are possible?

Solution:

Task 9: postal delivery
You have just been hired as a Post Delivery person for NUI Galway. On your first day, you must travel to seven buildings with letters. How many different routes are possible?

Solution:

3.2 Solve counting problems using permutations

A permutation is equivalent to counting without replacement.

Permutations :
A permutation is an arrangement of objects.

We have seen that arranging n distinct (different) objects can be done in $n(n-1)(n-2)...3.2.1$ different ways. This calculation is often written using the factorial symbol.
If n is an integer, the factorial symbol $n!$ is defined as $n! = n(n-1)(n-2)...3.2.1$
Eg: $3! = 3.2.1 = 6$ Eg: $2! = 2.1 = 2$
Note: There is only one way to arrange one item, $1! = 1$ and there is only 1 way to arrange no items, $0! = 1$.

Example:postal delivery - revisited..
How many routes to 7 buildings is equivalent to how many arrangements of the 7 buildings,
Solution: 7!=5040 routes.

Task 10:Airports - revisited...
How many different airport codes are possible if the same letter cannot be used more than once in the code?

Solution:

Example: Committee Problem
Three members from a 14-member committee are to be randomly selected to serve as chair, vice chair, and secretary. The first person selected is the chair, the second person selected is to be vice chair, and the third secretary. How many different committee structures are possible?
Solution: $14 \times 13 \times 12 = 2,184$

A **permutation** can also be an arrangement of r objects chosen from n distinct (different) objects where replacement in the selection is not allowed.
The symbol P_r^n represents the number of permutations of r objects selected from n objects.
The calculation is given by the formula:

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

Example: Committee Problem - revisited..
Write the solution to the committee problem in terms of permutation notation.
Solution: $P_3^{14} = \frac{14!}{11!} = 14.13.12 =$

Task 11:Airports - revisited...
How many different airport codes are possible if the same letter cannot be used more than once in the code?

Solution:

3.3 Solve counting problems using combinations

Task 12: Choosing teams
Peter, Mike, Rick and Jay are going to play golf. They will randomly select teams of two players. How many ways can the teams be chosen?
List all possible team combinations. That is, list all the combinations of the four people Peter, Mike, Rick and Jay taking two at a time.
Is the arrangement of players within a team important?

Solution:

A **Combination** is a collection, without regard of order, of n distinct objects without replacement. The symbol C_r^n represents the number of combinations of r objects chosen from n distinct objects, without arrangement. The calculation is given by the formula:

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

The symbols C_r^n and $\binom{n}{r}$ are equivalent. $\binom{n}{r}$ is referred to as a ***binomial coefficient***.

Example: Choosing teams - Revisited..
The solution could be written as: "from 4 choose 2" without arranging the 2 players or = ways of choosing 2 players from 4.
Solution: $C_2^4 = \binom{4}{2} = \frac{4!}{2!2!} = 6$

Task 13: Probabilistic Sampling
How many different simple random samples of size 4 can be obtained from a population whose size is 20?

Solution:

3.4 Solve counting problems using permutations with repetition

Example: WORDS.
How many distinguishable strings of letters can be formed by using all the letters in the word *REARRANGE* ?

Solution: Arranging 9 distinct letters is 9! = 362,880.
But the letters are not distinct... There are 3 R's , 2 A's, 2 E's.
Let the three R's be R_1, R_2 and R_3.
The word $R_1EAR_2R_3ANGE$ is the same as $R_2EAR_3R_1ANGE$.
There are 3!=6 ways to arrange the R_1, R_2 and R_3.
Therefore this arrangement is repeated 6 times in the count 9!=362,880.
To take this into account calculate $\frac{9!}{3!}$.
But there are also 2 A's, 2 E's to take into account,

$$\frac{9!}{3!2!2!}$$

The number of permutations of n of which
n_1 are of one kind,
n_2 are of a second kind,
...,
and n_k are of a kth kind
is given by

$$\frac{n!}{n_1!n_2!..n_k!}$$

where $n_1 + n_2... + n_k = n$.
The symbol $\binom{n}{n_1,n_2,...,n_k}$ is often used to describe this calculation and is referred to as a ***multinomial coefficient***

Task 14: Flags.
How many different vertical arrangements are there of 10 flags if 5 are white, 3 are blue and 2 are red?

Solution:

3.5 Solve counting problems with restrictions

Example: Schoolboys.

There are 8 boys in a class on their first day at school.

a) How many ways can they be arranged in assembly if they stand in a line?
b) If two of the boys are twins and must be kept together how many ways can they be arranged?
c) If the twins always fight how many ways can they be arranged if the twins must always be kept separate?

Solution:

Example: pincodes.

How many different pincodes are there? If all four numbers must be pressed down at the same time - how many arrangements are there?

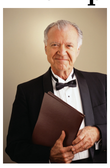
Solution:

3.6 Compute probabilities involving permutations and combinations

Recall :
For finite sample spaces the probability of an event is

$$P(A) = \frac{\#\{s|s \in A\}}{\#\{s|s \in S\}} = \frac{\#(A)}{\#(S)}$$

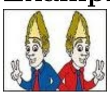
Example:



If all course menu options are equally likely to be chosen,
a) what is the probability that at random I will choose soup for my appetizer, baked chicken for my main course and cheese cake for dessert?
b) what is the probability of having soup, a beef main course and cheese cake?

Solution:

Example:



If the 8 boys stand in assembly in random order, what is the probability that the teacher will have to break up a fight.

Solution:

Example:

a) What is the probability you will win the lotto by playing 1 line of 6 numbers chosen randomly from 1 to 45 inclusive ?
b) What is the probability of winning the Irish Lotto with a 3 euro bet?

Solution:

***for more lotto problems see the section on the Hypergeometric distribution.*

4 Mutually Exclusive/Disjoint Events

Two events A and B are mutually exclusive/disjoint if $A \cap B = \emptyset$, i.e. $P(A \cap B) = 0$.

Example:Are these pairs of events mutually exclusive?
a) Experiment: Throw a die record number on face, $A = \{1, 2, 3, 4\}$ and event $B = \{5\}$.
b) Experiment: Throw a die record number on face, $A = \{3 \text{ or } 4\}$ and event $B = \{\leq 5\}$

Solution:
a) Yes, since $A \cap B = \emptyset$, $P(A \cap B) = 0$.
b) No, since $A \cap B = \{3, 4\}$, $P(A \cap B) = \frac{2}{6}$.

Task 15:Which of the following pairs of events are mutually exclusive?
a) Experiment: Throw a die record number on face, $A = \{\leq 2\}$, $B = \{> 2\}$
b) Experiment: Draw a card from pack 52, $A = redcard$, $B = blackcard$
c) Experiment: Draw a card from pack 52, $A = numbercard$, $B = Jack$
d) Experiment: Throw a die record number on face, $A = \{\leq 2\}$, $B = \{\geq 2\}$
e) Experiment: Draw a card from pack 52, $A = redcard$, $B = heart$
f) Experiment: Draw a card from pack 52, $A = redcard$, $B = Queen$

Solution:

5 Properties of Probability:

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. For disjoint (mutually exclusive) events A, B, i.e. where $\rightarrow A \cap B = \emptyset$,
   $P(A \text{ or } B) = P(A) + P(B)$.
   This can be written for any number of events.
   For a sequence of events $A_1, \cdots, A_n, \cdots$, all disjoint, $(A_i \cap A_j = \emptyset, i \neq j)$:
   $$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$
   which is called "countably additive".
4. $P(\emptyset) = 0$.
5. $P(\overline{A}) = 1 - P(A)$
6. If $A \subseteq B$ then $P(A) \leq P(B)$.

6 The Additive Rule - Union of events

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

To see this:

Example:Find $P(A \cup B)$ for each of the following pairs of events.
a) Experiment: Throw a die record number on face, $A = \{1, 2, 3, 4\}$ and event $B = \{5\}$.
b) Experiment: Throw a die record number on face, $A = \{3 \text{ or } 4\}$ and event $B = \{\leq 5\}$
Solution:
a) $P(A) = \frac{4}{6}$, $P(B) = \frac{1}{6}$, Since $A \cap B = \emptyset$, $P(A \cap B) = 0$,
$P(A \cup B) = \frac{4}{6} + \frac{1}{6} - 0 = \frac{5}{6}$.
b) $P(A) = \frac{2}{6}$, $P(B) = \frac{5}{6}$, since $A \cap B = \{3, 4\}$, $P(A \cap B) = \frac{2}{6}$,
$P(A \cup B) = \frac{2}{6} + \frac{5}{6} - \frac{2}{6} = \frac{5}{6}$.

Task 16:Find $P(A \cup B)$ for each of the following pairs of events.
a) Experiment: Throw a die record number on face, $A = \{\leq 2\}$, $B = \{> 2\}$
b) Experiment: Draw a card from pack 52, $A = \mathit{redcard}$, $B = \mathit{blackcard}$
c) Experiment: Draw a card from pack 52, $A = \mathit{numbercard}$, $B = \mathit{Jack}$
d) Experiment: Throw a die record number on face, $A = \{\leq 2\}$, $B = \{\geq 2\}$
e) Experiment: Draw a card from pack 52, $A = \mathit{redcard}$, $B = \mathit{heart}$
f) Experiment: Draw a card from pack 52, $A = \mathit{redcard}$, $B = \mathit{Queen}$

Solution:

Extension : union of three events....
$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$

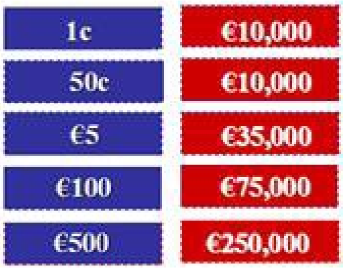
Extension : union of n events....
Theorem:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} P(A_1 \cap \cdots \cap A_n)$$

Example: Matching problem:
You have 4 letters and 4 envelopes, randomly stuff the letters into the envelopes.
What is the probability that at least one letter will match its intended envelope?
Hint: Let $A_i = \{ith \text{ position will match}\}$
Wish to find: $P(A_1 \cup A_2 \cup A_3 \cup A_4)$....

7 The word "given" in probability... Conditional Probability

Example: Deal or no Deal !!



Say two boxes (envelopes in our game), are opened in the first round What is the probability that the second was a red amount? This depends on the colour of the amount opened on the first box !!!
P(second blue given first red) = P(second blue | first red) =
P(second blue given first blue) = P(second blue| first blue) =
We use the symbol "|" when saying "GIVEN"
These are conditional probabilities.
Given that A happened, what is the probability that B also happened?
The sample space is narrowed down to the space where A has occurred: The sample size now only includes the determination that event A happened.

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

Example: In a study patients are randomly assigned to one of 3 treatments, T_1,T_2,T_3, or a placebo and the number of patients that relapsed or otherwise after two years are tabulated below:

	T_1	T_2	T_3	Placebo
Relapse	18	13	22	24
No relapse	22	25	26	10

Find the probability that a patient will relapse given they were on treatment T_1.
Find the probability that a patient will relapse given they were on the placebo.

Solution:
$P(R|T_1) = \frac{18}{18+22} = 0.7$
(Or $P(R \cap T_1) = \frac{18}{total}$, $P(T_1) = \frac{18+22}{total}$, so $P(R|T_1) = \frac{P(R \cap T1)}{P(T_1)} = 0.7$)
$P(R|Placebo) = \frac{24}{24+10} = 0.64$

Example: Roll 2 dice, let T denote the sum of the two faces. If the sum T is odd what is the probability that the sum has value less than 8?
Solution:
Find $P(T < 8|T \text{ is odd})$. Let A and B denote the events: $B = \{T < 8\}$ and $A = \{T \text{ is odd}\}$.
All possible values satisfying $T < 8$: $\{1, 2, 3, 4, 5, 6, 7\}$
All possible odd sums: $\{3, 5, 7, 9, 11\}$.
$P(A) = P(T \in \{3, 5, 7, 9, 11\}) = \frac{18}{36} = \frac{1}{2}$
$P(A \cap B) = P(T \in \{3, 5, 7\}) = \frac{12}{36} = \frac{1}{3}$
Therefore $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{1/2} = \frac{2}{3}$.

Example: Roll 2 dice until sum of 7 or 8 results (T = 7 or 8). What is the probability it was a 7 that was tossed?
Solution: Let $B = \{T = 7\}$ and $A = \{T = 7 \text{ or } 8\}$

$A \cap B = \{T = 7\}$
$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{6/36}{11/36} = \frac{6}{11}$

7.1 Calculating "and" probabilities using conditional probabilities

$$P(A \cap B) = P(A)P(B|A)$$

Example:
You have r red balls and b black balls in a bin.
Draw 2 without replacement, What is $P(\text{1st red} \cap \text{2nd black})$?
Solution:
$P(\text{1st red} \cap \text{2nd black}) = P(\text{1st red})P(\text{2nd black}|\text{1st red})$
$P(\text{1st red}) = \frac{r}{r+b}$
$P(\text{2nd black}|\text{1st red}) = \frac{b}{r+b-1}$
$P(\text{1st red} \cap \text{2nd black}) = \frac{r}{r+b}\frac{b}{r+b-1}$

Example: Draw 4 without replacement, What is $P(\text{1st red}\cap\text{2nd black}\cap\text{3rd black}\cap\text{4th red})$?
Solution:
$P(\text{1st red}\cap\text{2nd black}\cap\text{3rd black}\cap\text{4th red}) = P(\text{1st red})P(\text{2nd black}|\text{1st red})P(\text{3rd black}|(\text{1st red}\cap \text{2nd black}))P(\text{4th black}|(\text{1st red} \cap \text{2nd black} \cap \text{3rd black}))$
$P(\text{1st red}) = \frac{r}{r+b}$
$P(\text{2nd black}|\text{1st red}) = \frac{b}{r+b-1}$
$P(\text{3rd black}|(\text{1st red} \cap \text{2nd black})) = \frac{b-1}{r+b-2}$
$P(\text{4th red}|(\text{1st red} \cap \text{2nd black} \cap \text{3rd black})) = \frac{r-1}{r+b-3}$
$P(\text{1st red} \cap \text{2nd black} \cap \text{3rd black} \cap \text{4th red} = red) = \frac{r}{r+b}\frac{b}{r+b-1}\frac{b-1}{r+b-2}\frac{r-1}{r+b-3}$

In general :

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_2 \cap A_1) \times ... \times P(A_n|A_{n-1} \cap \cdots A_2 \cap A_1)$$

Task 17: Prove that $P(A_1 \cap A_2 \cap \cdots \cap A_n)$=$P(A_1 \cap A_2 \cap \cdots \cap A_n)$

Solution:

8 Independent events

Events A and B are **independent** if $P(B|A) = P(B)$

Example: Experiment: Throw a die twice, observe upper faces.
Are these pairs of events independent?
a) A= 2 on 1st throw, B = 4 on 2nd throw
b) A= 4 on 1st throw, B = 4 on 2nd throw
c) Let T denote the sum of the two throws, $A = \{T \text{ is odd}\}$ and $B = \{T < 8\}$
d) Let T denote the sum of the two throws, $A = \{T = 7 \text{ or } 8\}$ and $B = \{T = 7\}$
Solution:
a) Yes; $P(A) = 1/6, P(B) = 1/6$, $P(B \cap A) = 1/36$, $P(B|A) = \frac{1/36}{1/6} = 1/6 = P(B)$.
b) Yes; (as previous example)
c) No; $P(B) = P(T < 8) = 21/36$, $P(B|A) = \frac{2}{3}$ (from previous section) so $P(B|A) \neq P(B)$.
d) No; $P(B) = P(\{T = 7\}) = 6/36$, $P(B|A) = \frac{6}{11}$ (from previous section) so $P(B|A) \neq P(B)$.

Events A and B are **independent** if $P(A \cap B) = P(A)P(B)$

Example:Using a coin toss the coin 5 times,
a) what is the probability of getting five heads?
b) what is the probability of getting a H, then T, then T, then H, then T?
c) what is the probability of getting a 2 heads and 3 tails?
Solution:
a) $P(H \cap H \cap H \cap H \cap H) = P(H)P(H)P(H)P(H)P(H) = 0.5^5$
b) $P(H \cap T \cap T \cap H \cap T) = P(H)P(T)P(T)P(T)P(H) = 0.5^5$
c) $P(2 \text{ heads and } 3 \text{ tails}) = \binom{5}{2}P(H)P(H)P(T)P(T)P(T) = \binom{5}{2}0.5^5$

Task 18:Using an unfair coin, in which $P(H) = p$, toss the coin 5 times,
a) what is the probability of getting five heads?
b) what is the probability of getting a H, then T, then T, then H, then T?
c) what is the probability of getting a 2 heads and 3 tails?

Solution:

8.1 Pairwise independence

If you have several events, $A_1, A_2, ...A_n$, that you need to prove independent, it is necessary to show that ALL subsets are independent.
You could prove that any 2 events are independent, which is called **pairwise independence**, but this is not sufficient to prove that all events are independent.

Task 19: Pairwise independence:
Consider a tetrahedral die, equally weighted. Three of the faces are each colored red, blue, and green, but the last face is multicolored, containing red, blue and green.
Experiment: You toss the die and observe the colour in the upper face.
Define the following events $R = red, B = blue, G = green$.

a) Find $P(R), P(B), P(G), P(R \cap B), P(R \cap G), P(B \cap G), P(R \cap B \cap G), P(G|R), P(B|R), P(B|G), P(G|R \cap B)$

b) Are the events pairwise independent?

c) Are all the three events, R, B, G, fully independent?

Solution:

Task 20: Casino game -Craps.
Throw two dice. On first roll: if throw 7 or 11 you win; if throw 2, 3, 12 you lose; any other number, you continue playing. On subsequent rolls: If you eventually roll 7 you lose; if you roll the same number again that continued play, you win!
What's the probability of actually winning?Is the game fair?
(You can use the following table to help you answer the question.)

Dice Roll	Possible Dice Combinations
2	1-1
3	1-2, 2-1
4	1-3, 2-2, 3-1
5	1-4, 2-3, 3-2, 4-1
6	1-5, 2-4, 3-3, 4-2, 5-1
7	1-6, 2-5, 3-4, 4-3, 5-2, 6-1
8	2-6, 3-5, 4-4, 5-3, 6-2
9	3-6, 4-5, 5-4, 6-3
10	4-6, 5-5, 6-4
11	5-6, 6-5
12	6-6

Solution:

9 Bayes Theorem

If a sample space S is broken up into a set of disjoint partitions, $B_1, B_2, \cdots, B_k$, where $B_i \cap B_j = \emptyset$ for $i \neq j$, that is, $S = \bigcup_{i=1}^{k} B_i$, then for any event $A \in S$,

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

(all $A \cap B_i$ are disjoint, $\bigcup_{i=1}^{k} A \cap B_i = A$)

Example: There are two boxes, box 1 contains 60 short bolts and 40 long bolts and box 2 contains 10 short bolts and 20 long bolts. Experiment: Take a box at random, and pick a bolt.

What is the probability that you chose a short bolt?

Solution: Define the events B_1 = choose Box 1, B_2 = choose Box 2 and $A =$ a short bolt

$$\begin{aligned} P(A) =& P(A \cap B_1 \cup A \cap B_2) \\ =& P(A \cap B_1) + (A \cap B_2) \text{...since mutually exclusive} \\ =& (A|B_1)P(B_1) + P(A|B_2)P(B_2) \\ =& \frac{60}{100}\frac{1}{2} + \frac{10}{30}\frac{1}{2} \end{aligned}$$

Example:

A medical detection test is 90% accurate. The accuracy means, in terms of probability: $P(positive|disease) = 0.9$ and $P(positive|\overline{disease}) = 0.1$.

Say, in the general public, the chance of getting the disease is 1 in 10,000.

If the result comes up positive, what is the probability that you actually have the disease?

Solution: Let D be the event you have the disease, then $P(D) = 0.0001$, $P(\overline{D}) = 0.9999$.

$$\begin{aligned} P(D|positive) =& \frac{P(D \cap positive)}{P(positive)} \\ =& \frac{P(positive|D)P(D)}{P(positive|D)P(D) + P(positive|\overline{D})P(\overline{D})} \\ =& \frac{(0.9)(0.0001)}{(0.9)(0.0001) + (0.1)(0.9999)} = 0.0009 \end{aligned}$$

The probability is still very small that you actually have the disease.

In general:

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + ... + P(A|B_k)P(B_k)}$$

Task 21: Identify the source of a defective item.
There are 3 machines: M_1,M_2 and M_3. The percent of items made that come from each machine is: 20%, 30%, and 50%, respectively.
Given the item is from machine M_1,M_2 and M_3, the probability it is defective is 0.01, 0.02 and 0.03 respectively.
a) What is the probability the item selected is defective?
b) What is the probability a defective item came from machine M_1?
c) If the item was not defective what is the probability the item came from machine M_1?

Solution:

Task 22: A gene has 2 alleles: A, a.
The gene exhibits itself through a trait with two versions. The possible phenotypes are dominant, with genotypes AA or Aa, and recessive, with genotype aa. (Aa is considered the same genotype as aA).
Alleles travel independently, derived from a parent's genotype.
In a population, the probability of having a particular allele: P(A) = 0.5, P(a) = 0.5
What is the probability a random person in the population will have genotype;
i) AA? ii) Aa? iii) aa?

Partitions: genotypes of couples/parents are:
(AA, AA), (AA, Aa), (AA, aa), (Aa, Aa), (Aa, aa), (aa, aa).
Assume pairs match regardless of genotype.
What is the probability that a random couple in the population will have each of these combinations of genotypes?

Parent genotypes	Probabilities
(AA, AA)	
(AA, Aa)	
(AA, aa)	
(Aa, Aa)	
(Aa, aa)	
(aa, aa)	

Given the parents have a particular combination of genotype, what is is the probability their child will have a dominant phenotype?

event:	probability:
dom|(AA, AA)	
dom|(AA, Aa)	
dom|(AA, aa)	
dom|(Aa, Aa)	
dom|(Aa, aa)	
dom|(aa, aa)	

If you see that a person has a dominant trait, predict the genotypes of the parents.
If you see that a person has a recessive trait, predict the genotypes of the parents.
For example calculate : $P((AA, AA)|dom) =$

Task 23: You have 1 machine. If the machine is in good condition, then defective items are produced only produced 1% of the time. If the machine is not in a good condition, then defective items are produced produced 40% of the time.
At this stage in the warranty of the machine the company providing the machine claim that the probability that the machine is in good condition is 0.90, and so the probability that the machine is broken is 0.10.
If you sample 6 items, and find that 2 of the 6 are defective, is the machine broken?

Solution:

10 Random Variables

Probability Space: $(S, \mathcal{A}, P)$ where S -sample space, $\mathcal{A}$- set of all events, P -probability
Random variable: X is a function on S with values in real numbers, $X : S \to \mathbb{R}$. Transforms the outcome of an experiment into a number.
Examples:
Toss a coin 10 times, Sample Space = $S = \{HTH...HT, ...\}$, all configurations of H and T.
Let random Variable X = number of heads, $X : S \to \{0, 1, ..., 10\}$ for this example.

Examples:

- Roll a die observe uppermost face: 1,2,3,4,5,6
- Roll 2 dice sum of 2 faces: 2,3,...., 12
- Number of heads when toss coin 10 times: 0,1,2,3,4,...,10
- Number of cars using drive-through 6pm-8pm: 0,1,2...
- Time friend meets you during lunch hour: 0-60 mins

Note outcomes of random variables do not necessarily have to positive:
Examples:

- Mark awarded on a multiple choice question with negative marking : -1, 0, 2.
- Rate of return of a stock : ...-0.5%,...,0%,...2%.....

10.1 The probability distribution

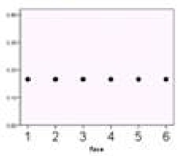
Probability distribution: If $X : S \to \mathbb{R}$ is a random variable, then the function $f(x)$ specifying $f(x) = P(X = x)$ for all outcomes of the variable $X = x$ is the *probability distribution.*
We can specify the probability distribution $f(x)$ using a table, a graph or a formula:
The following are some examples of probability distributions for some discrete random variables:

Example:A fair die.

X	1	2	3	4	5	6
$P(X = x)$	1/6	1/6	1/6	1/6	1/6	1/6

$P(X = x) = 1/6 \quad \forall x \in \{1, 2, 3, 4, 5, 6\}$

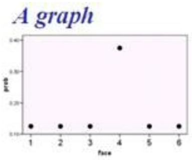
A graph



Example:A loaded die.

X	1	2	3	4	5	6
$P(X = x)$	0.125	0.125	0.125	0.375	0.125	0.125

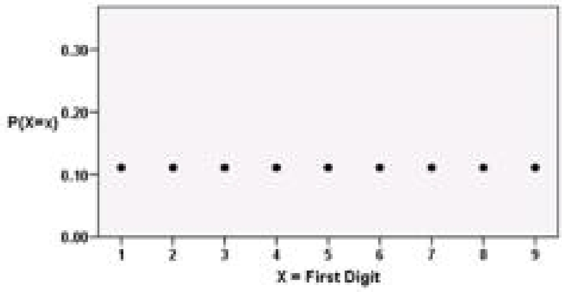
$$P(X = x) = \begin{cases} 0.125 & x \in \{1, 2, 3, 5, 6\} \\ 0.375 & x = 4 \end{cases}$$



Example:First digit in financial records.

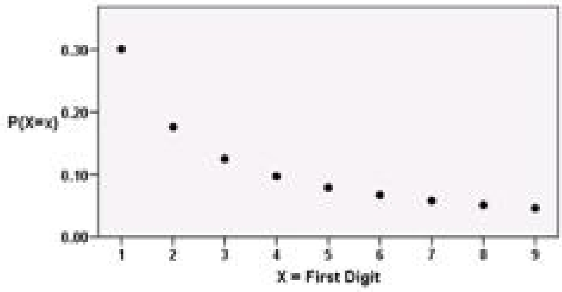
X	1	2	3	4	5	6	7	8	9
$P(X = x)$	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111

$P(X = x) = 0.111 \quad \forall x \in \{1, 2, 3, 5, 6, 7, 8, 9\}$



Example:First digit in financial records- Benford's Law.

X	1	2	3	4	5	6	7	8	9
$P(X = x)$	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046



Task 24:Deal or no Deal!

If X is the winning amount in our game of deal or no deal, describe $f(x)$, the probability distribution for X.

Solution:

Task 25:Multiple choice score.
Let $X =$ the mark you get on multiple choice question in which

$$X = \begin{cases} -1 & \text{if attempt question and get it wrong} \\ 0 & \text{if do not attempt question} \\ 2 & \text{if you attempt question and get it right} \end{cases}$$

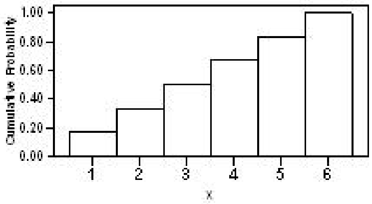
Say the probability that you attempt the question is 0.7.
Say the probability you get the question right given you attempted it is 0.5.
Find the probability distribution of the random variable X.

Solution:

10.2 The Cumulative distribution.

A cumulative probability at $X = x$ is the probability of getting all outcomes less than and equal to x, $P(X \leq x)$.
If $X : S \to \mathbb{R}$ is a random variable, then the function $F(x)$ specifying $F(x) = P(X \leq x)$ for all outcomes of the variable $X = x$ is the *cumulative distribution*.

Example:A fair die.



X	1	2	3	4	5	6
$P(X \leq x)$	0.167	0.333	0.5	0.667	0.833	1

10.3 Expectation and Variance of a discrete random variable

The expected value of a random variable X is defined as:

$$E(X) = \sum_{\text{all } x} xP(X = x)$$

The variance of a random variable X is defined as:

$$Var(X) = \sum_{\text{all } x} (x - \mu)^2 P(X = x)$$

The standard deviation of a random variable X is defined as:

$$SD(X) = \sqrt{Var(X)} = \sqrt{\sum_{\text{all } x} (x - \mu)^2 P(X = x)}$$

$\mu = E(X), \sigma^2 = Var(X), \sigma = SD(X)$

Task 26: Calculate the expected value, the variance and the standard deviation for the variable X the amount to be won in our game of deal or no deal. Interpret the meaning of these values.

Solution:

11 Common Discrete Random Variables

- discrete uniform distribution
- binomial distribution
- poisson distribution
- geometric distribution
- hypergeometric distribution
- negative binomial distribution

11.1 Discrete Uniform Distribution

Uniform distribution of a finite number of values $X \in \{1, 2, 3, ..., n\}$ each outcome has equal probability,

$$f(x) = P(X = x) = 1/n.$$

11.2 Binomial Distribution

Let $X =$ the number of *successes* in a sample of n bernoulli trials, each with probability of success p.

- Number of reds in 15 spins of roulette wheel
- Number of defective items in a batch of 5 items
- Number correct on a 33 question exam
- Number of customers who purchase out of 100 customers who enter store

When does a Binomial Distribution apply?

- A sample of individuals is of size n **individuals**.
- Each individual in the sample is a trial with **two possible outcomes** - a success and a failure (a bernoulli trial).
- The outcome observed for one trial is **independent** of outcomes of any other trial
- Each trial has the same **probability of a success** - let the probability of a success be p.

Write $X \sim Binomial(n, p)$

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{(n-x)} \text{ where } X \in \{0, 1, 2, 3..., n\}$$

Example: Let $X =$ Number of defective items in a batch of 5 manufactured on a production line. Say the probability an item is defective is 0.12.
Comment on the suitability of the binomial probability distribution for this random variable.
Write down the probability distribution.

Solution:

- A sample of $n = 5$ manufactured items.
- Two possible outcomes - a success and a failure - defective and non-defective. Each individual in the sample is a trial with **two possible outcomes** - a success and a failure (a bernoulli trial).

- The probability that the second item is defective does not depend on whether the first item was defective or not, items are independent.
- Each item has the same probability of being defective success be $p = P(defective) = 0.12$. Note: The success outcome is the outcome of interest - this is not necessarily the "good" outcome.

The distribution is :$X \sim Binomial(5, 0.12)$

Expectation and variance:
Using the formula provided for the expected value and variance it can be shown that

$$E(X) = np \text{ and } V(X) = np(1-p)$$

Task 27: For the variable described in the last example, X=Number of defective items in a batch of 5 manufactured on a production line, calculate the following :
a) the probability exactly 3 items out of the 5 are defective.
b) the probability all 5 items are defective.
c) the probability that no items are defective.
d) the probability of no more than 3 items are defective.
e) the probability at least one item is defective.
f) the expected value of X.

Solution:

11.3 Poisson Distribution

Let X = number of events that occur in a specified interval if in an interval the rate of success is λ.

- Number of customers arriving in 20 minutes
- Number of buses arriving at a stop in 30 minutes
- Number of times "trick or treat"ers ring your doorbell per hour on Halloween night

When does a Poisson Distribution apply ?

- An addition to the count is referred to as an event.
- An interval is specified - e.g. in Time, Length, Area, Space
- The probability that an event occurs in an interval is the same for all intervals of the same width.
- The number of events occurring in an interval is independent of the number of events occurring in any mutually exclusive (non-overlapping) interval (or portion of interval).
- The average/mean number of events occurring in the interval is λ, also referred to as a rate.

Write $X \sim Poisson(\lambda)$.

$$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \text{ where } X \in \{0, 1, 2, 3..., \infty\}$$

Example: Let X = "trick or treat"ers that ring your doorbell per hour on Halloween night. Say the rate per hour is 4.
Comment on the suitability of the Poisson probability distribution for this random variable. Write down the probability distribution.

Solution:

a)
- The count is the number of times doorbell is rang, each time is an event.
- An interval is specified - here the interval is 1 hour.
- The probability that a "trick or treat"er rings the doorbell in any hour is the same for all intervals of an hour. The probability between 6pm-7pm is the same as the probability between 7pm-8pm.
- The number of events between 6pm-7pm is independent of the number of events occurring between 7 and 7.15pm.
- The average/mean number of "trick or treat"ers in an hour is λ=4.

b) $X \sim Poisson(4)$.

Expectation and variance:
Using the formula provided for the expected value and variance it can be shown that

$$E(X) = \lambda \text{ and } V(X) = \lambda$$

Task 28: For the variable described in the last example calculate the following:
a) the probability exactly 3 "trick or treat"ers arrive in an hour.
b) the probability that no "trick or treat"ers arrive in an hour.
c) the probability at least 1 "trick or treat"ers arrive in an hour.

Solution:

Task 29: **Some more challenging questions...**

Calculate the following probabilities:

a) the probability exactly 3 "trick or treat"ers arrive in 30 mins.
b) the probability none arrive in 2 hours.
c) the probability that none arrive between 6pm and 7pm and 3 arrive between 7pm and 8pm.
d) the probability that 1 arrives between 6pm and 7pm but at least 2 arrives between 6pm and 8pm.
e) the probability that at least 3 arrive between 6pm and 8pm given 2 arrive in the first hour.

Solution:

11.4 Geometric Distribution

X = the number of Bernoulli trials, with probability of success p, on which the first success occurs $X \in \{1, 2, 3...\}$.

$$f(x) = P(X = x) = (1-p)^{(x-1)}p$$

Write $X \sim Geom(p)$.
When does a Geometric Distribution apply?

- Each individual is a trial with **two possible outcomes** - a success and a failure (a bernoulli trial).
- The outcome observed for one trial is **independent** of outcomes of any other trial
- Each trial has the same **probability of a success** - let the probability of a success be p.

Expectation and variance:
Using the formula provided for the expected value and variance it can be shown that

$$E(X) = \frac{1}{p} \text{ and } V(X) = \frac{1-p}{p^2}$$

Task 30: Let X be the number of items selected from a batch of manufactured product on a production line on which the first defective item is selected. Say the probability an item is defective is 0.12.
a) Write down the probability distribution.
b) What is the probability the first defective item will be the fourth item selected from the batch?
c) What is the probability the first defective item will be the tenth item selected from the batch?
d) Calculate the expected value of this variable and interpret its meaning.

Solution:

11.5 Negative Binomial

X is the number of trials before r successes are observed.
Write $X \sim NB(r, p)$, $X \in r, r+1, r+2 \cdots$

$$f(x) = P(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

Note that the geometric distribution can be viewed as a special case: $Geom(p) = NB(1, p)$
Take care: the negative binomial can be defined in a number of ways across texts!!
Expectation and variance:
Using the formula provided for the expected value and variance it can be shown that

$$E(X) = \frac{r}{p} \text{ and } V(X) = \frac{r(1-p)}{p^2}$$

When does a Negative Binomial Distribution apply?

- Each individual is a trial with **two possible outcomes** - a success and a failure (a bernoulli trial).
- The outcome observed for one trial is **independent** of outcomes of any other trial
- Each trial has the same **probability of a success** - let the probability of a success be p.

Task 31: Let X be the number of items selected from a batch of manufactured product on a production line on which the third defective item is selected. Say the probability an item is defective is 0.12.
a) Write down the probability distribution.
b) What is the probability the third defective item will be the third item selected from the batch?
c) What is the probability the third defective item will be the tenth item selected from the batch?
d) Calculate the expected value of this variable and interpret its meaning.

Solution:

11.6 Hypergeometric Distribution

$X=$ number of successes out of n trials drawn from a total of N in which there are a total of m successes.
*This is sampling without replacement. The probability of success on the *ith* trial is conditional on the outcome of previous trials.

$$f(x) = P(X = x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

Write $X \sim Hypergeometric(N, m, n)$
When does a Hypergeometric Distribution apply?

- A population of size N.
- The population contains two types of individuals, m *successes* and $N - m$ *failures.*
- A sample of size n is selected from the population without replacement.

Expectation and variance:
Using the formula provided for the expected value and variance it can be shown that

$$E(X) = \frac{nm}{N} \text{ and } V(X) = \frac{nm(N-n)(N-m)}{N^2(N-1)}$$

Task 32: An urn contains 30 black marbles and 20 red marbles. A sample of 10 marbles is selected at random from the urn. Marbles are not replaced before the next marble is selected. Let X be the number of black marbles that are in the sample of 10 marbles selected.
a) Write down the probability distribution.
b) What is the probability that exactly 3 black marbles are selected?
c) What is the probability that all 10 marbles selected are black?
d) What is the probability that no marbles selected are black?
e) What is that no more than 2 black marbles are selected?
f) What is the expected value of this random variable? Interpret the meaning of this value?

Solution:

12 Common Continuous Distributions

Task 33: Class discussion: Discuss the difference between a discrete random variable and a continuous random variable. Give some examples to illustrate the difference.

Solution:

For continuous random variables we generally don't look for the probability of an individual value, but instead look for the probability of a range of values occurring.
e.g. $P(X < x)$, $P(X > x)$, $P(x_1 < X < x_2)$
Probability density function: For a continuous random variable X, the function $f(x)$, such that,

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx$$

is the *probability density function.*
Properties:

- $f(x) \geq 0$ for $-\infty < x < \infty$
- The summation used in a discrete random variable is replaced by an integral for the continuous random variable:

$$\sum_{\text{all } x} P(X = x) = 1 \rightarrow \int_{\text{all } x} f(x)dx = 1$$

- Probabilities of ranges of values of X is equivalent to calculating the area under the density function within that range of values.
- The expectation and variance of a random variable X are found by

$$\mu = E(X) = \int_{\text{all } x} xf(x)dx \quad \text{and} \quad \sigma^2 = V(X) = \int_{\text{all } x} (x-\mu)^2 f(x)dx$$

- The **cumulative distribution** for a continuous random variable is defined as

$$F(x) = P(X < x) = \int_{-\infty}^{x} f(x)dx$$

The cumulative distribution for a continuous random variable is often referred to as simply *the distribution function.*

12.1 Continuous Uniform distribution

A random variable X is defined on a continuous range, $a \le X \le b$, such that the probability density function is the same for all values in the range $[a, b]$, i.e.

$$\text{p.d.f.: } f(x) = \frac{1}{b-a}, \text{ for } x \in [a, b] \text{ (and has value 0 elsewhere)}$$

Write $X \sim U(a, b)$.

Example: Say you arrange to meet your friend at lunchtime, where the time that your friend turns up is as equally likely to be any time in that 60 minute interval. Let X be the time from 0 to 60 minutes (inclusive) in which your friend turns up.

Task 34:

a) What is the density function for this random variable?
b) Sketch this density function.
c) What is the probability that your friend turns up anytime in the first half hour?
d) What is the probability that your friend turns up anytime in the first 20 minutes?
e) What is the probability that your friend turns up anytime between 30 minutes and 50 minutes into the hour?
f) Find the cumulative distribution function, $F(X)$, for this random variable.
g) What is the value of $E(X)$? Interpret the meaning of this value.

Solution:

Solution:

In general:

- Where $X \sim U[a, b]]$, and x_1 and x_2 are values in this range such that $a \leq x_1 < x_2 \leq b$, then
$$P(x_1 \leq X \leq x_2) = \frac{x_2 - x_1}{b - a}$$
- The cumulative distribution is:
$$F(x) = P(X < x) = \int_a^x \frac{1}{b-a} dx = \frac{x-a}{b-a}$$
- **Expectation and variance:**
Using the formula provided for the expected value and variance it can be shown that
$$\mu = E(X) = \frac{a+b}{2} \text{ and } \sigma^2 = V(X) = \frac{1}{12}(b-a)^2$$

12.2 Exponential Distribution

A random variable X is defined on a continuous range, $X \geq 0$, such that the probability density function is defined, for some value of the parameter $\lambda > 0$, as

$$\text{p.d.f.: } f(x) = \lambda e^{-\lambda x}, \text{ for } x \geq 0 \text{ (and has value 0 elsewhere, i.e. } x < 0)$$

Write $X \sim Exp(\lambda)$.

Some examples:

Let X be the life span of an electrical component.

Let X be the **waiting time** until an event occurs (e.g. waiting time between successes, where the number of successes follows a Poisson distribution).

The parameter λ can be interpreted as the inverse of the average life span/waiting time.

Expectation and variance:

Using the formula provided for the expected value and variance it can be shown that

$$\mu = E(X) = \frac{1}{\lambda} \text{ and } \sigma^2 = V(X) = \frac{1}{\lambda^2}$$

Example: Suppose the lifetime of a battery, in hours, is exponentially distributed with a rate of $\lambda = \frac{1}{10}$.

Task 35:

a) What is the density function for this random variable?

b) Sketch this density function.

c) What is the probability that the battery will last less than 5 hours?

d) What is the probability that the battery will last longer than 10 hours?

e) What is the probability that the battery will last between 5 and 10 hours?

f) Find the cumulative distribution function, $F(X)$, for this random variable.

g) What is the value of $E(X)$? Interpret the meaning of this value.

h) What is the probability that lifetime of the battery will be at least 7 hours given that it has already lasted 5 hours?

Solution:

Solution:

In general:

- The cumulative distribution for the random variable $X \sim Exp(\lambda)$ is:

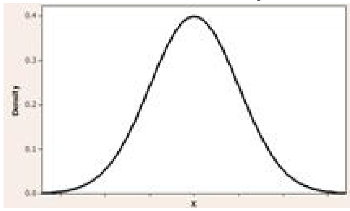
$$\begin{aligned} F(x) &= \int_0^x f(x)dx = \int_0^x \lambda e^{-\lambda x} dx \\ &= 1 - e^{-\lambda x} \end{aligned}$$

- The memoryless property: $P(X > s + t | X > t) = P(X > s)$

12.3 Normal Distribution

A random variable X is defined on a continuous range, $-\infty < X < \infty$, such that the probability density function is defined, as

Sketch of the density function:



$$\text{p.d.f.: } f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

where $\mu = E(X)$ and $\sigma^2 = V(X)$.

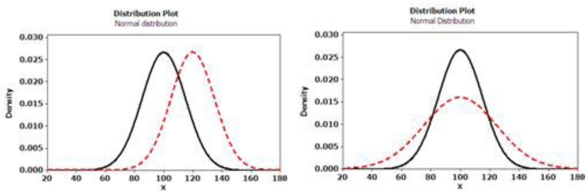
Write $X \sim Normal(\mu, \sigma^2)$.

Some examples:

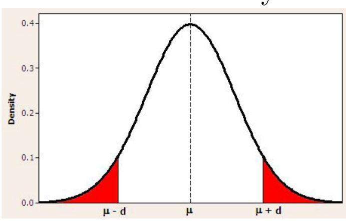
- Let X be the height of the male population (or female population).
- Let X be IQ with a mean of 100 and a standard deviation of 15. IQ is thought to be normally distributed, $X \sim Normal(\mu = 100, \sigma^2 = 15^2)$.
- A survey of per capita income indicated that the annual income for people in a certain country is normally distributed with a mean of 36,000 euro and a standard deviation of 1,600 euro. Let X be annual income, $X \sim Normal(\mu = 36000, \sigma^2 = 1600^2)$.

12.3.1 Properties

- A change in the parameters μ or σ results in;

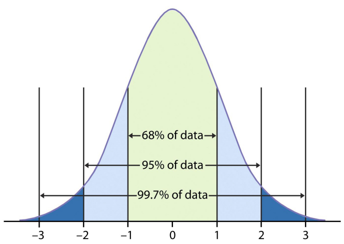


- The distribution is symmetrical, i.e. $P(X < \mu - d) = P(X > \mu + d)$



- Empirical rule:

$P(\mu - 1\sigma < X < \mu - 1\sigma) \approx 68\%$
$P(\mu - 2\sigma < X < \mu - 2\sigma) \approx 95\%$
$P(\mu - 3\sigma < X < \mu - 3\sigma) \approx 99.7\%$



12.3.2 Probabilities

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)dx$$

Cumulative probability:

$$P(X < x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)dx$$

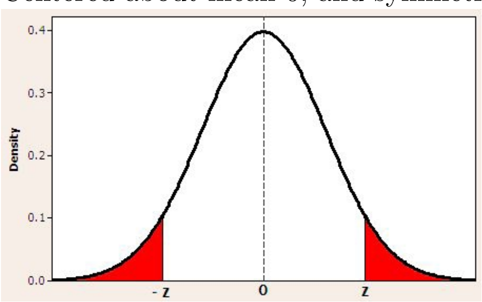
These can be found using a set of tabulated probabilities for the *standard* normal distribution....

12.3.3 Standard Normal Distribution

The *standard* normal distribution is a special case, in which the parameters have value $\mu = 0$, $\sigma = 1$; Write $Z \sim Normal(0, 1)$.
A random variable Z is defined on a continuous range, $-\infty < Z < \infty$, such that the probability density function is defined, as

$$\text{p.d.f.: } f(z) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}z^2\right)$$

Centered about mean 0, and symmetrical, $P(Z < -z) = P(Z > z)$



12.3.4 Probabilities of the standard normal distribution

Task 36: Use the cumulative standard normal table to find the following probabilities:
a) $P(Z < 0.6)$
b) $P(Z < 0.62)$
c) $P(Z < 1.96)$
d) $P(Z > 0.62)$
e) $P(Z < -0.62)$
f) $P(-1.96 < Z < 1.96)$
g) $P(-2.58 < Z < 2.58)$

Solution:

12.3.5 Z-scores

Values of X where $X \sim Normal(\mu, \sigma^2)$, can be converted to a corresponding value of the standard normal scale, $Z \sim Normal(0, 1)$, the *z-score*;

$$Z = \frac{X - \mu}{\sigma}$$

- A z-score of value 0 implies about average
- A z-score above 0 implies above average
- A z-score below 0 implies below average

Task 37: Mensa problem: Calculate the z-score for an individual with
a) IQ=110,
b) IQ=130,
c) IQ=70.

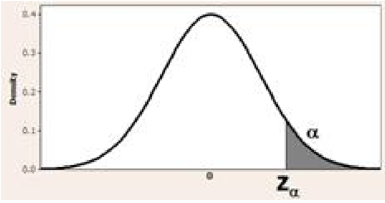
Solution:

Task 38: Mensa problem: Calculate the following probabilities;
a) What is the probability that a randomly selected person will have an IQ less than 100?
b) What is the probability that a randomly selected person will have an IQ less than 110?
c) What is the probability that a randomly selected person will have an IQ greater 125?
d) What is the probability that a randomly selected person will have an IQ between 70 and 130?

Solution:

12.3.6 Working in reverse - percentiles

Working in reverse on the standard normal scale, $Z \sim Normal(0,1)$;



Z_α : the value of the standard normal distribution at which the probability of being greater than that value is α.

Task 39: Find the following values:
a) $Z_{0.025} =$
b) $Z_{0.05} =$
c) $Z_{0.01} =$

Solution:

Working in reverse for any normal distribution parameters, $X \sim Normal(\mu, \sigma^2)$;

If $X \sim Normal(\mu, \sigma^2)$ for what value of X is a certain proportion of individuals in the population above/below.

Example: Members of Irish Mensa are in the top 2% of the population as regard to their IQ value. If it is assumed that IQ follows the distribution as provided previously, what IQ value would a person have to have to be a member?

Solution:

13 Functions of variables

Example: A scratch card costs 3 euro to buy. You can win 10 euro, with probability 0.2. Would you gamble?

Task 40: Let X be the variable "amount won".
a) What is the probability distribution?
b) What is the expected value, the variance and standard deviation of X?

Solution:

Task 41: Let Y be the variable "Gain/Loss".
a) What is the probability distribution?
b) What is the expected value, the variance and standard deviation of Y?

Solution:

The relationship between these two random variables is:
"Gain/Loss" = "Amount won" - 3. i.e. $Y = X - 3$.

Task 42: What is the relationship between the expected value, the variance and standard deviation of Y and the expected value, the variance and standard deviation of X?

Solution:

Example: There are four balls in an urn, two of which are red and two are black. You select two without replacement winning money for each black ball you select.

Task 43: Let X be the number of black balls that you select.
a) What is the probability distribution?
b) What is the expected value, the variance and standard deviation of X?

Solution:

Task 44: Say you win 2 euro for each black ball selected and pay 3 euro to play the game. Let Y be the variable "Gain/Loss".
a) What is the probability distribution?
b) What is the expected value, the variance and standard deviation of Y?

Solution:

The relationship between these two random variables is:
"Gain/Loss" = 2("Number black balls selected") - 3. i.e. $Y = 2X - 3$.

Task 45: What is the relationship between the expected value, the variance and standard deviation of Y and the expected value, the variance and standard deviation of X?

Solution:

In general: If $Y = aX + b$ where a and b are constants, then
$E(Y) = aE(X) + b$,
$V(Y) = a^2V(X)$,
$SD(Y) = aSD(X)$

13.1 Functions of Binomial random variables

If X and Y are independent random variables with distributions,
$X \sim Binomial(n, p)$ and $Y \sim Binomial(m, p)$, then
$X + Y \sim Binomial(n + m, p)$.

13.2 Functions of Poisson random variables

If X and Y are independent random variables with distributions,
$X \sim Poisson(\lambda_1)$ and $Y \sim Poisson(\lambda_1)$, where λ_1 and λ_2 are averages/rates defined on the same interval, then
$X + Y \sim Poisson(\lambda_1 + \lambda_2)$.

Example: The manager of a fast food drive-through with two entrances knows that cars arrive at the rate of 3 per hour at entrance A and 8 per hour through entrance B. Let N be the total number of cars arriving at the drive-through per hour. Assuming that the number of cars that arrive through entrance A and B are independent,
a) what is the distribution of N?
b) Calculate the probability that there are between 10 and 14 cars, inclusive, arriving in a given hour.

Solution:

13.3 Functions of normally distributed random variables

If X and Y are independent random variables with distributions,
$X \sim Normal(\mu_1, \sigma_1^2)$ and $Y \sim Normal(\mu_2, \sigma_2^2)$, and a and b are constants, then
$aX + bY \sim Normal(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$.

Example: The time taken for a truck to drive from town A to town B (in that direction) is known to be normally distributed with a mean of 5 hours and a standard deviation of 1 hour. The time taken for a truck to make the return journey from town B to town A is known to be normally distributed with a mean of 4 hours and a standard deviation of 2 hours.
A man is to drive a truck from town A to town B, wait in a depot in town B for the truck to be loaded, and then return immediately to town A. The time taken to load the truck is normally distributed with a mean of 1 hour and a standard deviation of 1 hour.
Assuming that the lengths of time taken for the outward and return journeys, and the time taken to load the truck are independent, what is the probability that:
a) the outward journey from A to B will take longer than 5.5 hours,

b) the length of time the truck will be gone from town A will be less than 12 hours,
c) the time taken to drive from A to B will be less than that taken to drive from B to A.

Solution:

Example: The population of Statsville has weights that are normally distributed with a mean of 70kg and a standard deviation of 10kg. A lift installed in a building in Statsville will function to a total weight capacity of 480kg.

a) If 6 people get in the lift what is the probability distribution for the total weight in the lift?
b) What is the probability that the lift will not function due to excess weight?

Solution:

If T is a random variable representing the sum of n independent identically distributed random variables, $T = X_1+X_2+\cdots+X_n$, for which each X_i has distribution, $X_i \sim Normal(\mu, \sigma^2)$, then the total will have distribution

$$T = \sum_{i=1}^{n} X_i \sim Normal\left(n\mu, n\sigma^2\right).$$

Example: A soft-drink vending machine is set so that the amount of drink dispensed into a cup is a random variable with a mean of 200ml, a standard deviation of 15ml and is normally distributed.

a) If 10 cups are sampled from this machine, what is the probability distribution for the average amount dispensed in this sample of 10 cups?

b) What is the probability that the average amount dispensed of these 10 cups will be less than 190ml?

Solution:

If $\overline{X}$ is a random variable representing the average of n independent identically distributed random variables, $\overline{X} = \frac{X_1+X_2+\cdots+X_n}{n}$, for which each X_i has distribution $X_i \sim Normal(\mu, \sigma^2)$, then the average will have distribution

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim Normal\left(\mu, \frac{\sigma^2}{n}\right).$$

14 The Central Limit Theorem

Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables each with mean μ and standard deviation σ.
Then as $n \to \infty$ the distribution of $\sum_{i=1}^{n} X_i$ tends to

$$\sum_{i=1}^{n} X_i \approx Normal\left(n\mu, n\sigma^2\right)$$

14.1 Application: Normal approximation to Binomial

Example: Let X be the number of heads observed when a fair coin is tossed 40 times.
a) What is the probability distribution for this random variable?
b) Find the probability that exactly 20 heads occur.
c) Find the probability that at least 22 heads occur.

The Binomial distribution is a sum of n identically distributed independent bernoulli trials with probability of success p.

$$X \sim Binomial(n, p) = \sum_{i=1}^{n} X_i$$

where $X_i = \begin{cases} 0 & \text{with probability } p \\ 1 & \text{with probability } 1-p \end{cases}$.
$\mu = E(X_i) = 0(1-p) + 1(p) = p$,
$\sigma^2 = V(X_i) = (0-p)^2(1-p) + (1-p)^2(p) = p(1-p)$.

$$X = \sum_{i=1}^{n} X_i \approx Normal\,(np, np(1-p))$$

$$X \sim Binomial(n, p) \approx Normal\,(np, np(1-p))$$

However, there is a catch or two...
Catch 1:
Provided the expected number of successes, np and the expected number of failures $n(1-p)$ are large enough. Rule of thumb: both greater than 5.
Catch 2:
The binomial distribution is a discrete random variable but the normal distribution is a continuous random variable - Apply a continuity correction.

Discrete		Continuous
$P(X \geq x)$	$\to$	$P(X > x - \frac{1}{2})$
$P(X \leq x)$	$\to$	$P(X < x + \frac{1}{2})$
$P(x_1 \leq X \leq x_2)$	$\to$	$P(x_1 - \frac{1}{2} < X < x_2 + \frac{1}{2})$

Solution:

14.2 Normal approximation to Poisson

$X \sim Poisson(\lambda) \approx Normal(\mu = \lambda, \sigma^2 = \lambda)$

Example: The manager of a fast food drive-through with two entrances knows that cars arrive at the rate of 3 per hour at entrance A and 8 per hour through entrance B. Let N be the total number of cars arriving at the drive-through per hour. Assuming that the number of cars that arrive through entrance A and B are independent,

a) what is the distribution of N?

b) Calculate the probability that there are 6 or more cars arriving in a given hour

c) Calculate the probability that there are between 10 and 14 cars, inclusive, arriving in a given hour.

Solution:

14.3 Application to the Sampling Distribution of the Mean

For a random variable X, which has mean μ and standard deviation σ,
the mean of a sample of n observations will have distribution

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \approx Normal\left(\mu, \frac{\sigma^2}{n}\right).$$

This approximation being better for larger sample sizes, typically $n \geq 30$. Note that this is true for *any* population distribution X.

Simulation demonstrated in class.

Q: Explain the difference between the statement of the Central Limit Theorem and the distribution for $\overline{X}$ outlined in section 13 ?