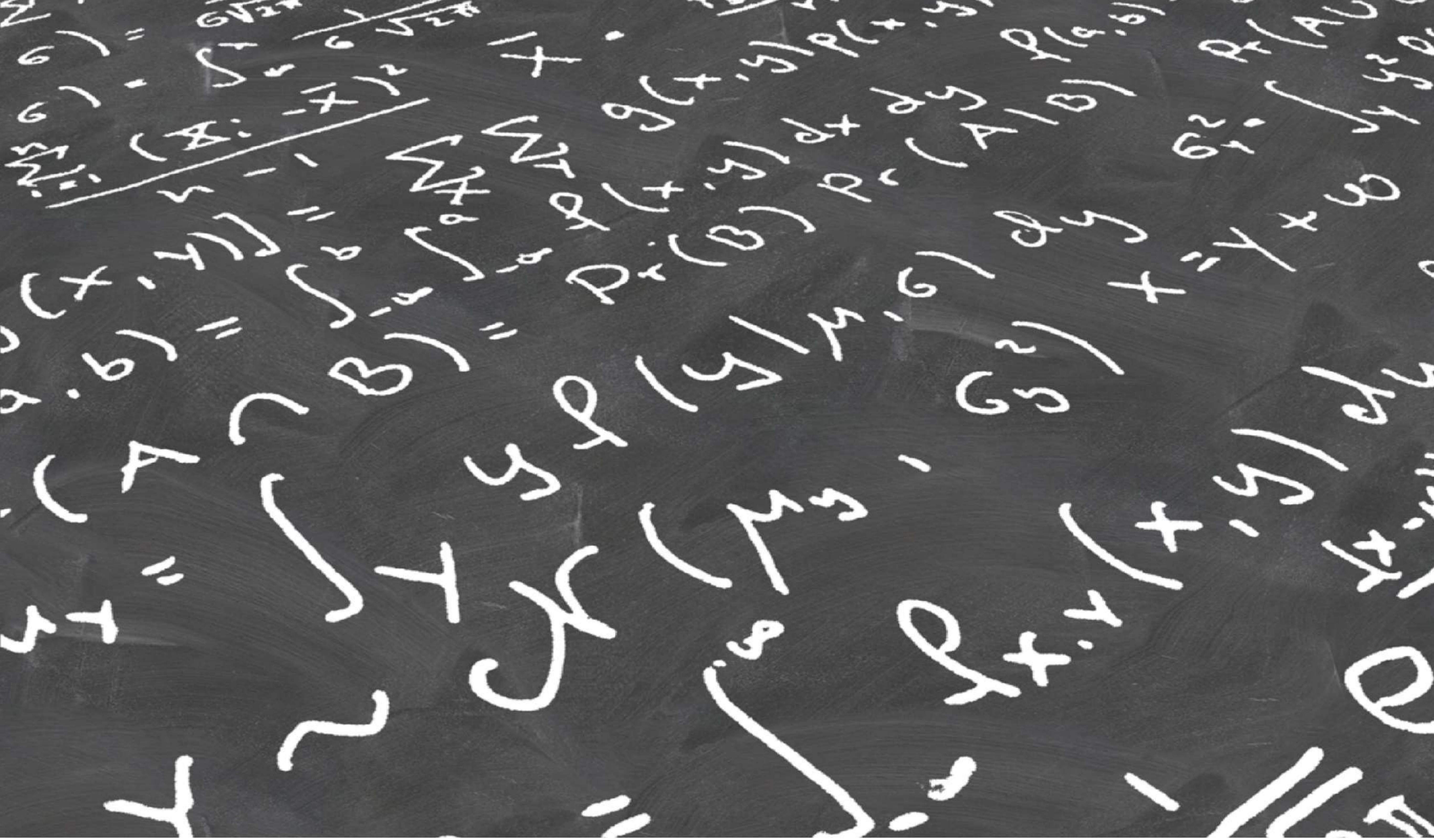


1. Introduction to Probability Theory

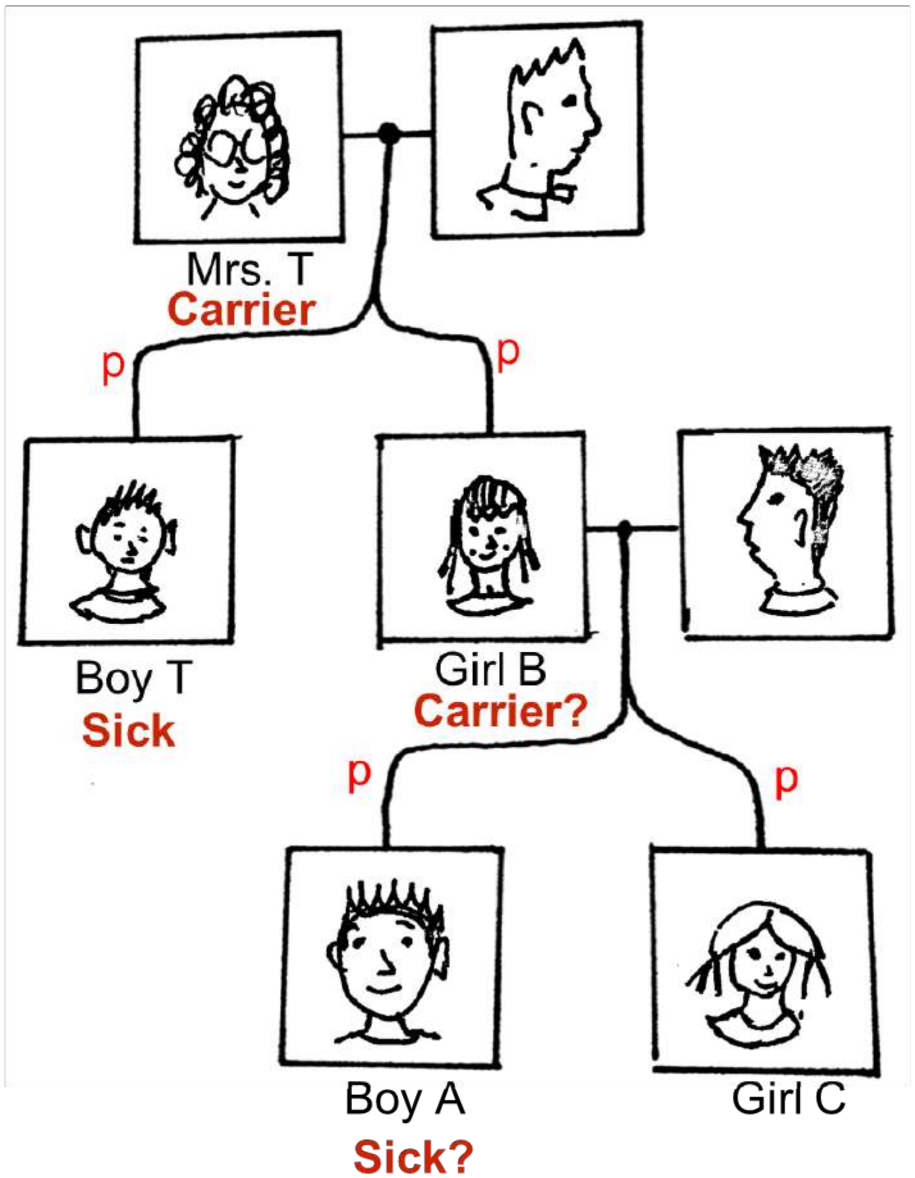
Gunvor Elisabeth Kirkelund
Lars Mandrup

Todays Content

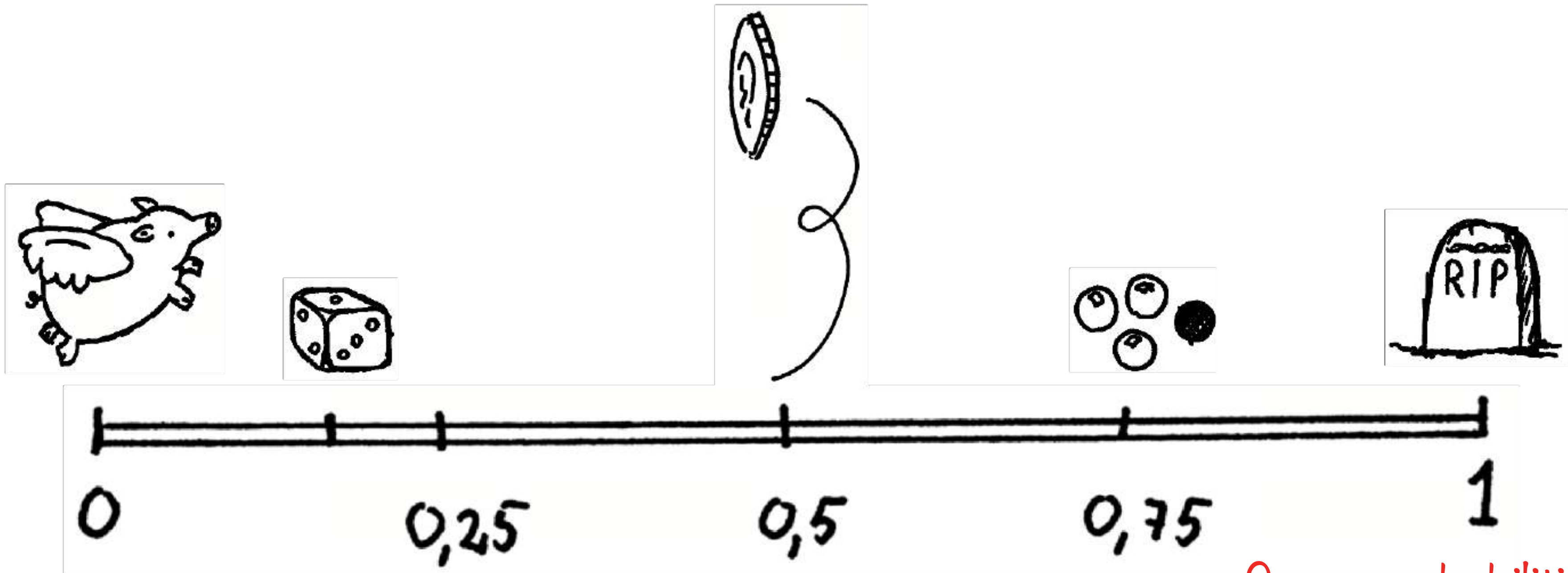
- Introduction to Probability Theory
- Definitions, concepts and notation
- Relative Frequency Approach
- Set theory
- Basic Axions on probabilities

Example - X linked recessive disease

→ Conditional probabilities
→ Don’t (only) rely on logic
→ Systematic calculations



Probability Line



Some probabilities you know

- All probabilities are numbers between 0 and 1.
- In percentage, between 0% to 100%.
- We begin with one sample point.

Words to Know

- Experiment/trial (*Forsøg/test*) — Roll a dice
- Sample space (*Udfaldsrum*) — $S=\{1,2,3,4,5,6\}$
- Sample point (*Bestemt udfald*) — $a=\{4\}$
- Event (*Hændelse*) — $A=\{2,4,6\}$ (even number)
  - Elementary event — Event that has <u>one</u> possible outcome
  - Joined event — Event that has <u>many</u> possible outcomes
  - Simultaneous event — Event with <u>two or more sub trials</u>

Relative Frequency Approach

- The number of times event A occurs: N_A

- The number of times that all events occur (sample space):

$$N = N_A + N_B + N_C + \cdots$$

- Then we have the relative frequency:

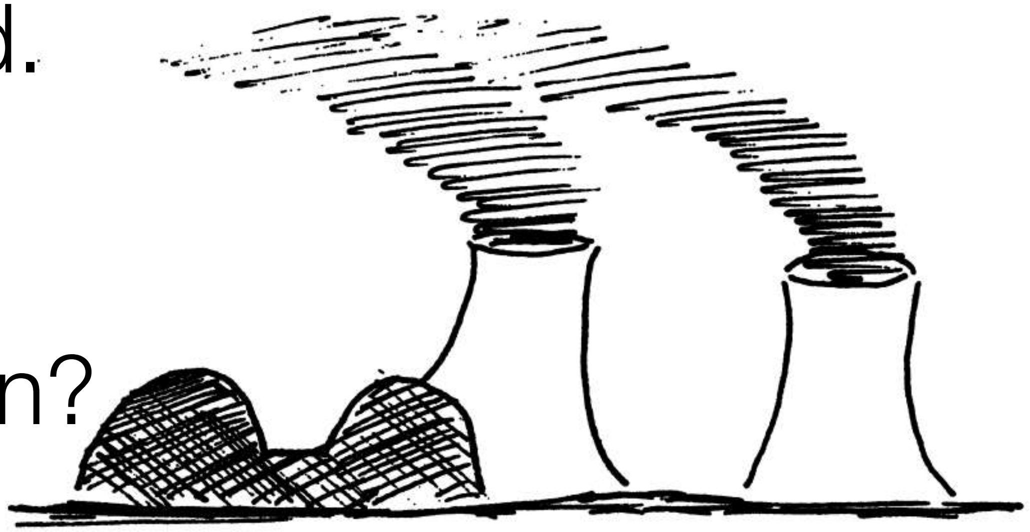
$$\boxed{Pr(A) \sim r(A) = \frac{N_A}{N}}$$

All sample points should have the same a priori probability

- Where: $Pr(A) = \lim_{N \to \infty} r(A)$

Risk of a Meltdown

- There are 437 reactors in the world.
- ~153M operating reactor hours.
- ~Four reactor meltdowns.
- What are the chance of a meltdown?



$$\frac{4}{153M} \text{ pr. reactor pr. hour}$$

$$\sum_{n=1}^{437} \frac{4}{153M} = \frac{1}{87600} \text{ pr. hour}$$

$$\frac{24*365}{87600} = \frac{1}{10} \text{ pr. year}$$

→ Be carefull: Small samples, small probabilities, circumstances, etc.

→ Very uncertain: If number of reactors > 4370 → Pr(Meltdown)>1

Set Theory (*Mængdelære*)

A set:

- A collection of things.
- Elements of sets are not ordered.

Name of set

Some more elements

$$E = \{\alpha_1, \ldots, \alpha_n\}$$

Element

Example:

- The set of all persons in a drug trial group.
- The number of cars i DK.
- All numbers.
- All colours.

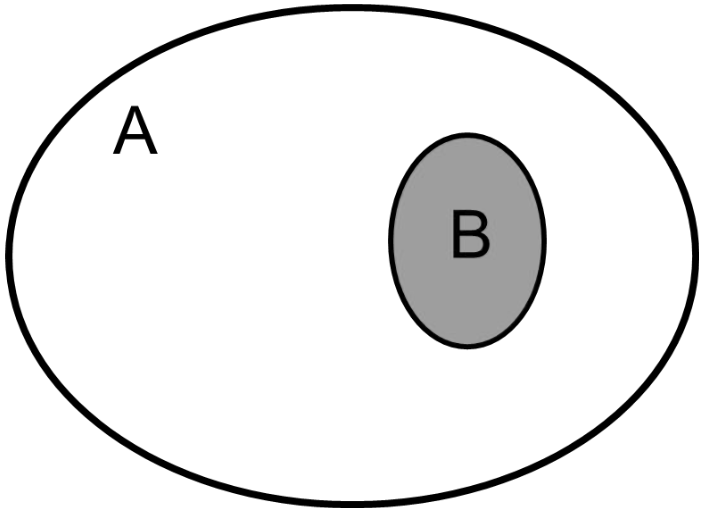
A Subset to a Set *(Delmængder)*

- A subset is any set, where all elements are included in the original set

Notation:

B is a subset to A:

$$B \subset A$$



Example:

For a set $A = \{blue, red, green\}$

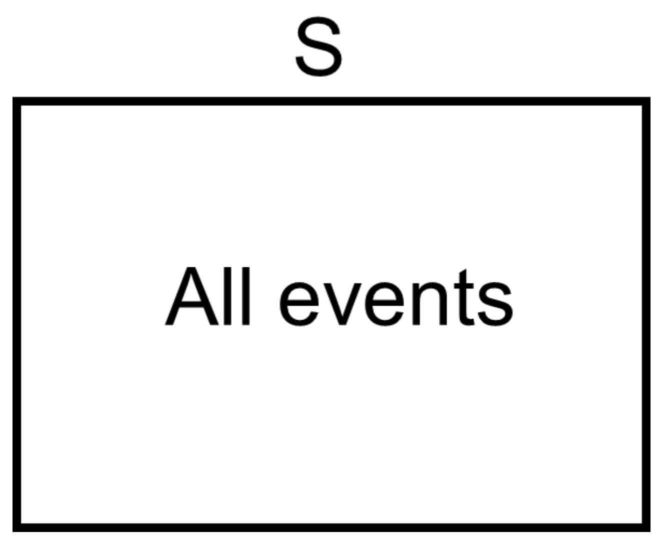
we have a subset $B \subset A$ if B is in A,

e.g. $\{blue, red\}$,$\{blue, red, green\}$,$\{green\}$, $\{\}$

The Sample Space (*Udfaldsrum*)

- The sample space contains <u>all possible events</u>.
- The probability that a sample is from the sample space is 1.

Pr(S)=1

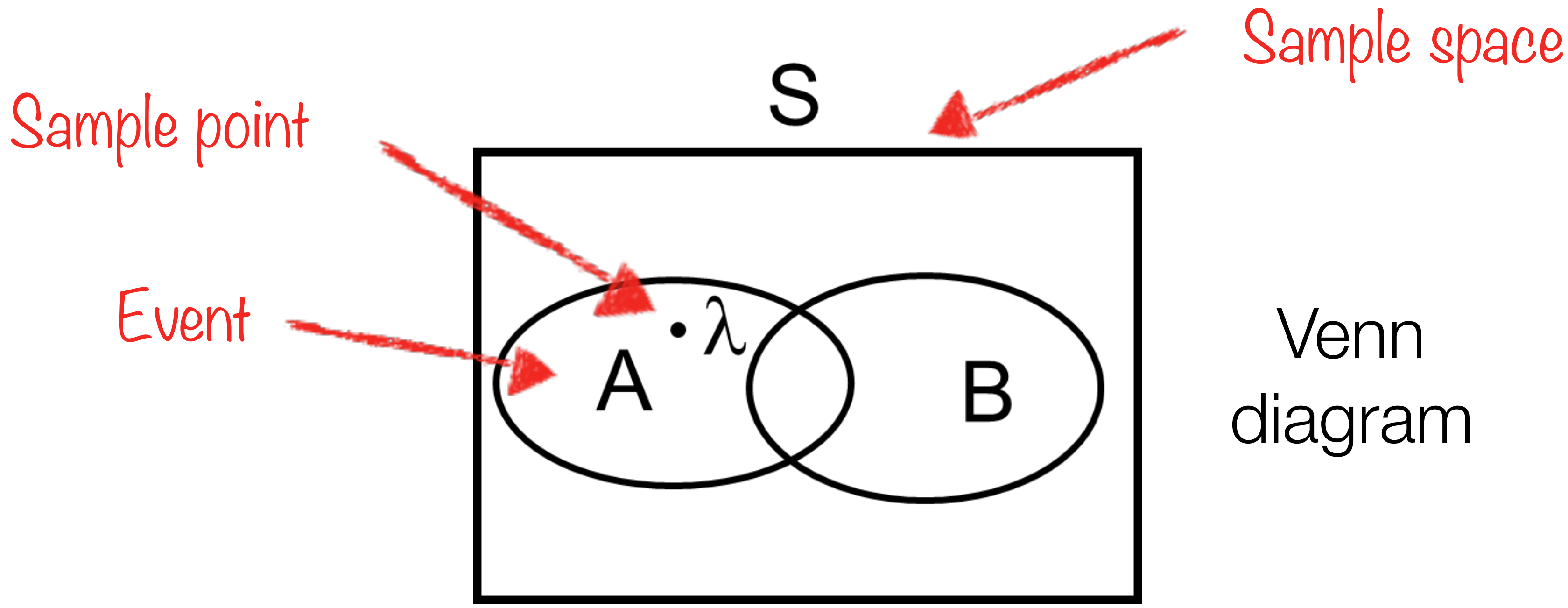


Example:

- A sample space contains 20 people
- 8 has a given disease, 12 is healthy
- Draw a random person.
- What is the chance that a person is a person?
- What is the chance that a person is sick?

A Sample Point *(Udfald)*

- An elementary event.
- Events are collections of sample points.
- Sample space is the collection of all possible sample points.
- Sample points are not ordered.



Venn diagram

Example:

Throw of a dice:

Possible outcomes: 1,2,3,4,5,6 → S={1,2,3,4,5,6}

Events: A={1,2,3} and B={2,4,6}; $A \subset S$; $B \subset S$

Basic Axions of Probability

- The probability of a sample point (element of a sample space).
- The probability of a event (E) (collection of sample points).
- All sample points of a probability space (S) sums up to 1.

➢ Basic Axions of Probability:

$$\boxed{\begin{aligned} &\textbf{Axion 1:}\quad 0 \le Pr(E) \le 1 \\ &\textbf{Axion 2:}\quad Pr(S) = 1 \end{aligned}}$$

The Empty Set (*Den tomme mængde*)

- The empty set is always a subset of any set.
- This corresponds to the impossible event.

$$\varnothing = \{\}$$ The null set

- The probability of the impossible event is 0.

Example:

- The set of boys in an all girlschool.
- The chance of pigs growing wings and fly.
- To get an 8 when rolling a dice.

Summary

S is the certain set.

The certain event

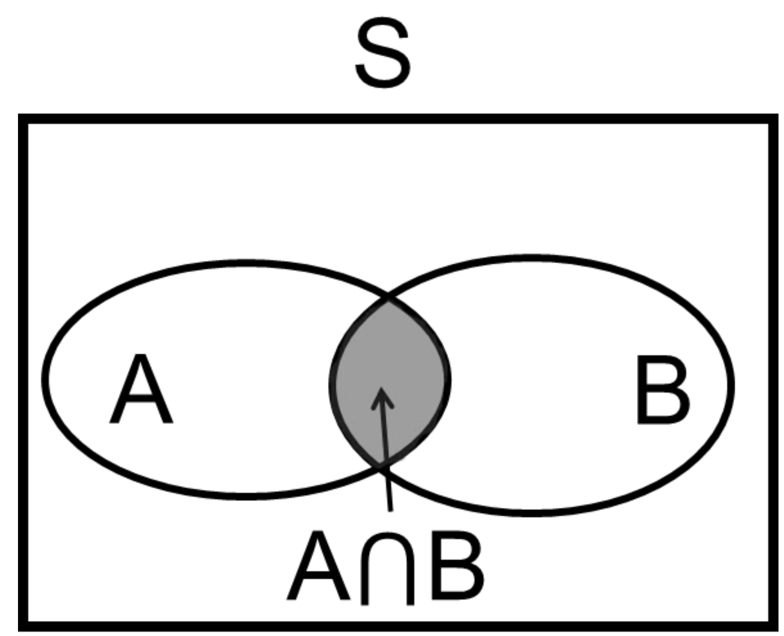
$\varnothing$ is the empty set.

The impossible event

Joint Events (*Fællesmængde*)

- The intersection $A \cap B$ are the common elements of the events A and B
- $A \cap B$ means A and B.

Venn diagram

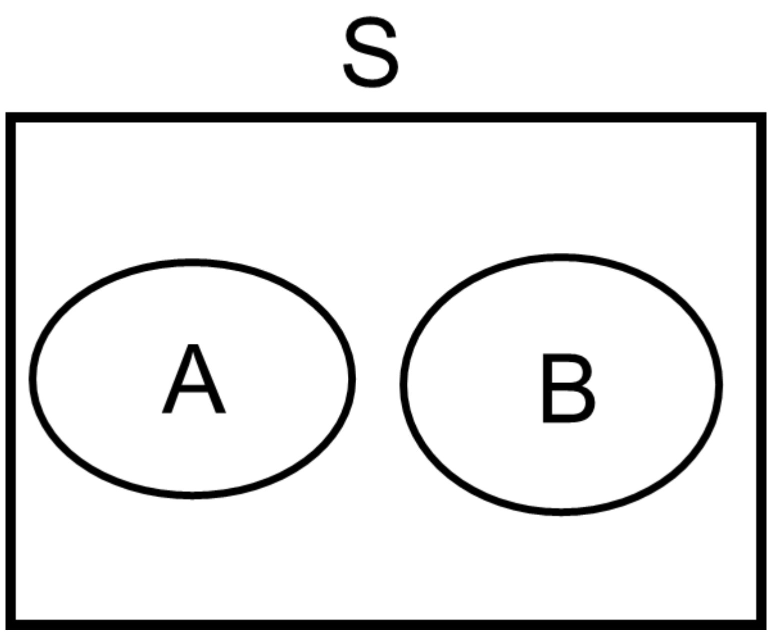


Example:

- A is the event of VW cars i DK
- B is the event of red cars in DK
- The intersection of the events is all red VW in DK.

Mutually Exclusive (Disjoint) Events *(Disjunkte)*

- The sets of A and B are disjoint if: $A \cap B = \emptyset$



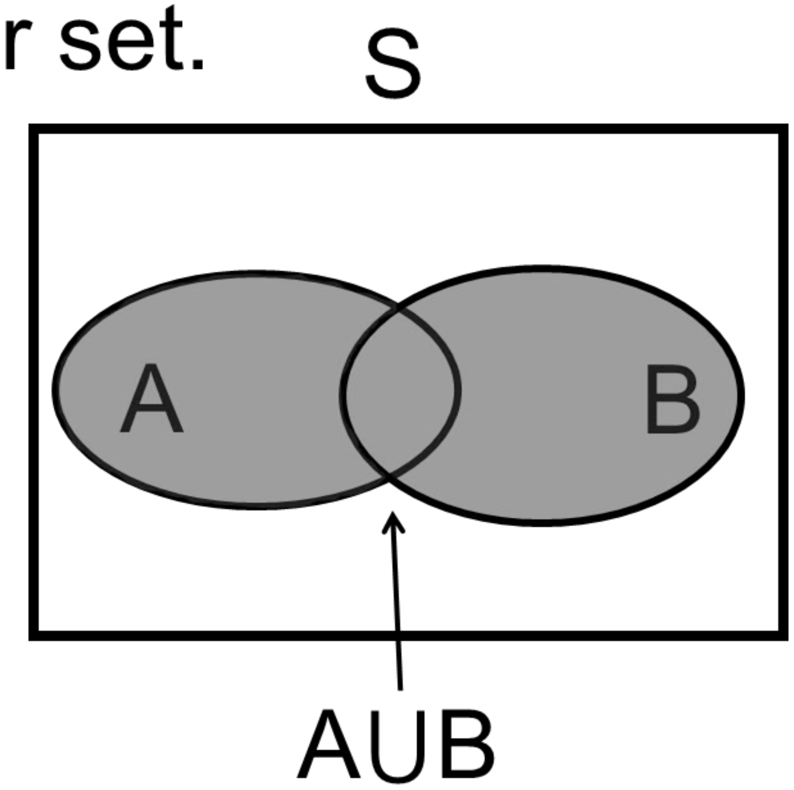
$A \cap B = \varnothing$

Example:

- Event A: The child is a girl.
- Event B: The child is a boy.

Union of Events (*Foreningsmængde*)

- The union of events $A \cup B$ are all the events in one set 'plus' the events in the other set.
- $A \cup B$ means A <u>or</u> B.
- $A \cup B = A + B - A \cap B$

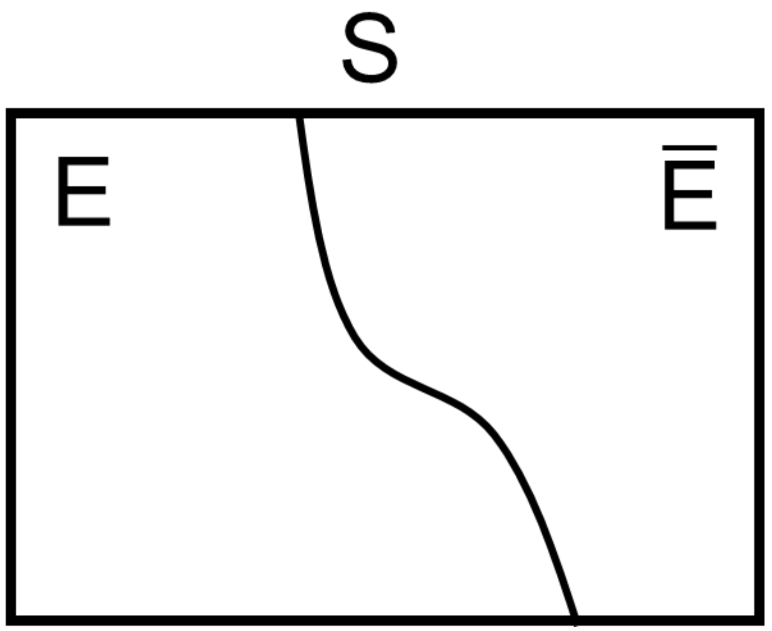


Example:

- I can choose between oatmeal (A) and cornflakes (B) for breakfast.
- The union of the events is that I had breakfast.

The Complement Event (*Komplementær*)

Notation: $S \setminus E = \overline{E} = E^c$ "not-E"



Notice:

$E \cup \bar{E} = S$

The certain event

$E \cap \bar{E} = \varnothing$

The impossible event

Example:

- The complement of having a disease is not having a disease

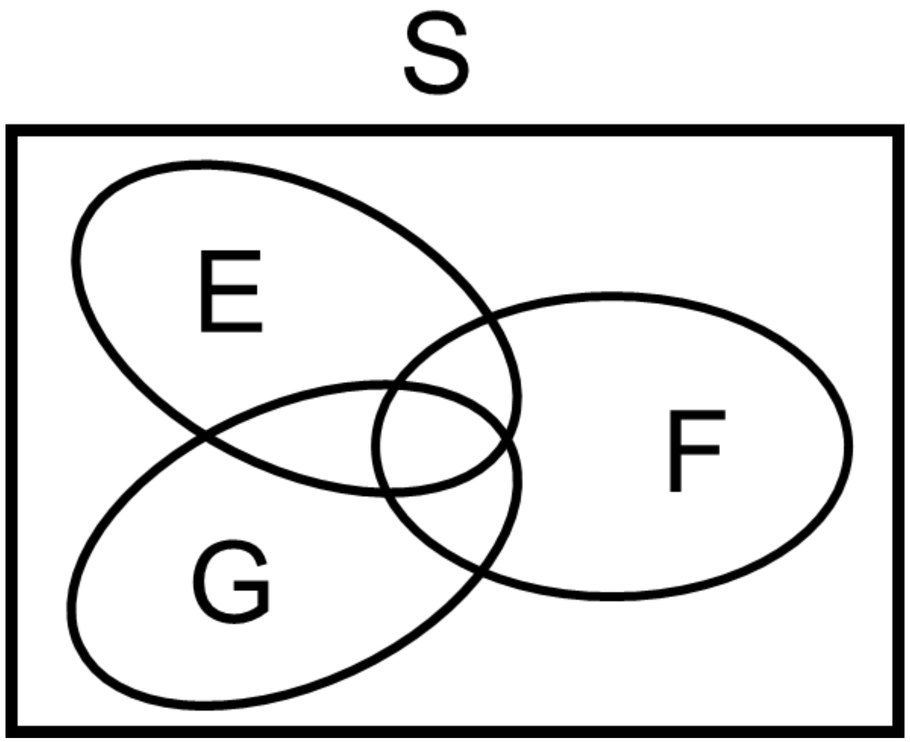
Calculation Rules for Set Theory

Commutative law

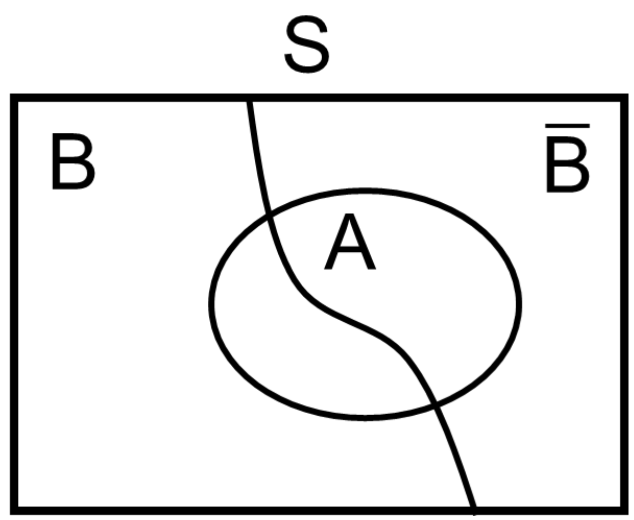
Associative law

$$E \cup F = F \cup E$$
$$E \cup (F \cup G) = (E \cup F) \cup G$$
$$(E \cup F) \cap G = (E \cap G) \cup (F \cap G)$$

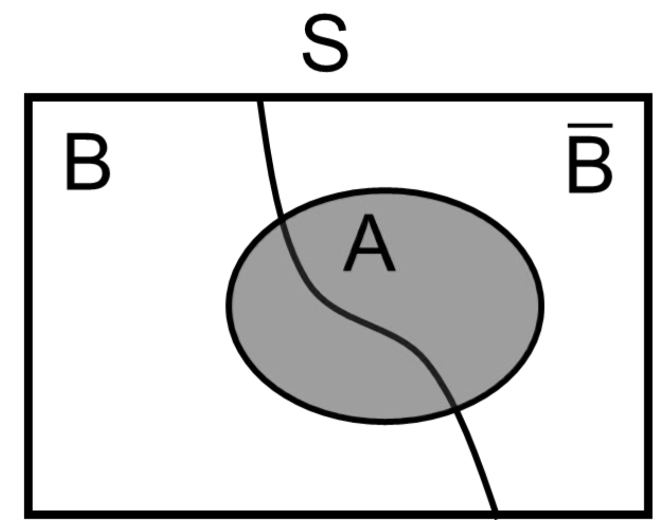
Distributive law



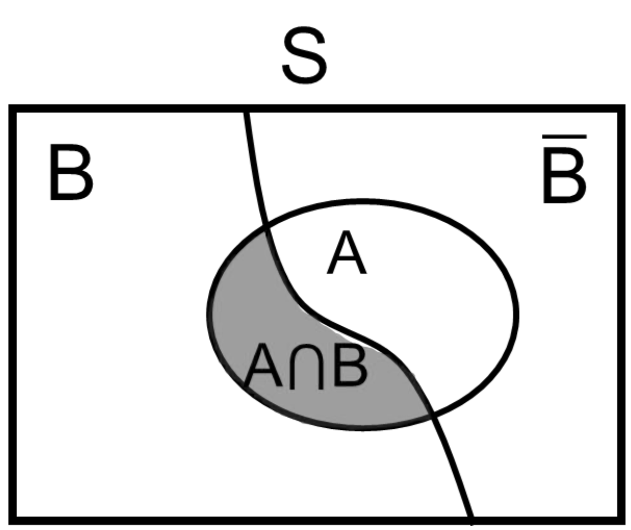
Probability of joint events



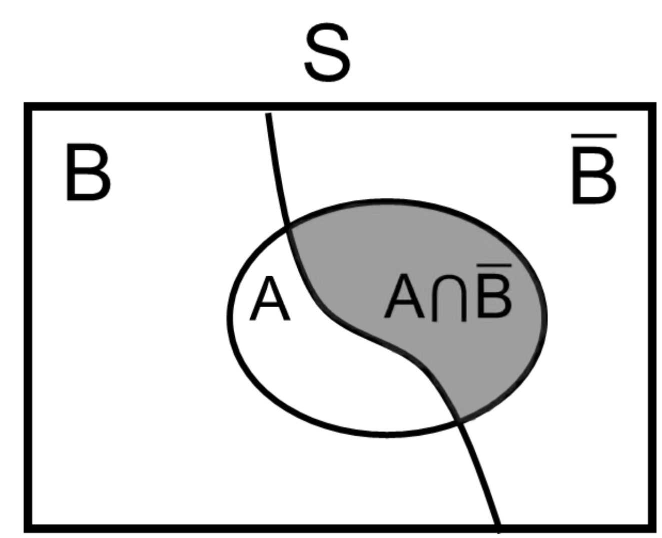
Venn diagram



$$\Pr(A) = \frac{N_A}{N_S}$$



$$\Pr(A \cap B) = \frac{N_{A \cap B}}{N_S}$$



$$\Pr(A \cap \bar{B}) = \frac{N_{A \cap \bar{B}}}{N_S}$$

Independence *(Uafhængighed)*

- We define that two events are independent if and only if:

$$\boxed{Pr(A \cap B) = Pr(A) \cdot Pr(B)}$$

Notice:

- This does not apply if the events A and B are dependent.

Example:

- Two throws with a dice
- The gender of two siblings

Conditional Probability (*Betingede sandsynligheder*)

- We write a conditional probability as:

$$\boxed{Pr(A|B)}$$

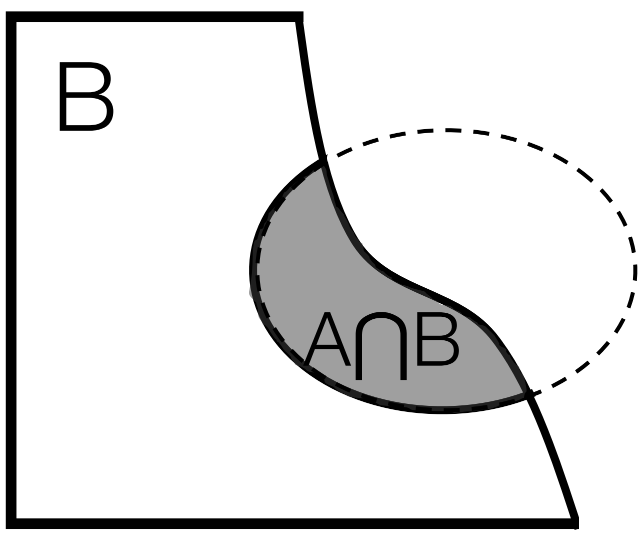
"A given B"

- This means that if the event B has already happened, what is the probability of the event A.
- Reduction of the sample space (possible events) from S to B

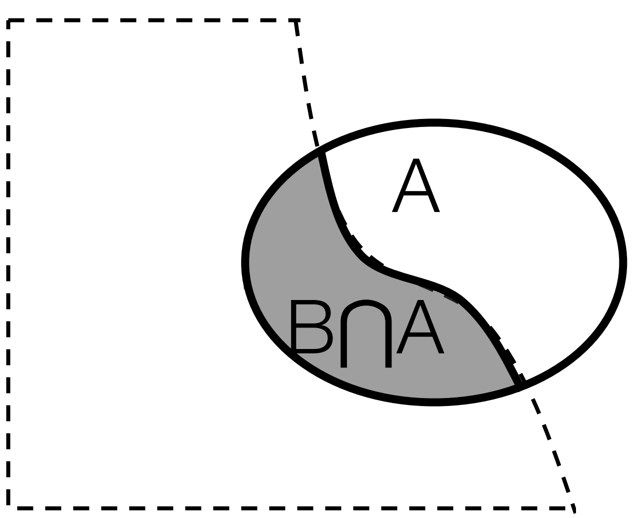
Example:

- From a population, I have selected a female.
- What is the chance that the selected person is below 1.6 m in height?

Conditional Probabilities – Bayes Rule



$$\Pr(A|B) = \frac{N_{A\cap B}}{N_B} = \frac{N_{A\cap B}/N_S}{N_B/N_S} = \frac{\Pr(A \cap B)}{\Pr(B)}$$



$$\Pr(B|A) = \frac{N_{B\cap A}}{N_A} = \frac{N_{B\cap A}/N_S}{N_A/N_S} = \frac{\Pr(B \cap A)}{\Pr(A)}$$

Probabilities of a Joint Event

- We can calculate the probability of a joint event

$$\boxed{Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = Pr(B|A) \cdot Pr(A)}$$

Notice:

- We can extend this rule to multiple events.
- Joint events are not the same as conditional events

- If A and B independent:

$$\boxed{Pr(B|A) = Pr(B) \quad \text{and} \quad Pr(A|B) = Pr(A)}$$

Very important!

Bayes Rule

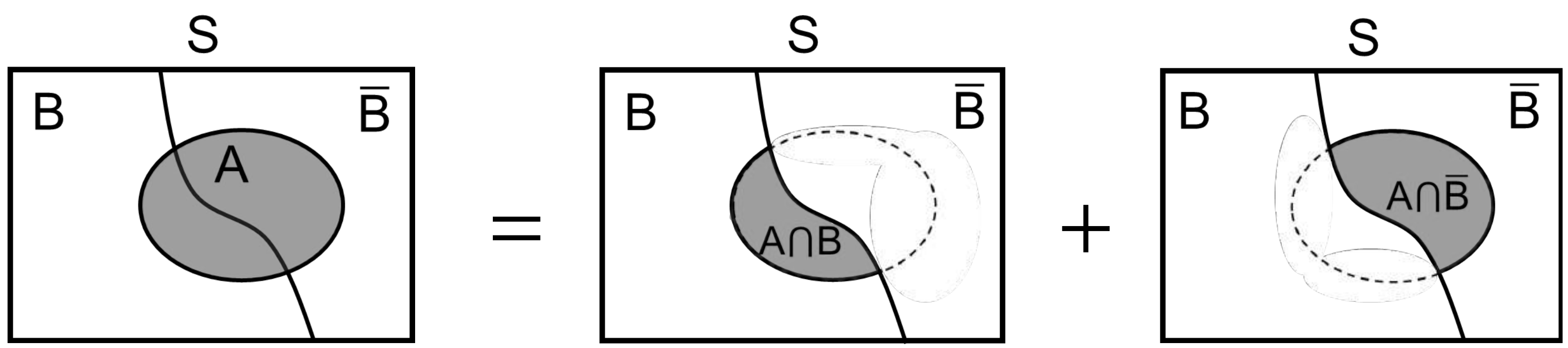
- We can write Bayes rule for two events as:

$$\boxed{Pr(B) \cdot Pr(A|B) = Pr(A) \cdot Pr(B|A)}$$

or

$$\boxed{Pr(A|B) = \frac{Pr(B|A) \cdot Pr(A)}{Pr(B)} = \frac{Pr(A \cap B)}{Pr(B)}}$$

Conditional Probabilities – Total Probability



$$\Pr(A) = \frac{N_A}{N_S} = \frac{N_{A\cap B}}{N_S} + \frac{N_{A\cap \bar{B}}}{N_S} = \frac{N_{A\cap B}}{N_B}\cdot\frac{N_B}{N_S} + \frac{N_{A\cap \bar{B}}}{N_{\bar{B}}}\cdot\frac{N_{\bar{B}}}{N_S}$$

$$\boxed{\begin{aligned}\Pr(A) &= \Pr(A\cap B) + \Pr(A\cap \bar{B}) \\ &= \Pr(A|B)\cdot\Pr(B) + \Pr(A|\bar{B})\cdot\Pr(\bar{B})\end{aligned}}$$

Conditional Probabilities - Example

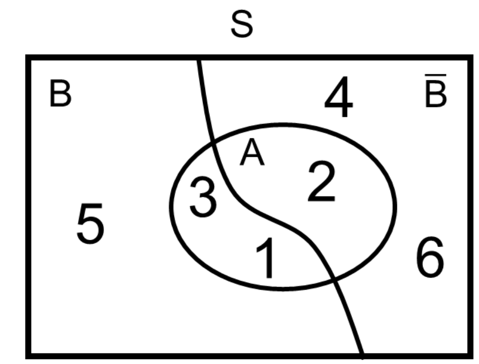
Rolling a dice:

Sample space: $S=\{1,2,3,4,5,6\}$

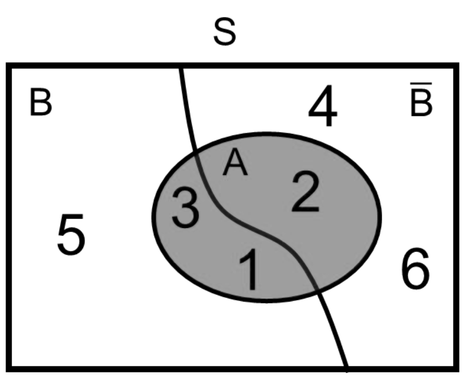
Events: $A=\{1,2,3\}$

$B=\{1,3,5\}$

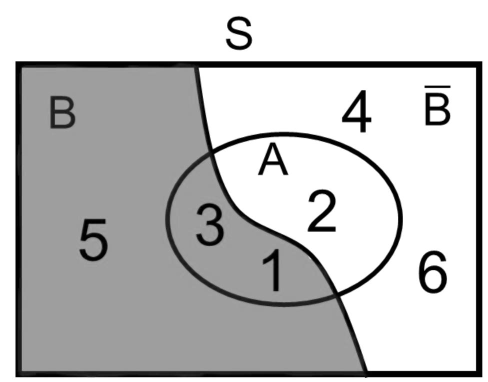
$\bar{B}=\{2,4,6\}$



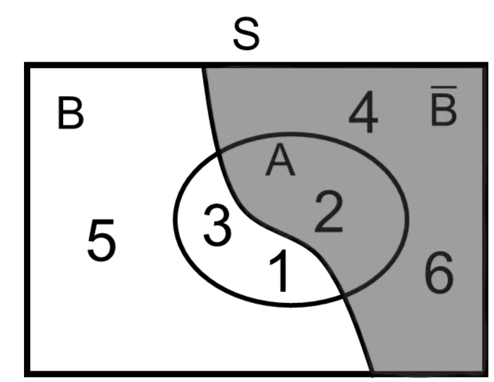
Venn diagram



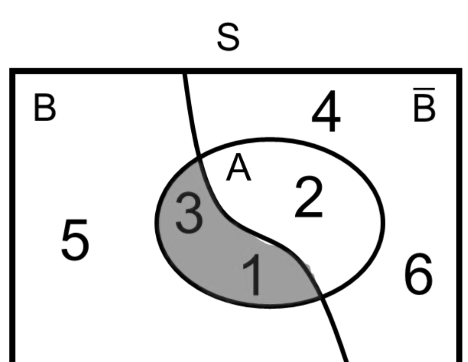
$$\Pr(A) = \frac{N_A}{N_S} = \frac{3}{6} = \frac{1}{2}$$



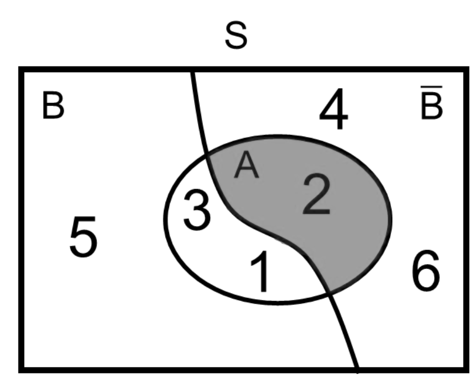
$$\Pr(B) = \frac{N_B}{N_S} = \frac{3}{6} = \frac{1}{2}$$



$$\Pr(\bar{B}) = \frac{N_{\bar{B}}}{N_S} = \frac{3}{6} = \frac{1}{2}$$

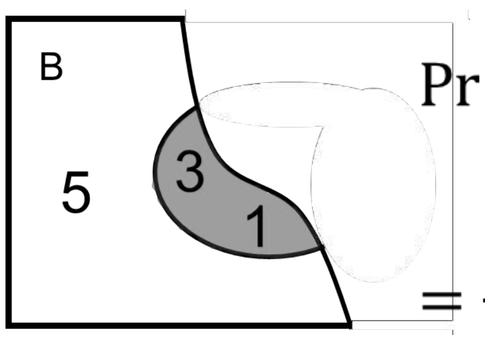


$$\Pr(A \cap B) = \frac{N_{A\cap B}}{N_S} = \frac{2}{6} = \frac{1}{3}$$



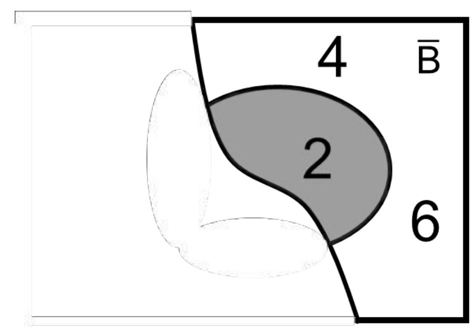
$$\Pr(A \cap \bar{B}) = \frac{N_{A\cap \bar{B}}}{N_S} = \frac{1}{6}$$

Conditional Probabilities - Example



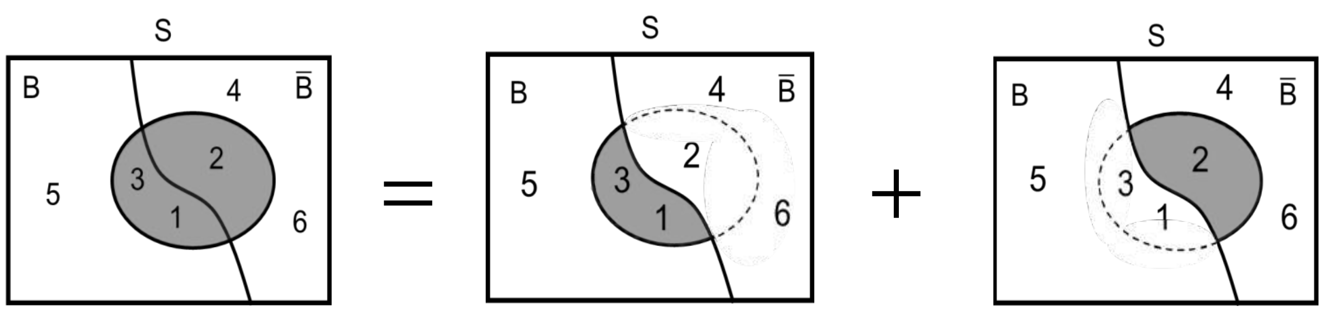
$$\Pr(A|B) = \frac{N_{A\cap B}}{N_B} = \frac{2}{3}$$

$$= \frac{\Pr(A\cap B)}{\Pr(B)} = \frac{1/3}{1/2} = \frac{2}{3}$$



$$\Pr(A|\bar{B}) = \frac{N_{A\cap \bar{B}}}{N_{\bar{B}}} = \frac{1}{3}$$

$$= \frac{\Pr(A\cap \bar{B})}{\Pr(\bar{B})} = \frac{1/6}{1/2} = \frac{2}{6} = \frac{1}{3}$$



$$\Pr(A) = \frac{N_A}{N_S} = \frac{N_{A\cap B}}{N_S} + \frac{N_{A\cap \bar{B}}}{N_S} = \frac{N_{A\cap B}}{N_B}\cdot\frac{N_B}{N_S} + \frac{N_{A\cap \bar{B}}}{N_{\bar{B}}}\cdot\frac{N_{\bar{B}}}{N_S}$$

$$= \frac{3}{6} = \frac{2}{6} + \frac{1}{6} = \frac{2}{3}\cdot\frac{3}{6} + \frac{2}{3}\cdot\frac{3}{6} = \frac{1}{2}$$

$$\Pr(A) = \Pr(A\cap B) + \Pr(A\cap \bar{B}) = \Pr(A|B)\cdot\Pr(B) + \Pr(A|\bar{B})\cdot\Pr(\bar{B}) = \frac{2}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot\frac{1}{2} = \frac{3}{6} = \frac{1}{2}$$

Markov Properties

- Bayes rule for tree events:

$$\boxed{Pr(A \cap B \cap C) = Pr(A) \cdot Pr(B|A) \cdot Pr(C|B, A)}$$

- For a Markov chain, it holds that:

$$\boxed{Pr(A \cap B \cap C) = Pr(A) \cdot Pr(B|A) \cdot Pr(C|B)}$$

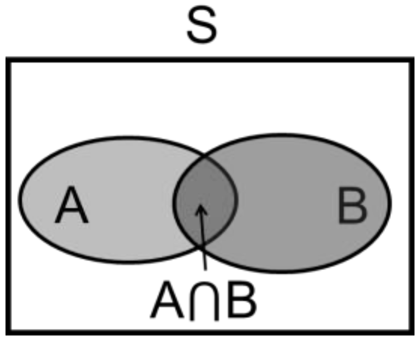
i.e. $$\boxed{Pr(C|B, A) = Pr(C|A \cap B) = Pr(C|B)}$$

(*A* don't give new information)

Probabilities of a Union of Event

- We can calculate the probability of a union of events:

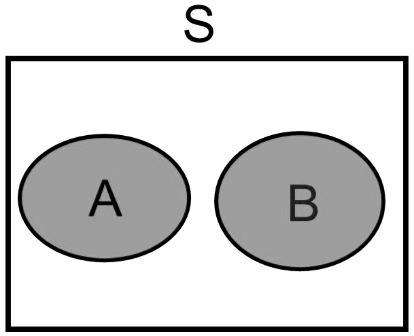
$$\boxed{Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)}$$



Notice:

- If the events are mutually exclusive

$$\boxed{Pr(A \cup B) = Pr(A) + Pr(B)}$$



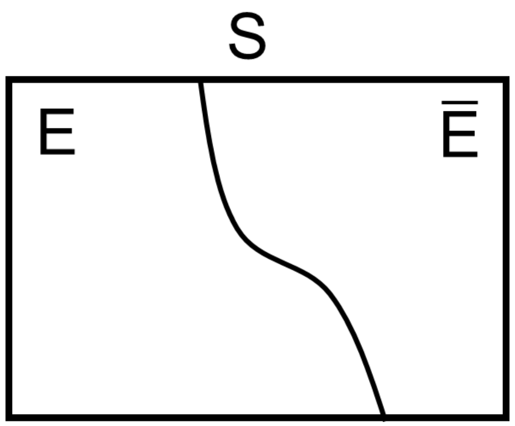
Probabilities of Complement Events

- We can write some rules for the probabilities of a complement event

$$Pr(E \cup \bar{E}) = Pr(S) = 1$$

$$Pr(E) + Pr(\bar{E}) = Pr(S) = 1$$

$$\boxed{Pr(E) = 1 - Pr(\bar{E})}$$



Example:

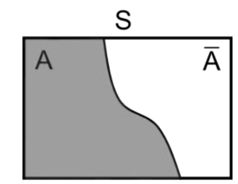
- The probability of not hitting 2 eyes on dice.

$$Pr(\{1,3,4,5,6\}) = 1 - Pr(\{2\}) = 1 - \frac{1}{6} = \frac{5}{6}$$

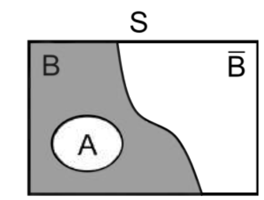
Summary of Probability

Relative frequency: $Pr(A) = \frac{N_A}{N_S}$

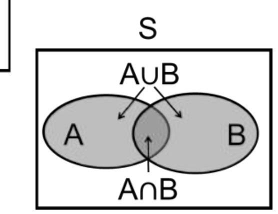
Complement: $Pr(\overline{A}) = 1 - Pr(A)$



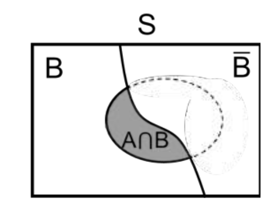
Exclusive: $Pr(\overline{A} \cap B) = Pr(B) - Pr(A)$ if $A \subset B$



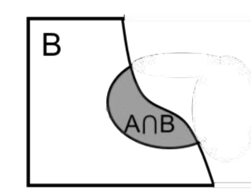
Union: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$



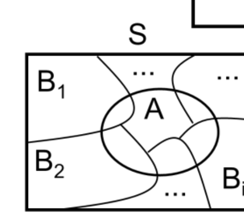
Joint: $Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = Pr(B|A) \cdot Pr(A)$



Conditional: $Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}$ if $Pr(B) \neq 0$



Total probability: $Pr(A) = \sum_{i=1}^{n} Pr(A|B_i) \cdot Pr(B_i)$



Bayes rule: $Pr(B|A) = \frac{Pr(A|B) \cdot Pr(B)}{Pr(A)}$

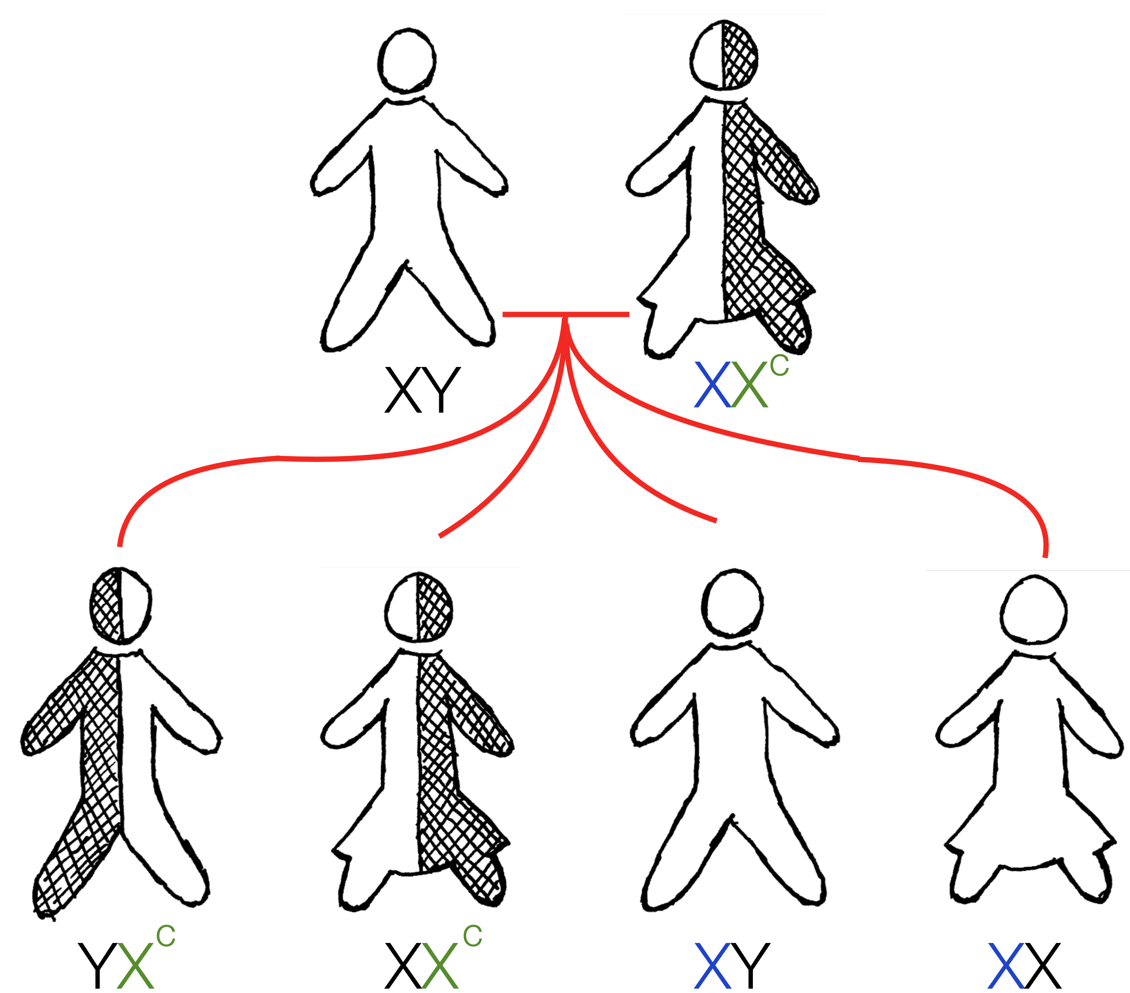
Bayes formula: $Pr(B_i|A) = \frac{Pr(A|B_i) \cdot Pr(B_i)}{\sum_{i=1}^{n} Pr(A|B_i) \cdot Pr(B_i)}$

Independence: $Pr(A \cap B) = Pr(A) \cdot Pr(B)$

Neutralized by a healthy X-gene

Example - X linked recessive

- A mother has a sick X gene.
- The chance of giving the sick X gene to a child is 50%.
- A boy with the sick X gene has the disease.
- A girl with the sick X gene is a carrier of the disease.
- Of cause the chance of not giving the sick X gene to a child is also 50%.



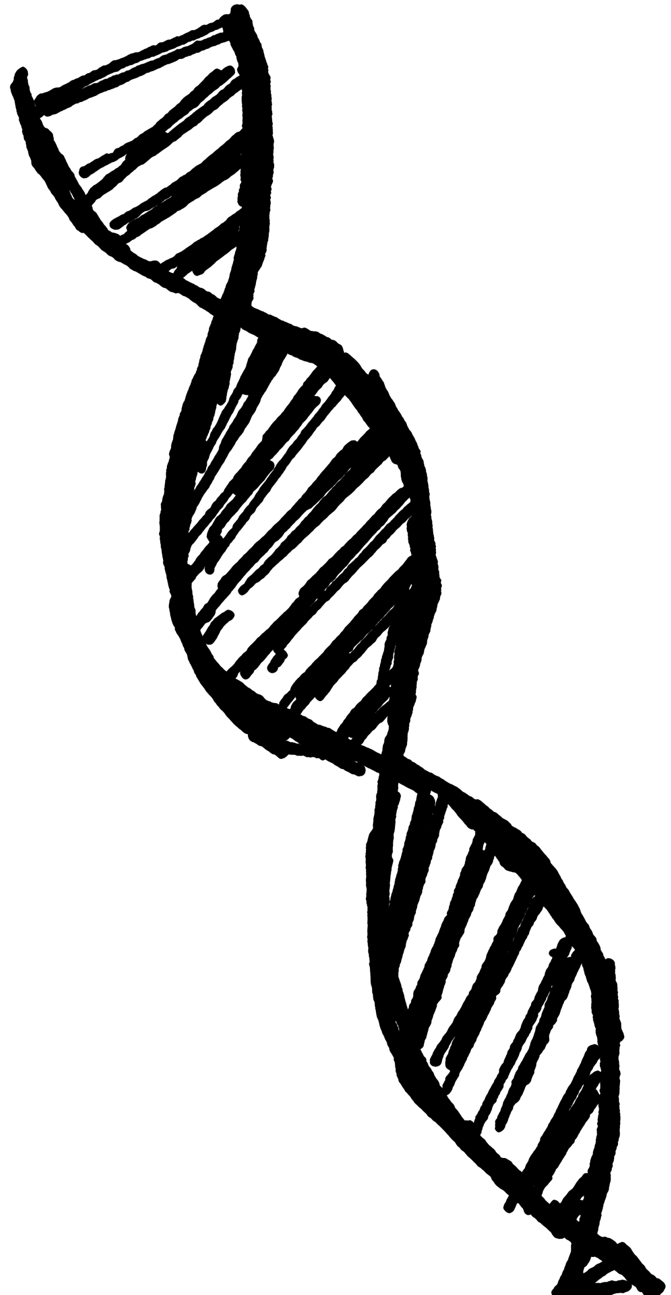
Hunter Syndrome (MPS II)

- X linked recessive
- 1:130.000 male births

- What are the probability that a boy have Hunter?

- Event A: The boy has Hunter.

$$Pr(A) = \frac{1}{130.000} = 7{,}69 \cdot 10^{-6}$$



Genetic Risk Assessment Hunters Syndrome (MPS II)

Information:

- X linked recessive.
- Boy T has Hunter.

Events:

- Event B: Mrs. B is a carrier.
- Event A: Boy A has Hunter.

Find:

- What is Pr(A)?

$Pr(B)=\frac{1}{2}; \quad Pr(A|B)=\frac{1}{2}$

$\Downarrow$

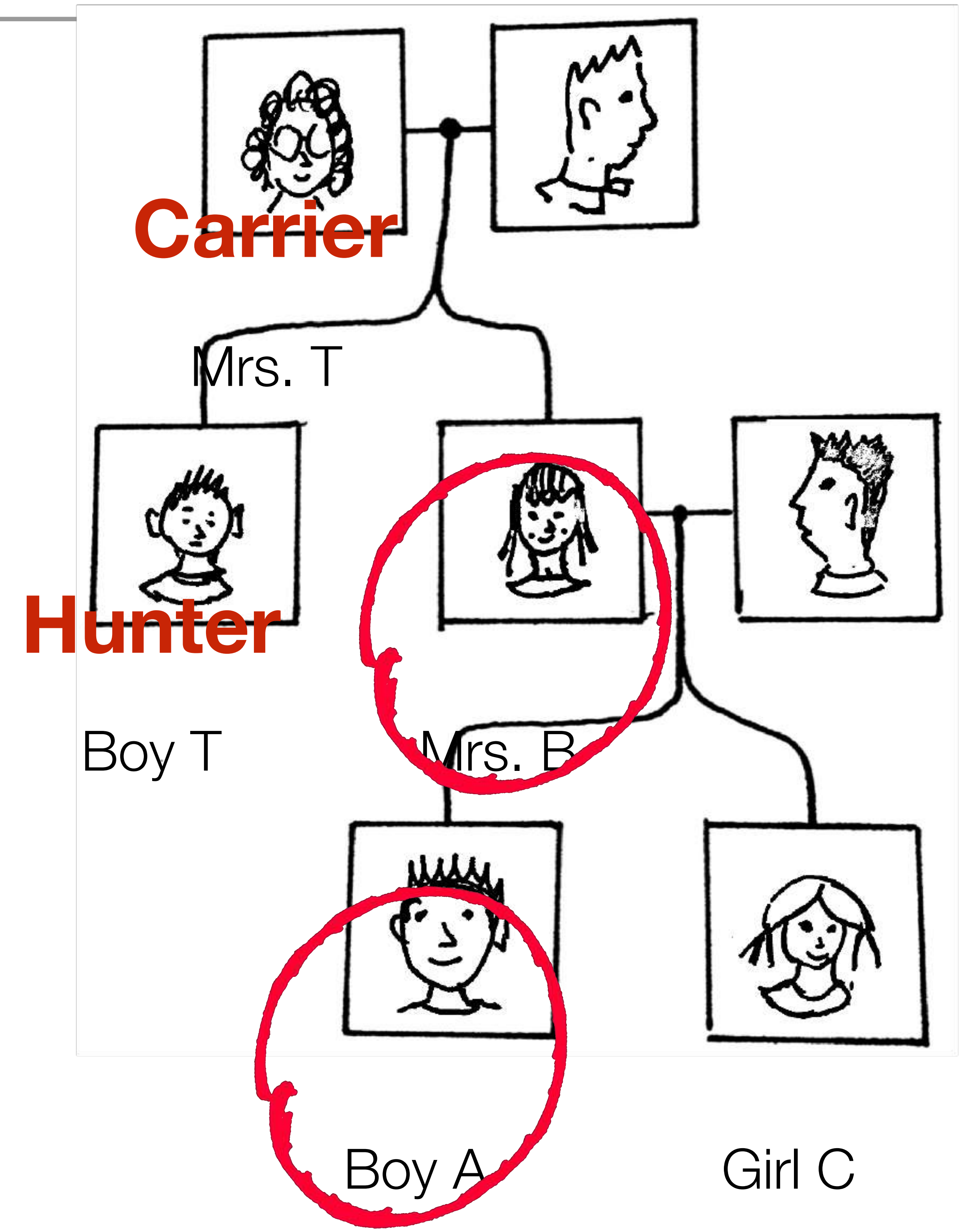
Bayes: $Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = \frac{1}{4}$

$Pr(\bar{B}) = 1 - Pr(B) = \frac{1}{2}; \quad Pr(A|\bar{B}) = 0$

$\Downarrow$

Bayes: $Pr(A \cap \bar{B}) = Pr(A|\bar{B}) \cdot Pr(\bar{B}) = 0$

$Pr(A) = Pr(A \cap B) + Pr(A \cap \bar{B}) = \frac{1}{4} + 0 = \frac{1}{4}$



Genetic Risk Assessment Hunters Syndrome (MPS II)

Information:

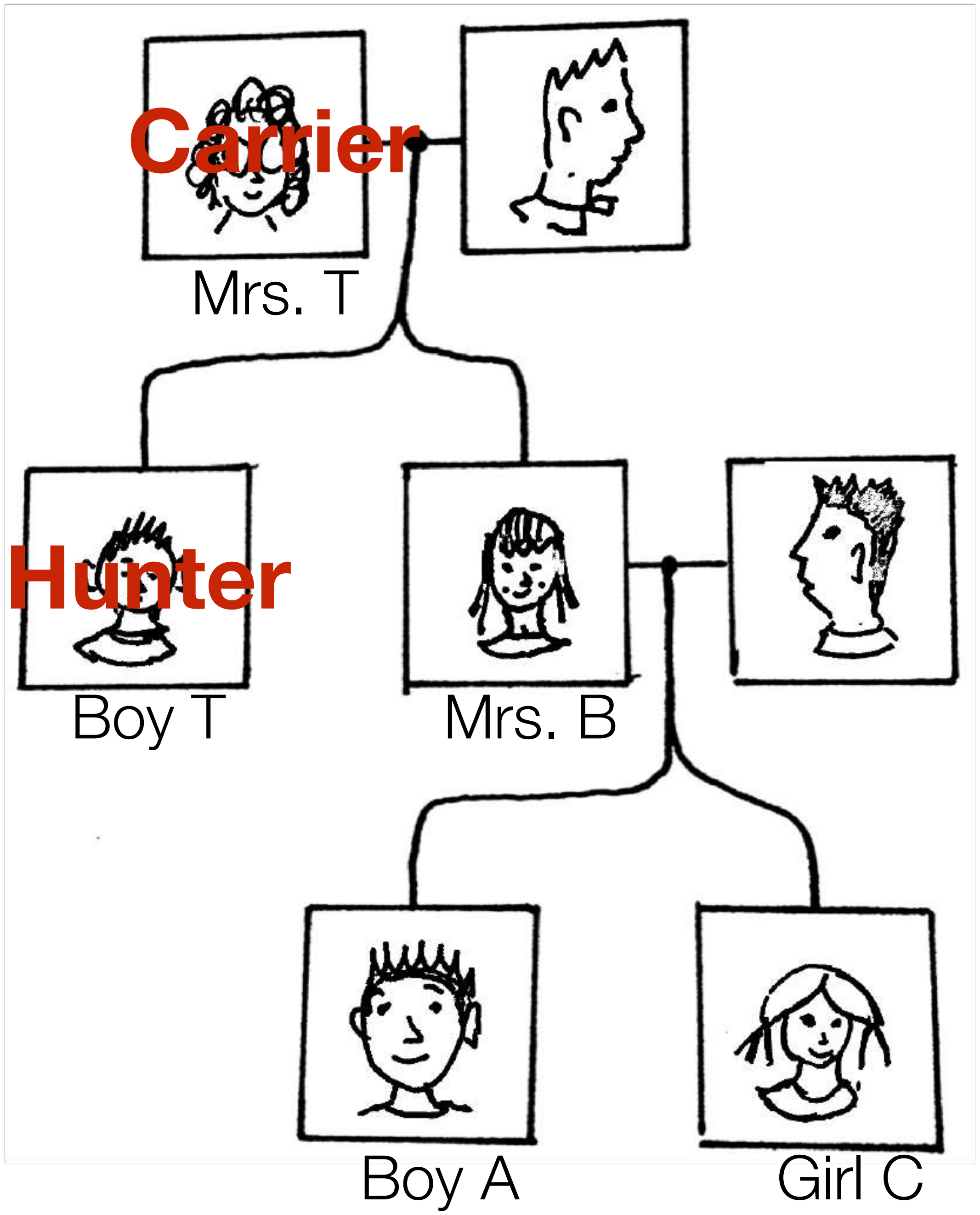
- X linked recessive.
- Boy T has Hunter.

Events:

- Event A: Boy A has Hunter.
- Event B: Mrs. B is a carrier.
- Event C: Girl C is a carrier.

Find:

- What is $\Pr(C|\bar{A})$?



Genetic Risk Assessment Hunters Syndrome (MPS II)

Events:

- Event A: Boy A has Hunter.
- Event B: Mrs. B is a carrier.
- Event C: Girl C is a carrier.

$Pr(B)=\frac{1}{2}$; $Pr(A)=\frac{1}{4}$; $Pr(\bar{A})=1-Pr(A)=\frac{3}{4}$;

$Pr(A|B)=Pr(\bar{A}|B)=\frac{1}{2}$; $Pr(A|\bar{B})=0$: $Pr(\bar{A}|\bar{B})=1$;

$Pr(C|B)=Pr(\bar{C}|B)=\frac{1}{2}$; $Pr(C|\bar{B})=0$; $Pr(\bar{C}|\bar{B})=1$;

$\Downarrow$

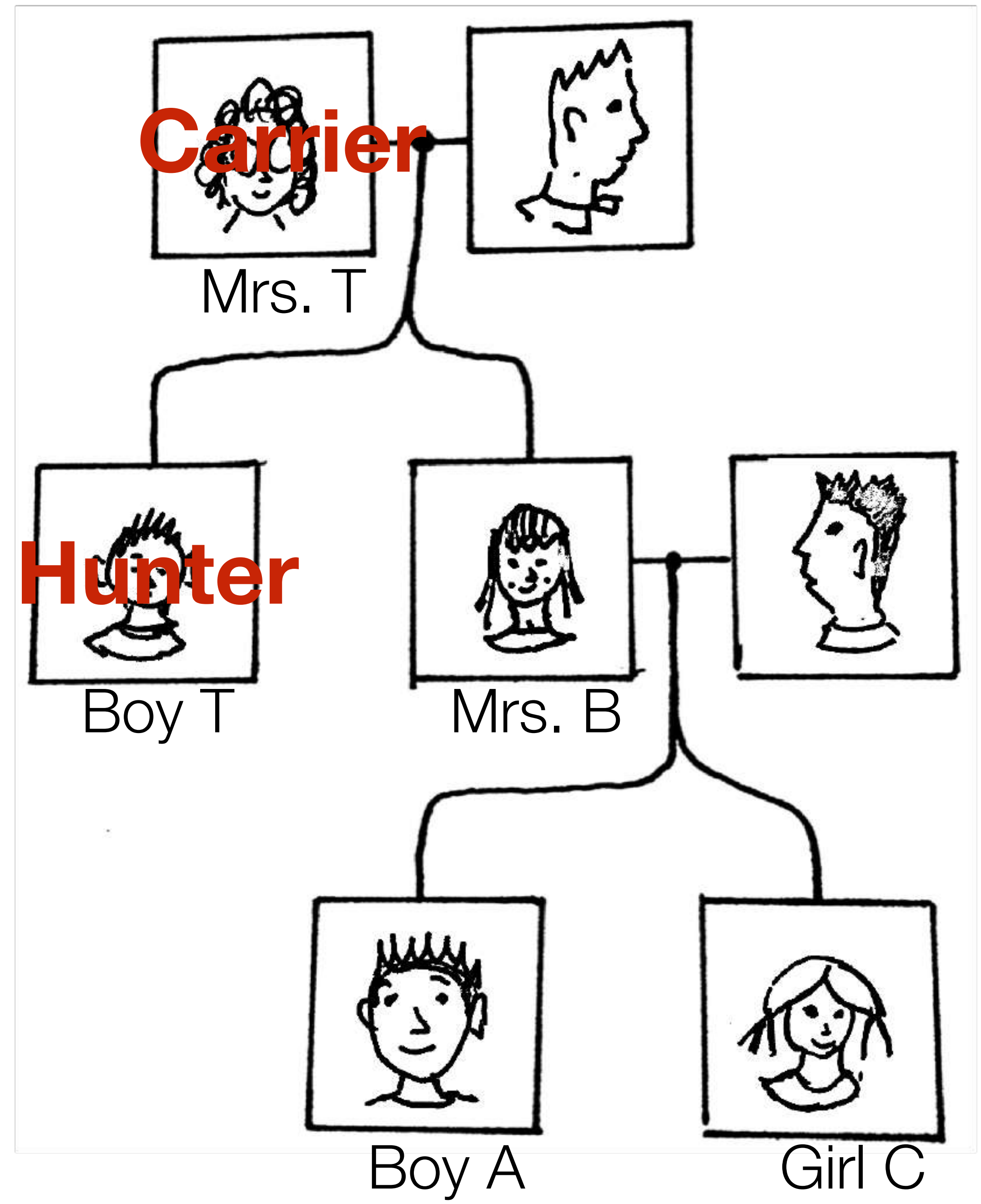
$$\text{Bayes: } Pr(B|\bar{A}) = \frac{Pr(\bar{A}|B)\cdot Pr(B)}{Pr(\bar{A})} = \frac{\frac{1}{2}\cdot\frac{1}{2}}{\frac{3}{4}} = \frac{1}{3}$$

$$\text{Markov: } Pr(C \cap B|\bar{A}) = Pr(C|B)\cdot Pr(B|\bar{A}) = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6}$$

$$Pr(C \cap \bar{B}|\bar{A}) = Pr(C|\bar{B})\cdot Pr(\bar{B}|\bar{A}) = 0$$

$\Downarrow$

$$Pr(C|\bar{A}) = Pr(C \cap B|\bar{A}) + Pr(C \cap \bar{B}|\bar{A}) = \frac{1}{6}+0=\frac{1}{6}$$



Words and Concepts to Know

Experiment/Trial

Intersection

Markov chain

Sample space

Mutually Exclusive/Disjoint

Event

Sample point

Union

Complement/not

Relative frequency

Independence

Set

Subset

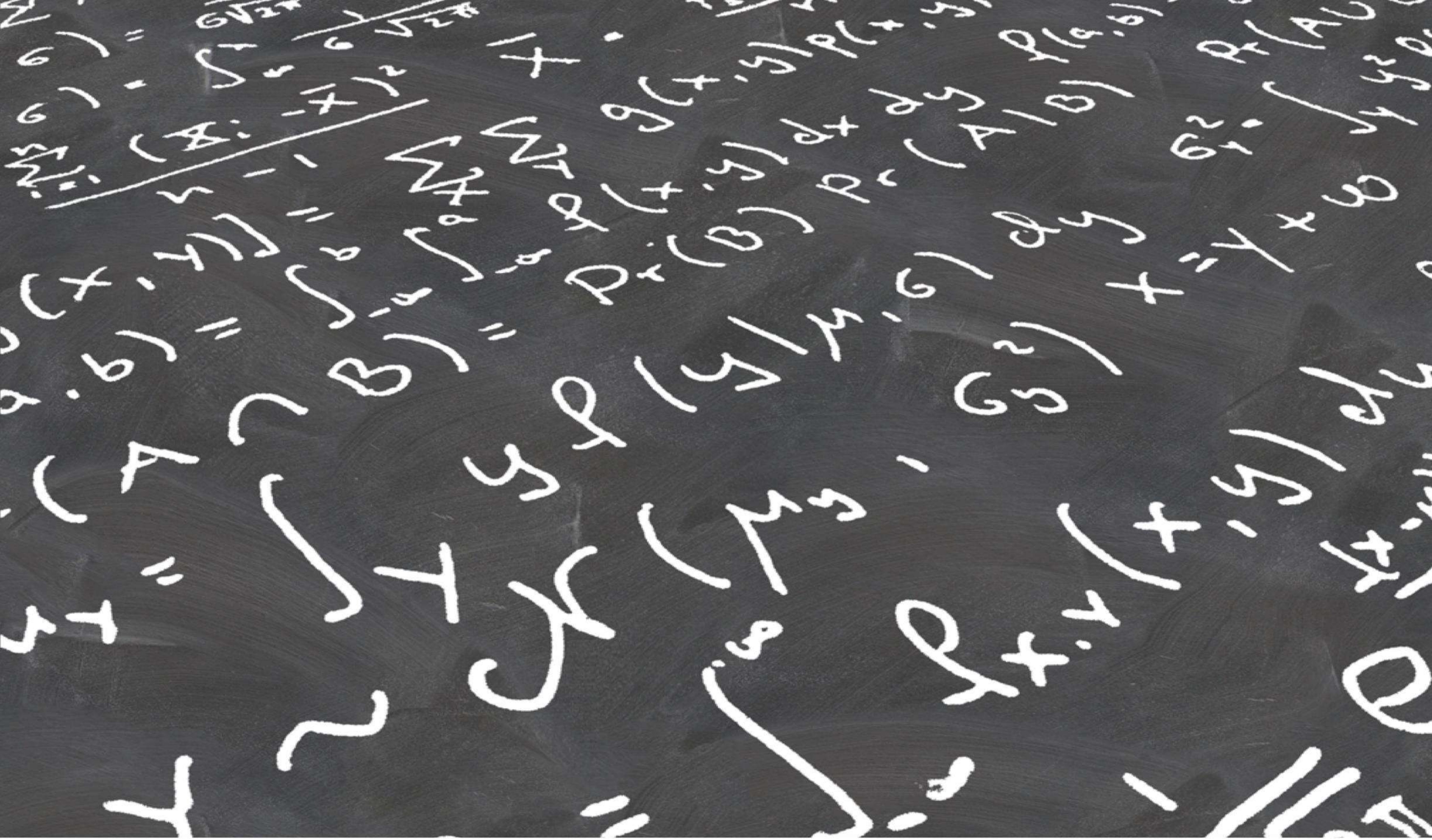
Bayes Rule

Empty set/Null set

Conditional probability

Total probability

Joint events



2. Probability Theory and Combinatorics

Gunvor Elisabeth Kirkelund
Lars Mandrup

Agenda for Today

- Repetition from last time
- Bayesian probability calculations and total probability
- Bernoulli trials
- Combinatorics
- An experiment

Basic Probability

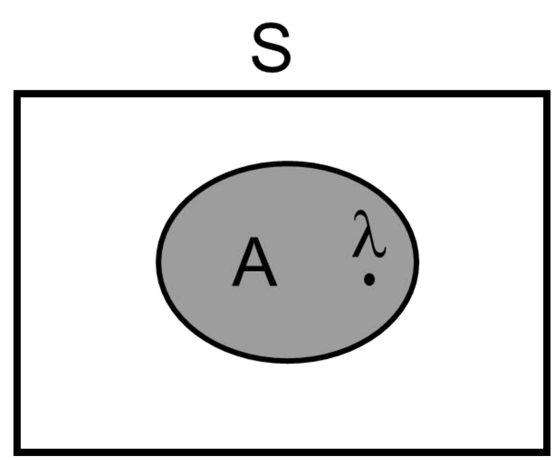
- Probability theory tells us what is in the sample given nature

- Basic Axions:

Axion 1: $\mathbf{0 \leq \mathit{Pr}(\mathit{A}) \leq 1}$

Axion 2: $\mathbf{\mathit{Pr}(\mathit{S}) = 1}$

S: Sample space
A: Event
λ: Sample point

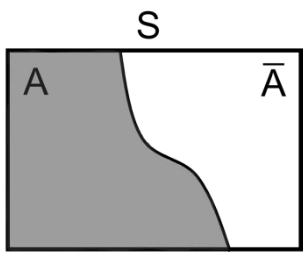


- Often (but not always) we use the relative frequency:

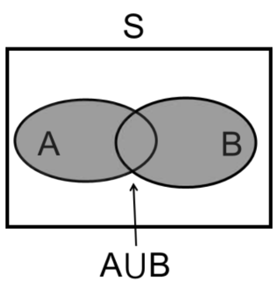
$$\Pr(A) = \frac{N_A}{N}$$

Basic Probability

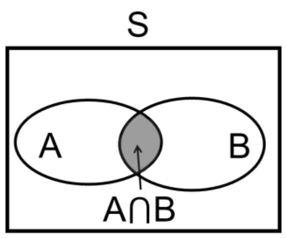
- Complement: $Pr(A) = 1 - Pr(\bar{A})$



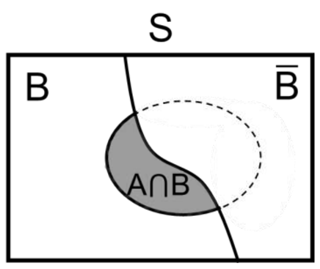
- Union: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$



- Joint: $Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = Pr(B|A) \cdot Pr(A)$



- Conditional: $Pr(A|B)$



Bayes Rule and Independence

- Bayes Rule:

$$\boxed{Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(B|A) \cdot Pr(A)}{Pr(B)}}$$

- A and B independent:

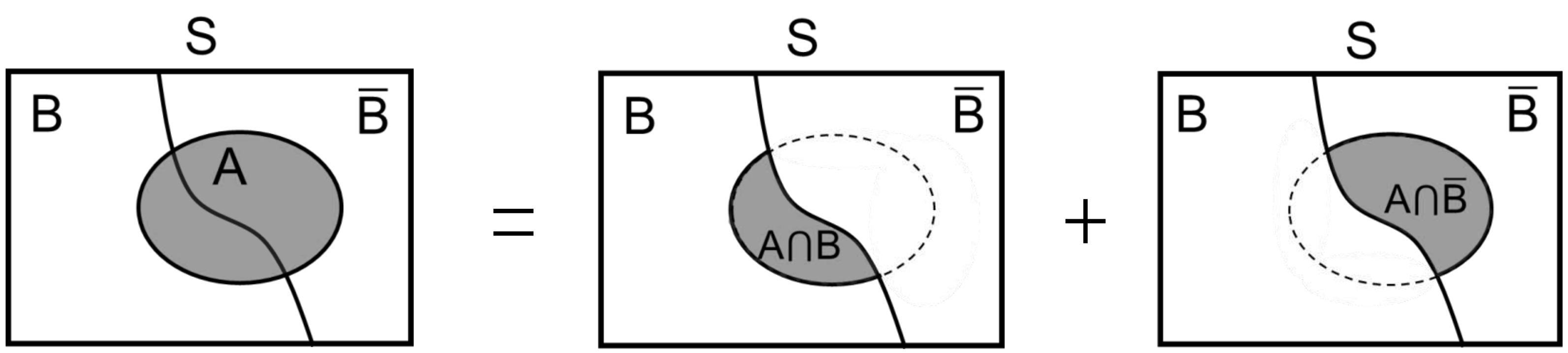
$$\boxed{Pr(A \cap B) = Pr(A) \cdot Pr(B)}$$

$$\boxed{Pr(B|A) = Pr(B) \quad \text{and} \quad Pr(A|B) = Pr(A)}$$

Total Probability

We sometime call it the marginal

- Pr(A) of an event is the total probability of that event.

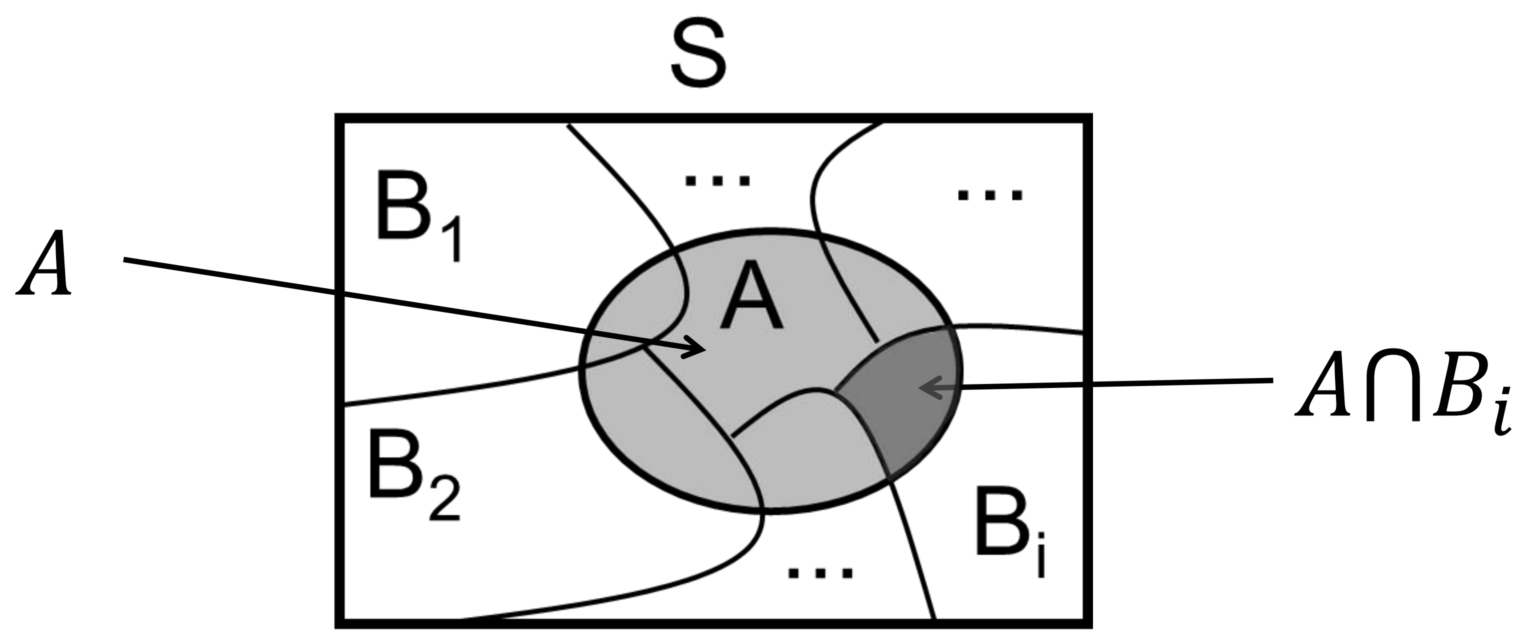


$$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap \bar{B})$$
$$= \Pr(A|B) \cdot \Pr(B) + \Pr(A|\bar{B}) \cdot \Pr(\bar{B})$$

Total Probability

We sometime call it the marginal

- Pr(A) of an event is the total probability of that event.



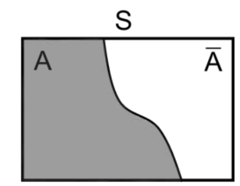
$$Pr(A) = Pr(A \cap B_1) + Pr(A \cap B_2) + \cdots + Pr(A \cap B_i) + \cdots$$
$$= Pr(A|B_1) \cdot Pr(B_1) + Pr(A|B_2) \cdot Pr(B_2) + \cdots$$

where the B_i's are mutually exclusive ($B_i \cap B_j = \varnothing$ for $i \neq j$)
and $S = B_1 \cup B_2 \cup \ldots \cup B_i \cup \ldots$

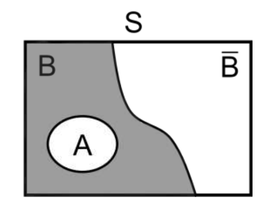
Summary of Probability

Relative frequency: $Pr(A) = \frac{N_A}{N_S}$

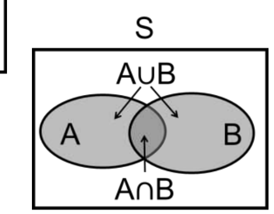
Complement: $Pr(\bar{A}) = 1 - Pr(A)$



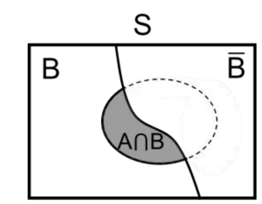
Exclusive: $Pr(\bar{A} \cap B) = Pr(B) - Pr(A)$ if $A \subset B$



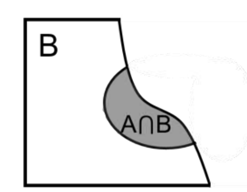
Union: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$



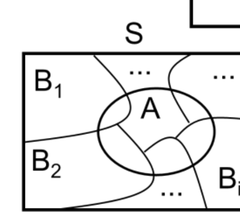
Joint: $Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = Pr(B|A) \cdot Pr(A)$



Conditional: $Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}$ if $Pr(B) \neq 0$



Total probability: $Pr(A) = \sum_{i=1}^{n} Pr\left(A \middle| B_i\right) \cdot Pr(B_i)$

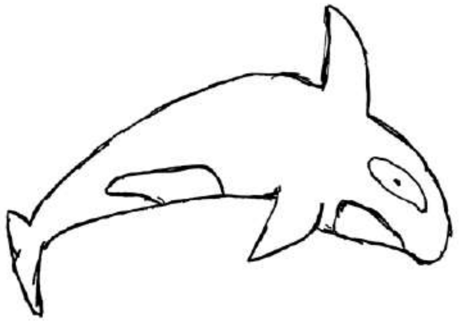


Bayes rule: $Pr(B|A) = \frac{Pr(A|B) \cdot Pr(B)}{Pr(A)}$

Bayes formula: $Pr(B_i|A) = \frac{Pr\left(A \middle| B_i\right) \cdot Pr(B_i)}{\sum_{i=1}^{n} Pr\left(A \middle| B_i\right) \cdot Pr(B_i)}$

Independence: $Pr(A \cap B) = Pr(A) \cdot Pr(B)$

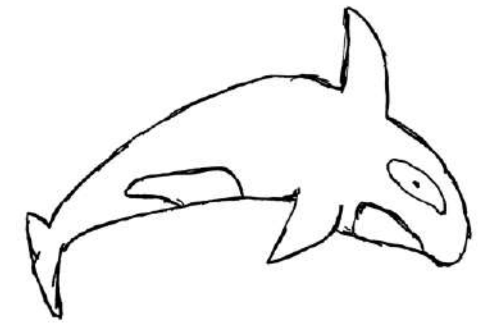
Orca Example



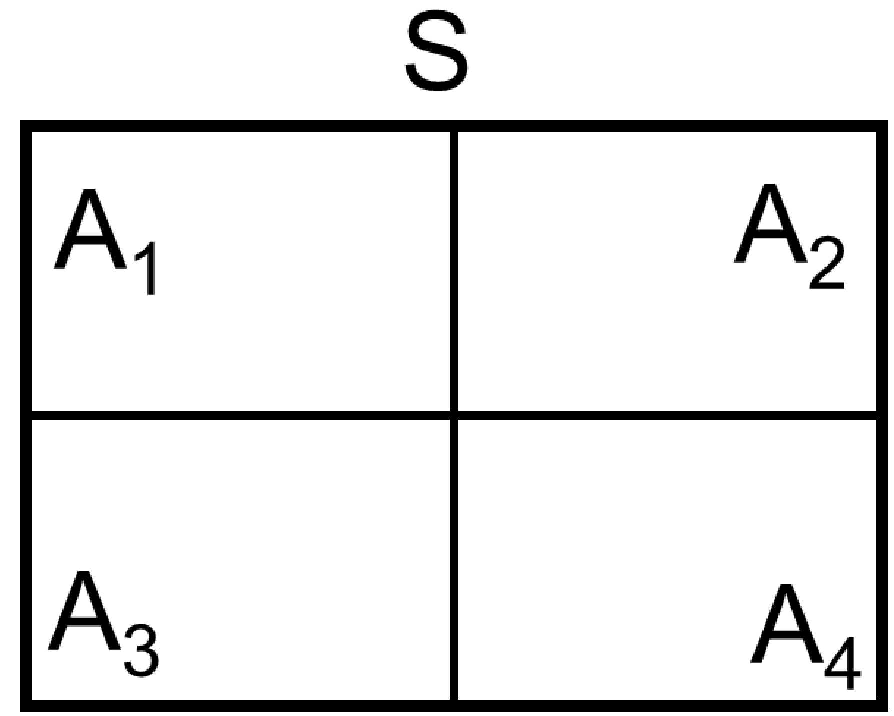
- In a conversation effort, we look for dead orcas when we are visiting an ocean.
- Given (conditioned) that <u>we have selected an ocean to examine</u>, how many males and females orcas will we observe?

Gender\ location	Atlantic (A_1)	Antartica (A_2)	Pacific (A_3)	Seaworld (A_4)
Female ($\overline{\mathbf{B}}$)	2	7	11	9
Male ($\mathbf{B}$)	8	3	1	19
Total	10	10	12	28

Orca Example (Cont'd)



- The probability selecting an ocean is identical.



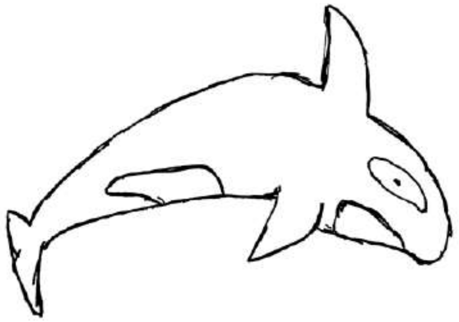
- Event A_1 : Atlantic
- Event A_2 : Antartica
- Event A_3 : Pacific
- Event A_4 : Seaworld

$$Pr(A_1) = Pr(A_2) = Pr(A_3) = Pr(A_4) = \frac{1}{4}$$

$$Pr(A_1) + Pr(A_2) + Pr(A_3) + Pr(A_4) = 1$$

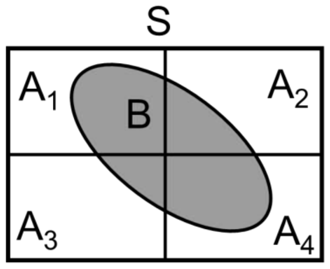
The events $A_1 - A_4$ are mutually exclusive.

Orca Example Total Probability



- The event B, that the orca is a male, can then be written as:

$$B = (B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3) \cup (B \cap A_4)$$



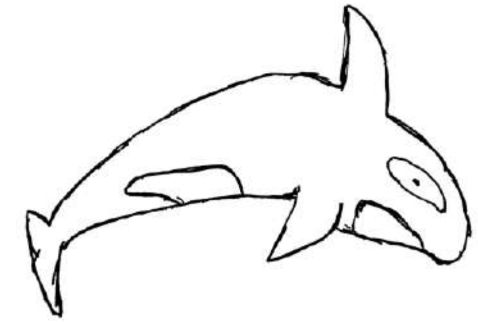
- The total probability of a found killer whale, being a male, since event $A_1 - A_4$ are mutually exclusive (sum rule):

$$\boxed{Pr(B) = Pr(B \cap A_1) + Pr(B \cap A_2) + Pr(B \cap A_3) + Pr(B \cap A_4)}$$

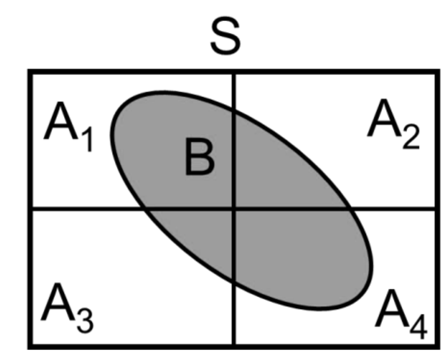
- We rewrite with Bayes rule:

$$\begin{aligned} Pr(B) = {} & Pr(A_1)\, Pr(B|A_1) + Pr(A_2)\, Pr(B|A_2) \\ & + Pr(A_3)\, Pr(B|A_3) + Pr(A_4)\, Pr(B|A_4) \end{aligned}$$

Orca Example Cont'd



- Total Probability:

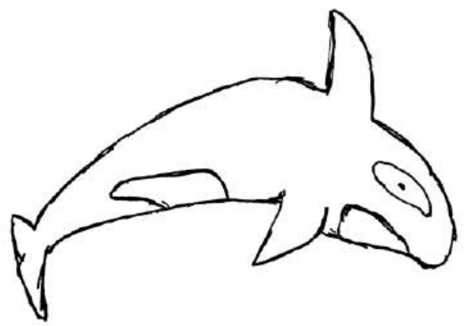


$$Pr(B) = Pr(A_1)\,Pr(B|A_1) + Pr(A_2)\,Pr(B|A_2) + Pr(A_3)\,Pr(B|A_3) + Pr(A_4)\,Pr(B|A_4)$$

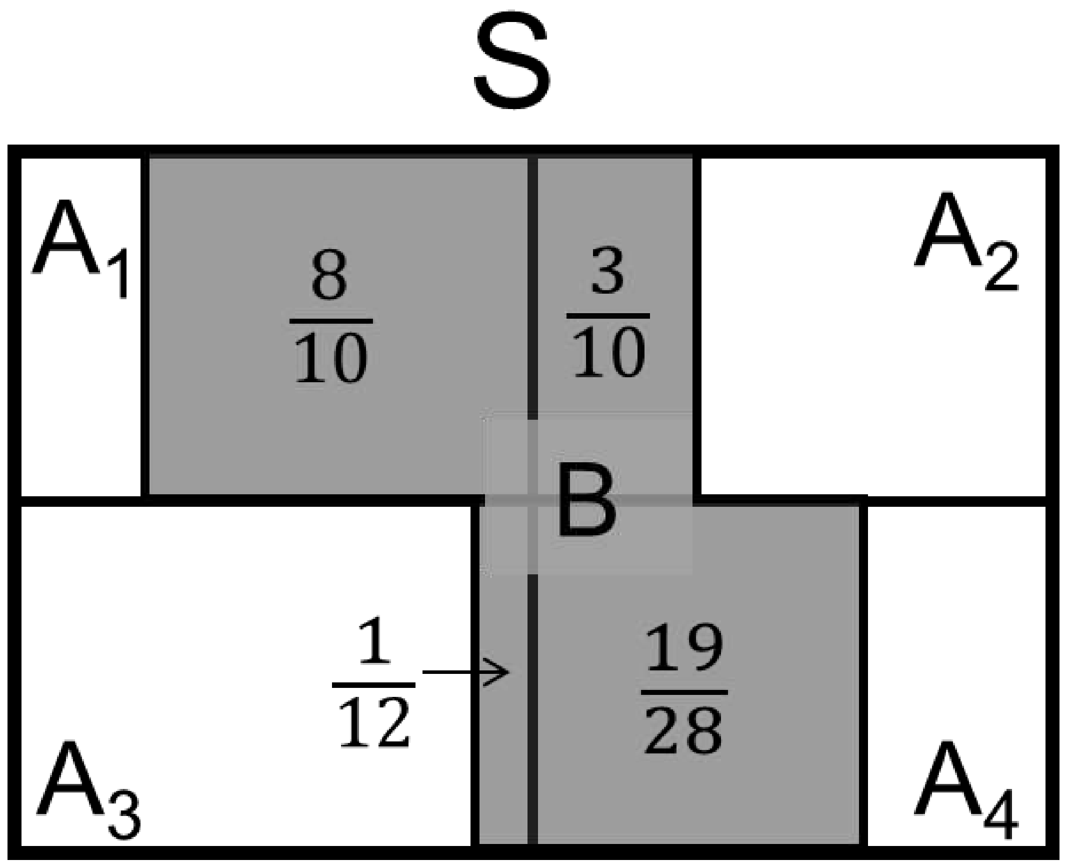
Gender\ location	Atlantic (A_1)	Antartica (A_2)	Pacific (A_3)	Seaworld (A_4)
Female ($\bar{\mathbf{B}}$)	2	7	11	9
Male ($\mathbf{B}$)	8	3	1	19
Total	10	10	12	28

$$Pr(B) = \frac{8}{10}\cdot\frac{1}{4} + \frac{3}{10}\cdot\frac{1}{4} + \frac{1}{12}\cdot\frac{1}{4} + \frac{19}{28}\cdot\frac{1}{4} = 0,465$$

Orca Example Graphical

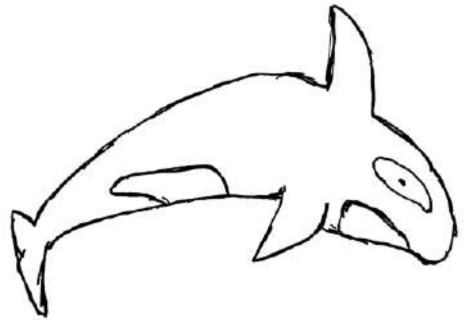


- We can also use a Graphical approach with Venn diagrams.



- The total probability of B is given by the marked area divided by the area of S.

Orca Example



- If an orca found is a male, what is the probability of us being in the Antartica?

$$Pr(A_2|B)$$

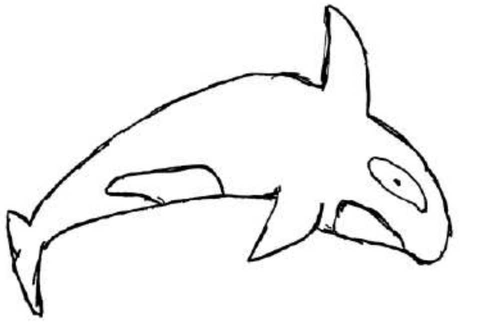
- We use Bayes rule:

$$Pr(A_2|B) = \frac{Pr(A_2 \cap B)}{Pr(B)} = \frac{Pr(B|A_2)Pr(A_2)}{Pr(B)}$$

- $Pr(B) = 0{,}47;\ \ Pr(A_2) = 0{,}25;\ \ Pr(B|A_2) = 0{,}3$

$$Pr(A_2|B) = \frac{Pr(B|A_2)Pr(A_2)}{Pr(B)} = \frac{0{,}3 \cdot 0{,}25}{0{,}47} = 0{,}16$$

Orca Example

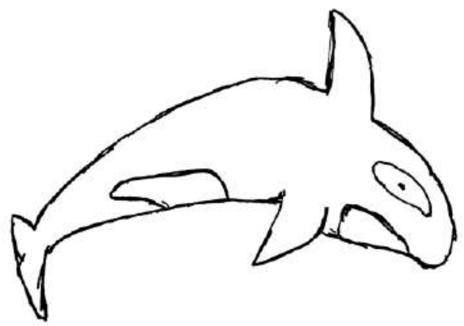


- Is locations of the found orca independent of gender?
- How would you test it?

Gender\ location	Atlantic (A_1)	Antartica (A_2)	Pacific (A_3)	Seaworld (A_4)
Female ($\bar{\mathbf{B}}$)	2	7	11	9
Male ($\mathbf{B}$)	8	3	1	19
Total	10	10	12	28

$$Pr(A_2 \mid \bar{B}) = \frac{Pr(\bar{B}|A_2)Pr(A_2)}{Pr(\bar{B})} = \frac{0{,}7 \cdot 0{,}25}{1 - 0{,}47} = 0{,}33 \neq 0{,}16 = Pr(A_2|B)$$

Orca Example Conclusion



- Prior: What is the probability of us being in the Antartica?

$$Pr(A_2) = 0{,}25$$

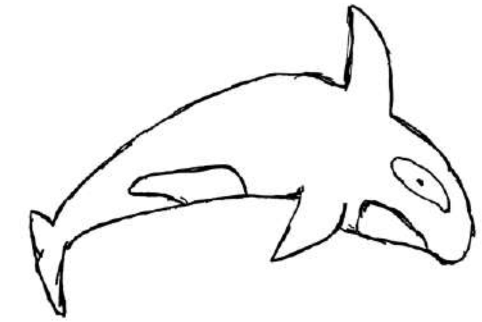
- Likelihood: A tacked orca is found dead in Antartica, what is the probability of it being male?

$$Pr(B|A_2) = 0{,}3$$

- Posterior: A tacked orca whale is found dead and is a male, what is the probability of us being in Antartica?

$$Pr(A_2|B) = 0{,}16$$

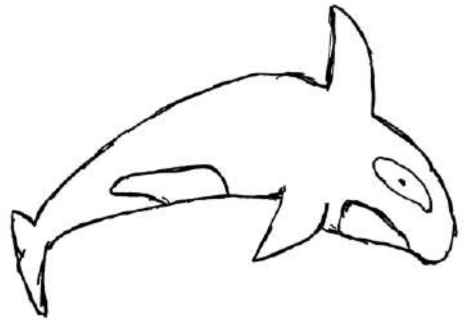
Orca Example – Another test method



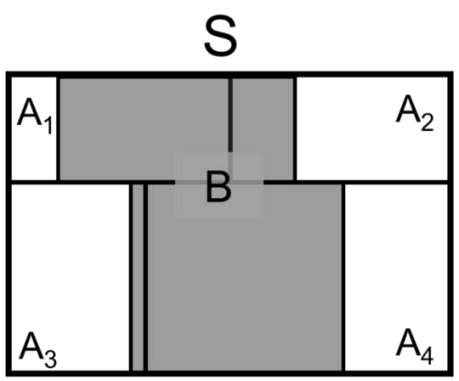
- In a conversation effort, we pick up dead orcas from different oceans.
- The dead orcas are marked with the ocean and collected in the same container.
- A dead orca is randomly picked from the container: What is the probability that the orca is a male?

Gender\ location	Atlantic (A_1)	Antartica (A_2)	Pacific (A_3)	Seaworld (A_4)
Female ($\overline{\mathbf{B}}$)	2	7	11	9
Male ($\mathbf{B}$)	8	3	1	19
Total	10	10	12	28

Orca Example – Another test method



- Total Probability:

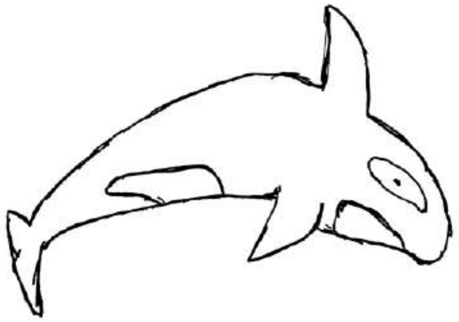


$$Pr(B) = Pr(A_1)\,Pr(B|A_1) + Pr(A_2)\,Pr(B|A_2) + Pr(A_3)\,Pr(B|A_3) + Pr(A_4)\,Pr(B|A_4)$$

Gender\ location	Atlantic (A_1)	Antartica (A_2)	Pacific (A_3)	Seaworld (A_4)	Total
Female ($\overline{\mathbf{B}}$)	2	7	11	9	29
Male ($\mathbf{B}$)	8	3	1	19	31
Total	10	10	12	28	60

$$Pr(B) = \frac{10}{60}\cdot\frac{8}{10}+\frac{10}{60}\cdot\frac{3}{10}+\frac{12}{60}\cdot\frac{1}{12}+\frac{28}{60}\cdot\frac{19}{28} = \frac{8+3+1+19}{60} = \frac{31}{60} = 0{,}517$$

Orca Example – Another test method



- If an orca found is a male, what is the probability that it is from the Antartica?

$$Pr(A_2|B)$$

- We use Bayes rule:

$$Pr(A_2|B) = \frac{Pr(A_2 \cap B)}{Pr(B)} = \frac{Pr(B|A_2)Pr(A_2)}{Pr(B)}$$

- $Pr(B) = 0{,}517;\ \ Pr(A_2) = 0{,}167;\ \ Pr(B|A_2) = 0{,}3$

$$Pr(A_2|B) = \frac{Pr(B|A_2)Pr(A_2)}{Pr(B)} = \frac{0{,}3 \cdot 0{,}167}{0{,}517} = \frac{3}{31} = 0{,}097$$

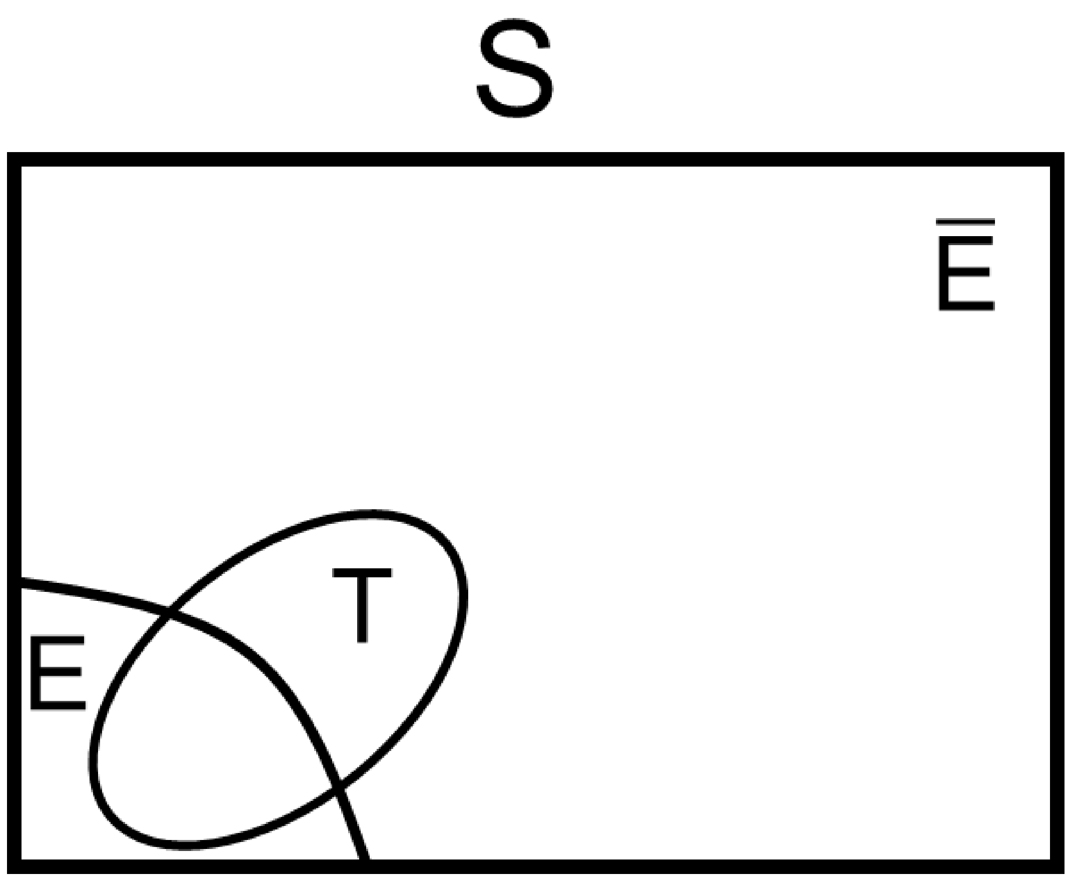
Tests and Types of Errors

- We can classify testing with two outcomes as:

Given / Result	Disease (True)	No disease (False)
Positive test	Sensitivity	Type I Error
Negative test	Type II Error	Specificity

Example: Ebola Test

- Event E: Patient are infectious with Ebola.
- Event T: The Ebola test is positive.



Example: Ebola Test

- Prior: What are the probability of a patient having Ebola?

$$Pr(E)$$

- Likelihood: What are the probability of a positive test given infectious with Ebola? Or of a negative test given not infectious with Ebola?

$$Pr(T|E) \quad \text{Sensitivity} \qquad\qquad Pr(\bar{T}|\bar{E}) \quad \text{Specificity}$$

- Posterior: What are the probability of being infectious given that a test is positive?

$$Pr(E|T)$$

Example: Ebola Test — Total Probability

- Prior: What are the probability of a patient having ebola?

Complement of E

$$Pr(E) = 0,01 \qquad Pr(\bar{E}) = 1 - 0,01 = 0,99$$

- Likelihood: What are the probabilities of the tests?

$$Pr(T|E) = 0,9$$ Sensitivity

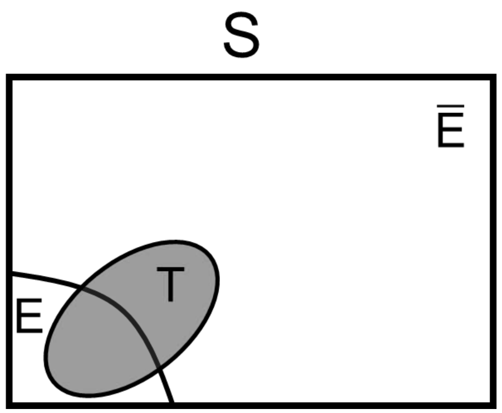
$$Pr(\bar{T}|\bar{E}) = 0,8$$ Specificity

- Complement: What are the probability of a patient having a positive test without being infectious?

$$Pr(T|\bar{E}) = 1 - Pr(\bar{T}|\bar{E}) = 0,2$$

Example: Ebola Test — Total Probability

- **Total Probability with the Sum Rule**: What are the probability of a patient having a positive test?



$$Pr(T) = Pr(T \cap E) + Pr(T \cap \bar{E})$$

- **The Product Rule**: We can with Bayes rule find

$$Pr(T) = Pr(T|E)\,Pr(E) + Pr(T|\bar{E})\,Pr(\bar{E})$$
$$= 0{,}9 \cdot 0{,}01 + 0{,}2 \cdot 0.99$$
$$= 0{,}207$$

Ebola Example — Posterior

- We have: We now know the probabilities:

$$P(E) = 0,01$$ Prior

$$P(T) = 0,207$$ Total probability

$$P(T|E) = 0,9$$ Likelihood

- Product Rule: What are the probability of being infectious given that a test is positive?

Bayes rule

$$Pr(E|T) = \frac{Pr(T|E)Pr(E)}{Pr(T)} = \frac{0,9 \cdot 0,01}{0,207} = 0,043$$

Ebola Example — Posterior

- What are the probability of being infectious given that a test is positive?

$$Pr(E|T) = \frac{Pr(T|E)Pr(E)}{Pr(T)} = \frac{0,9 \cdot 0,01}{0,207} = 0,043$$

- What are the probability of not being infectious given that a test is positive?

$$Pr(\bar{E} \mid T) = 1 - Pr(E|T) = 0,957$$

- What are the probability of not being infectious given a negative test?

$$Pr(\bar{E}|\bar{T}) = \frac{Pr(\bar{T} \mid \bar{E})Pr(\bar{E})}{Pr(\bar{T})} = \frac{0,8 \cdot 0,99}{0,793} = 0,999$$

- What are the probability of being infectious given that a test is negative?

$$Pr(\mathrm{E} \mid \bar{T}) = 1 - Pr(\bar{E}|\bar{T}) = 0,001$$

Ebola Example — Conclusion

- If the test is negative, it is allmost certain (99,9%) that you're not being infectious:

$$Pr(\bar{E}|\bar{T}) = 0{,}999$$

- If the test is positive, there is still only a small risk (4,3%) that you actually are being infectious:

$$Pr(E|T) = 0{,}043$$

Monty Hall Dilemma

- We have three doors
- Behind two of the doors is a goat
- Behind one door is a million dollars ($)

- What is the chance of guessing behind which door the money is?

$$\Pr(\$|1) = \Pr(\$|2) = \Pr(\$|3) = \frac{1}{3}$$

Monty Hall Dilemma cont'd

- We make a selection of a door, say door 2, without open it.
- The quizmaster eliminates one of the doors ($\bar{\$}$) ,which we did not select, based on his knowledge on the goat situation, say door 1.
- We can now reselect between door 2 and 3.
- What are the probabilities of the money being behind the two doors? Should we switch door?

Monty Hall Dilemma cont'd

- What are the probabilities of the money being behind the two doors? Should we switch door?

$$Pr(\$|1+3) = \frac{2}{3} = Pr(\$|1) + Pr(\$|3) = 0 + Pr(\$|3)$$

$$\Downarrow$$

$$Pr(\$|3) = \frac{2}{3} > \frac{1}{3} = Pr(\$|2)$$

The Binomial Distribution

- We have n repeated trials.
- Each trial has two possible outcomes

Bernoulli trial

  - Success — probability p
  - Failure — probability q=1-p

- What is the probability of having k successes out of n trials?
- We write this question as:

$$\boxed{Pr_n(k) = \frac{n!}{k!\,(n-k)!} p^k q^{n-k} = \binom{n}{k} p^k q^{n-k}}$$

- Faculty: $n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$
  $0! = 1$

Bernoulli Trial

Definition: The binomial coefficient is defined as:

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

Number of ways to select k objects out of a collection of n objects

Example: Out of 10 children, what is the probability that exactly 2 are girls?

$$Pr_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

$$= \frac{10!}{2!(10-2)!} (0,5)^2 (1-0,5)^{10-2} = 0,044$$

Combinatorics

- Take an object from a collection of n objects.
- Repeat the test k times.

Types of Experiments:

- With or without replacement
- Ordered or unordered

Example:

What is the probability that if I have two children that the oldest is a girl and the youngest is a boy?

- Ordered.
- With replacement.

Ordered with Replacement

- Take an object from a collection of n objects.
- Put it back each time.
- Repeat the test k times.
- The sequence of the objects matters.

- The number of combinations is: n^k
  - Each trial has n possible outcomes
  - All the trials are independent

Ordered without Replacement

- Take an object from a collection of n objects.
- Do not put it back each time.
- Repeat the test k times.
- The sequence of the objects matters.

- The number of combinations is:

$$ {}_nP_k = P_k^n = \frac{n!}{(n-k)!} = \mathrm{n} \cdot (n-1) \dots (n-k+1) $$

➢ The 1st trial has n possible outcomes, the 2nd trial has n-1 possible outcomes, … , the k'th trial has n-k+1 possible outcomes

Unordered without Replacement

- Take an object from a collection of n objects.
- Do not put it back each time.
- Repeat the test k times.
- The sequence of the objects do not matter.

- The number of combinations is:

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

  - ➢ The k ordered draws can be shuffled in k! different ways (sequences)

Unordered with Replacement

- Take an object from a collection of n objects.
- Put it back each time.
- Repeat the test k times.
- The sequence of the objects do not matter.

- The number of combinations is:

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!\,(n-1)!}$$

> ➢ Each time we draw an object, we should replace an object (except for the last draw). This correspond to we start with n+k-1 object and draw k objects unordered without replacement.

Summary of Combinatorics

- We can summarise the number of possible outcomes of k trials, sampled from a set of n objects.

		Replacement	
		With	Without
Sam-pling	Ordered	n^k	$P_k^n = \frac{n!}{(n-k)!}$
	Unordered	$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!\,(n-1)!}$	$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$

Experiment: Birthday Example

- k=35 students
- n=365 (number of days in the year)
- What are the probability that at least two have birthday on the same day (E)?

All have different birthdays

Ordered sampling without replacement
(k unique birthsdays in n days)

Complement rule

$$\Pr(E) = 1 - \Pr(\bar{E}) = 1 - \frac{\frac{n!}{(n-k)!}}{n^k} = 1 - \frac{\frac{365!}{(365-35)!}}{365^{35}} > 80\%$$

Ordered sampling with replacement
(all possible combinations of k students birthdays in n days)

- k=50 students: $\Pr(E) > 97\%$
- k=75 students: $\Pr(E) > 99{,}97\%$

Words and Concepts to Know

Prior

Type I Error

Binomial coefficient

Sampling

Unordered

Replacement

Specificity

Likelihood

Combinatorics

Bernoulli Trial

Sensitivity

Posterior

Ordered

Type II Error

Binomial distribution