Bayesian Statistics
Part I: History, Philosophy, and Motivation
Part II: Introduction to Probability

Michael Anderson, PhD
Hélène Carabin, DVM, PhD

Department of Biostatistics and Epidemiology
The University of Oklahoma Health Sciences Center

May 18, 2016

Outline

Who were Bayes, Price, and Laplace?

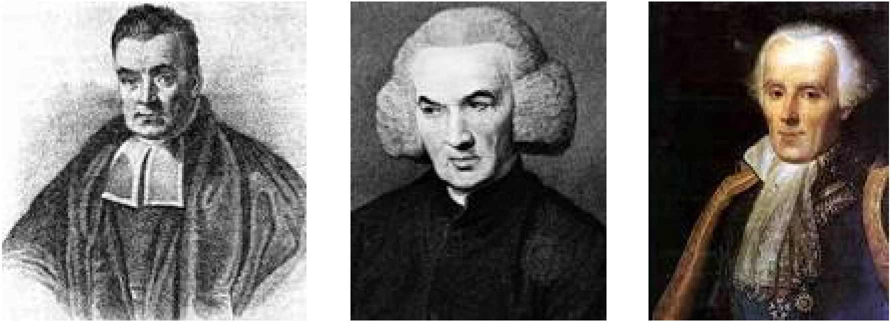


Figure: Major players in creation of Bayes' rule

Sharon Bertsch McGrayne †

- **McGrayne Bios (5:41-9:55)**

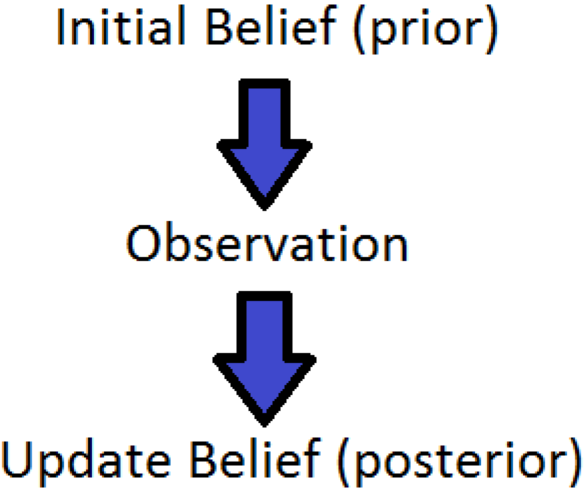
† "The Theory That Would Not Die" How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy"

Bayesian Paradigm

Initial Belief (prior)

Bayesian Paradigm

Bayesian Paradigm



Philosophy of Science

- What is a scientific theory?
  - Is a universal statement
    - Applies to all events in all places and time
  - Explains the behavior/happening of all things (describes reality).
  - Predicts what will happen in the future.
  - Typically assumes that an objective reality exists and can be explored

Philosophy of Science

- How can one come up with a scientific theory
  - Deductive Method
    - General propositions (positive or normative) lead to specific logical implications.
    - Developed as early as 400 BC by Aristotle.
  - Inductivism (method), empiricism (knowledge).
    - From observations to general conclusions.
    - All science must start from unbiased and non-informed (no prior knowledge) observations.
    - There are many forms of inductive reasoning.

Philosophy

- Inductivism
  - Empirism: concept developed by Hume 1777.
  - Causes of events can be determined by observation.
    - Generalization from observations.

  - Unlike deduction, the conclusions of inductive reasoning are *probable* given the evidence, in contrast to being certain.
    - When events not following the rule occur the theory becomes a probability rather than a certainty.
    - This is where Bayesian statitics can play a huge role.
  - Principle of the Uniformity of Nature
    - “The future will resemble the past”

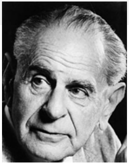
Philosophy

- Some forms of Inductivism
  - Positivism (19^{th} Century)
    - Science can rise above superstition by specializing in the description and analysis of observable phenomena, leading to discovery of natural laws.
  - Logical positivism
    - Science progresses toward truth by oservation, formulation of hypotheses, empirical verification, leading to additional hypotheses.
    - Scientific questions are referred to as "positive" while unscientific ones are "normative."

Philosophy

- Problems with inductivism
  - The theory chosen is not necessarily the right one.
    - Russel's turkey
  - The observed data can be biased
  - Difficult to come up with an experiment without an underlying theory.
    - In other words, we are born Bayesian. . ..

Philosophy



- Incuction has no place in the logic of science.
  - Largely developed by Karl Popper in "Conjectures and Refutations" (1963).
  - Science is deductive where scientists fomulate hypotheses and theories that they test by deriving observations.
  - Theories are not confirmed or verified, *but they may be falsified*.
  - Every "good" scientific theory is a prohibition: it forbids certain things to happen . . .
  - One can try to "corroborate" a hypothesis through refuting that it is true.
    - So a scientific theory becomes something that can be falsifiable, not verifiable.
  - Nothing is certain though, and Popper's philosophy is close to that of Bayesians.

Philosophy

- Problems with refutationism/falsification
    - There may be an infinity of scientific theories to be tested.
    - Not all theories can be verified with observations and many result in probabilistic outcomes (ie movement of planets results from combining several theories)
    - The issue of measurement error remains.

Philosophy

- Hypothetico-deductive and falsification
  - Fisher with p-values (nil-null hypothesis testing).
  - Pearson with hypothesis testing (alternative hypothesis).
  - Neyman with 95% CI.
- Applies Popper's theory to statistics
  - *"The null hypothesis is never proved or established, but is possibly disproved, in the course of experimentation. Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis."* (Fisher 1947).
- Issues with this theory
  - The hypothesis is true or false, there is no probability. . .
  - *"Scientific hypotheses can be rejected (i.e. falsified), but never established or accepted the same way."* (Gelman and Shalizi, 2013)
  - Many more during this workshop. . .

Philosophy

- The Bayesian approach
  - Refutation and hypothetico-deductive approaches were developed to avoid the concept that theories in science can be appraised in terms of their "probabilities."
  - Theories usually lie between being certainly right or certainly wrong, we cannot tell...
  - *"It is often stated that one should experiment without preconceived ideas. This is simply impossible; not only would it make every experiment sterile, but even if we were ready to do so, we could not implement this principle. Everyone stands by his own conception of the world, which he cannot get rid of so easily."* (Poincaré, 1905)

Philosophy

- Inductivism: requires 3 conditions
  - Make SEVERAL observations of the events resulting from the theory (for example, all heated metals expand).
  - Make obsevations under SEVERAL CONDITIONS
    - The theory must hold under ALL conditions.
  - NO observation can go against the theory.
  - SO if A has been observed under various conditions, and if ALL A have characteristic B, then all A lead to B.

Philosophy

- Bayesian statistics have an inductive approach
  - Start with a prior distribution (prior knowledge, multitude of observations).
  - Get data (observe under different conditions).
  - Obtain a posterior distribution (update prior knowledge).
- However, Bayesian statistics do not have to follow the hypothetico-deductive and refutation theories.
  - Preconceived idesa are allowed and even encouraged.

Motivation and History

The allegory of our statistical lives

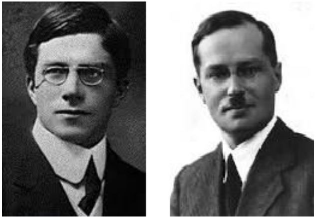
- Most of us are born Bayesians.

"It is remarkable that this science (probability), which originated in the consideration of games of chance, should have become the most important object of human knowledge."
$\sim$ Pierre-Simon de Laplace (1749-1827)

- **McGrayne WWII (16:02-30:23)**
- **McGrayne Air France Flight 447 (2:30-4:00)**

Motivation and History

- But by adolescence...



  - Fisher's work at Rothamstad advanced experimental design.
    McGrayne Fisher (12:44-16:00)
    McGrayne Obscurity (29:57-36:31)
  - Are these methods appropriate for observational studies?
- Today's message- "It's never too late to have a happy childhood."

Motivation and History

Why go Bayesian?

- Another powerful tool for your tool kit.
  - measurement error from misclassification
  - complex dependencies among observations
  - missing data
- Frequentist probabilities don't always make intuitive sense.
  - **McGrayne Frequentist (9:56-12.44)**
  - Example: Probability of an H-bomb accident.
- Bayesian "credible intervals" can be correctly interpreted by intro stat students.
- P-values and confidence intervals are somewhat ill defined.
  - Example: A study finds that out of 24 subjects with lung cancer 7 are female.

Motivation and History

Example: 7 out of 24 subjects with lung cancer are female.

Gender	Gender
0	0
1	0
0	0
0	0
1	0
0	0
1	0
1	0
1	0
0	1
0	0
1	0

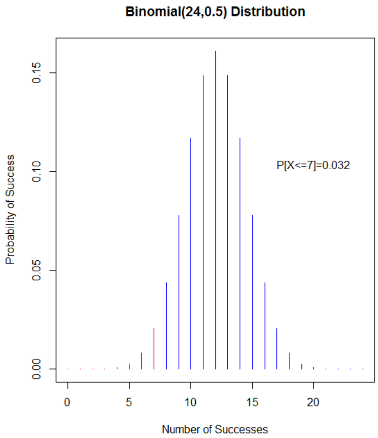


Figure: One-tailed Binomial p-value.

Motivation and History

Example: 7 out of 24 subjects with lung cancer are female.

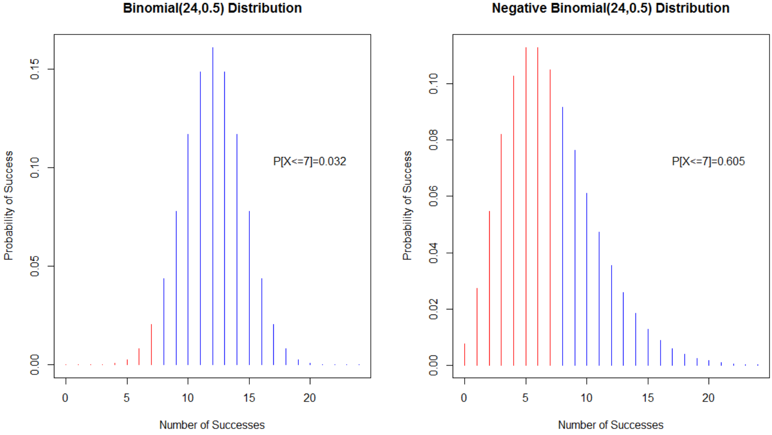


Figure: P-values depend on the researcher's intention.

Experiment, Outcome, Sample Space, and Event

Experiment
A process or procedure for which there is more than one possible outcome.

Outcome
The results of a single trial of an experiment.

Sample Space S
The collection of all possible outcomes of an experiment.

Event
A subset of the S. A collection of outcomes of interest.

Examples of Sample Space

- Experiment 1: Tossing one coin one time: $S = \{H, T\}$
- Experiment 2: Tossing one coin 4 times:
- How many ways to get
  - all Tails?
  - at least 2 Tails?
  - at least 1 Head?

$$S = \left\{ \begin{array}{cccc} H & H & H & H \\ H & H & H & T \\ H & H & T & H \\ H & T & H & H \\ H & H & T & T \\ H & T & T & H \\ H & T & H & T \\ H & T & T & T \\ T & H & H & H \\ T & H & H & T \\ T & H & T & H \\ T & T & H & H \\ T & H & T & T \\ T & T & H & T \\ T & H & T & T \\ T & T & T & T \end{array} \right\}$$

Examples of Sample Space

- Rolling one die: $S = \{1, 2, 3, 4, 5, 6\}$
- Rolling two dice:

S

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Figure: The sample space for two dice.

Example of an Event

Let A be the event that the sum of the dice is 6.

$\mathcal{S}$

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Figure: The sample space for two dice.

Example of an Event

Let A be the event that the sum of the dice is 6.

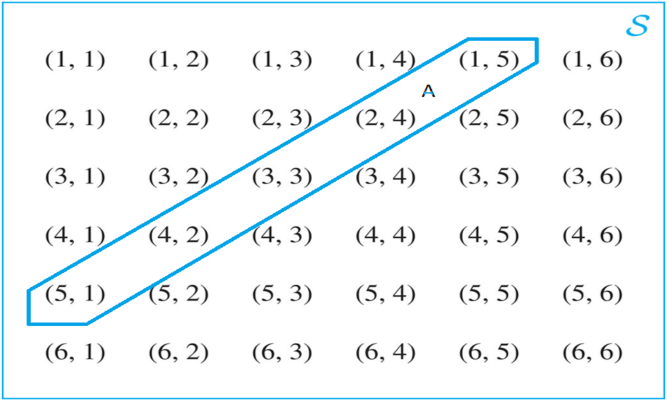


Figure: The sample space for two dice.

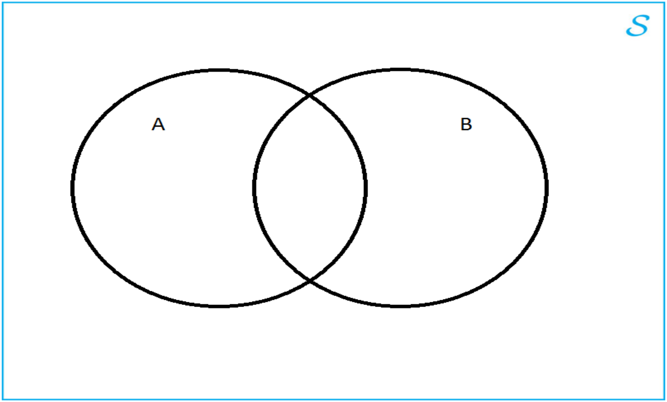
Axioms of Probability

For a sample space S

1. the probability of the i^{th} outcome is p_i
2. $0 \leq p_i \leq 1$
3. $\sum_{i=1}^{n} p_i = 1$

Venn Diagrams

It is often useful to express events in S as a Venn Diagram

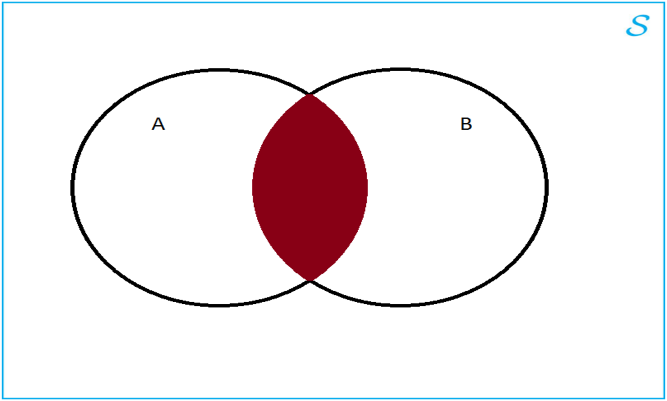


- gives you a feel for the outcomes common to the events.
- depicts event/S ratio.

Combinations of Events

Intersection of Events ($A \cap B$)

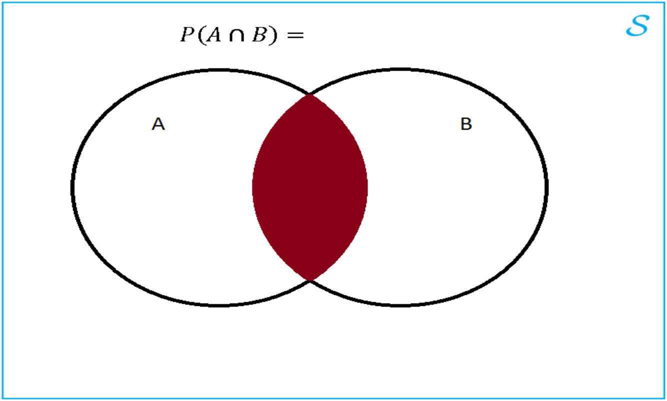
The set of outcomes that belong to both A "AND" B.



Combinations of Events

Intersection of Events ($A \cap B$)

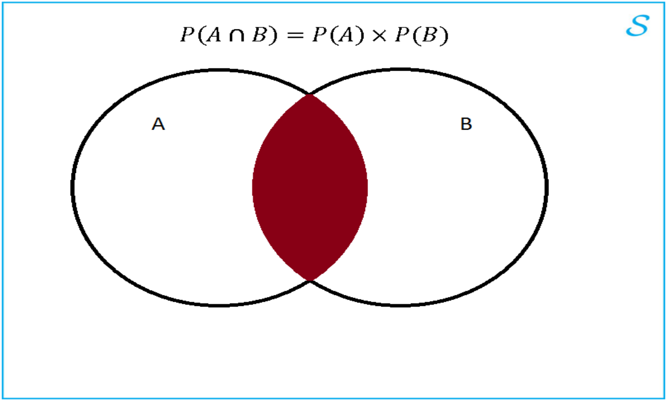
The set of outcomes that belong to both A "AND" B.



Combinations of Events

Intersection of Events ($A \cap B$)

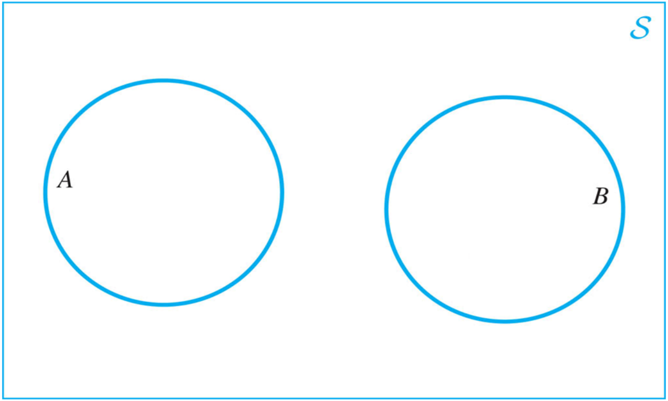
For **independent** A and B



Combinations of Events

Mutually Exclusive (disjoint) Events

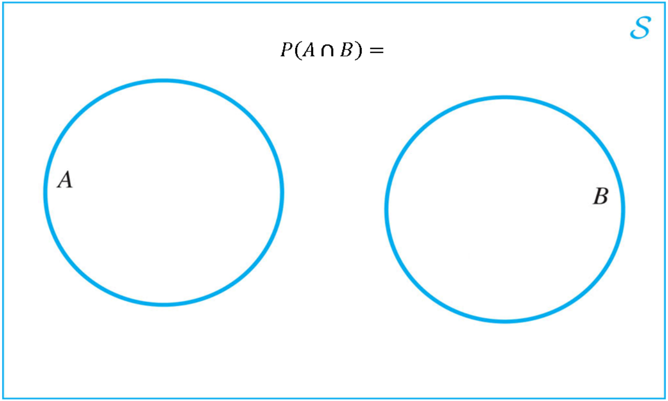
Two events A and B are said to be mutually exclusive if $(A \cap B) = \emptyset$.



Combinations of Events

Mutually Exclusive (disjoint) Events

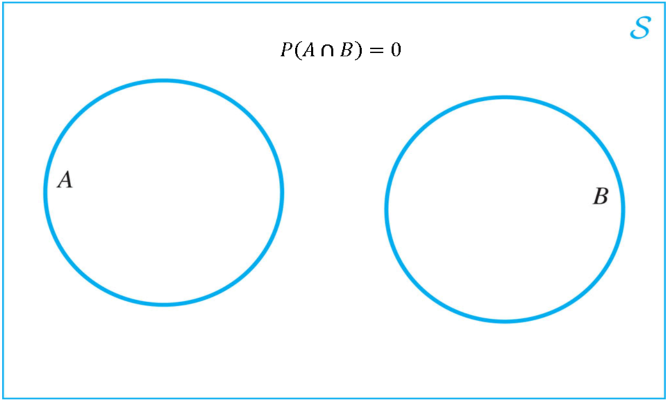
Two events A and B are said to be mutually exclusive if $(A \cap B) = \emptyset$.



Combinations of Events

Mutually Exclusive (disjoint) Events

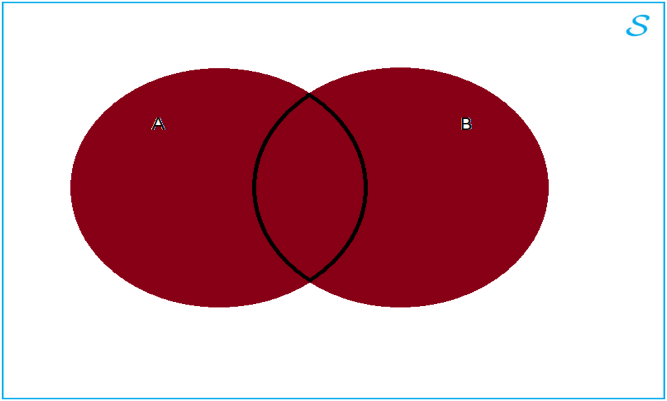
Two events A and B are said to be mutually exclusive if $(A \cap B) = \emptyset$.



Combinations of Events

Union of Events ($A \cup B$)

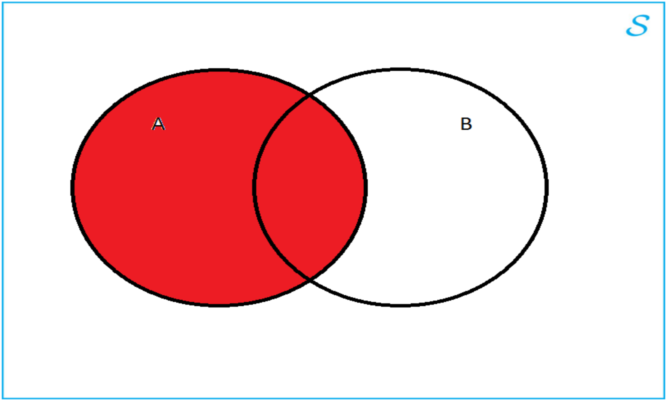
The set of outcomes that belong to either A "OR" B.



Combinations of Events

Union of Events ($A \cup B$)

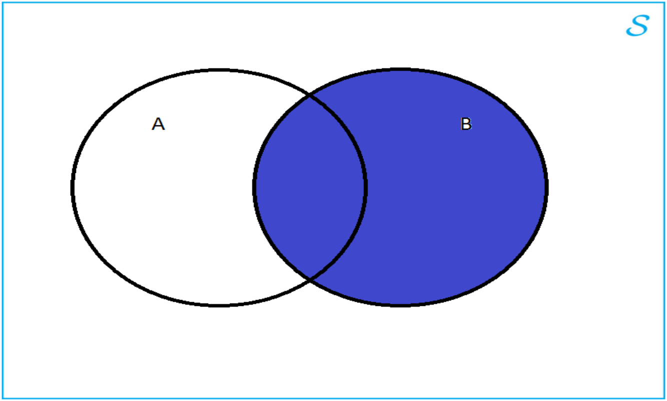
The set of outcomes that belong to either A "OR" B.



Combinations of Events

Union of Events ($A \cup B$)

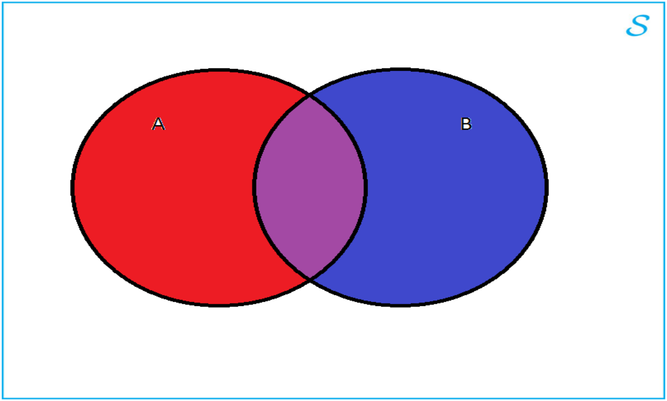
The set of outcomes that belong to either A "OR" B.



Combinations of Events

Union of Events ($A \cup B$)

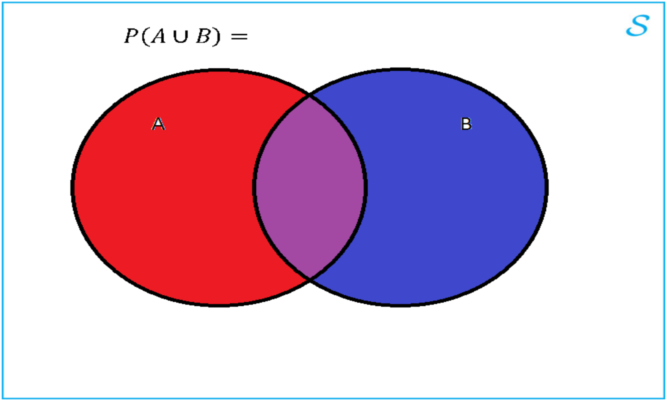
The set of outcomes that belong to either A "OR" B.



Combinations of Events

Union of Events ($A \cup B$)

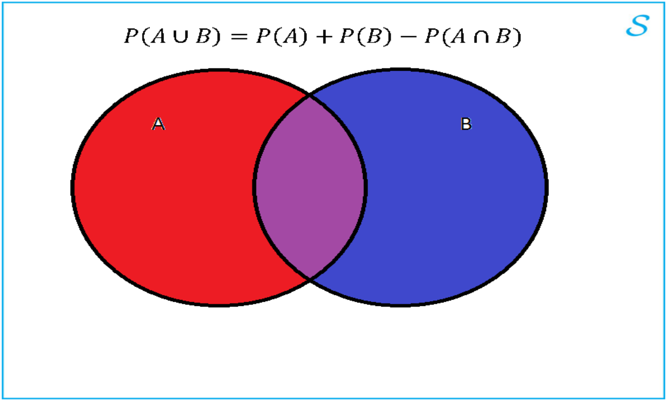
The set of outcomes that belong to either A "OR" B.



Combinations of Events

Union of Events ($A \cup B$)

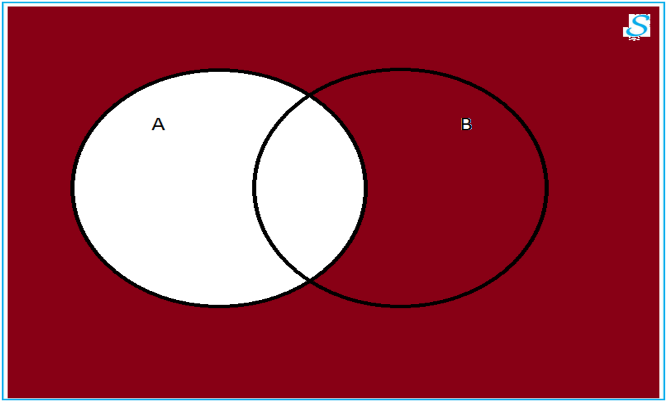
The set of outcomes that belong to either A "OR" B.



Combinations of Events

Complement of A, A^c

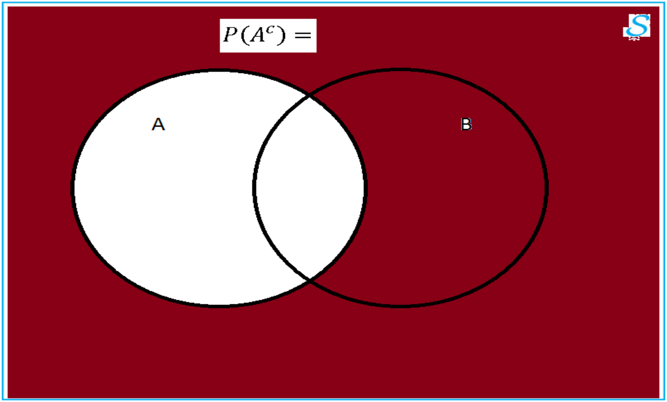
The set of all outcomes that are not in A.



Combinations of Events

Complement of A, A^c

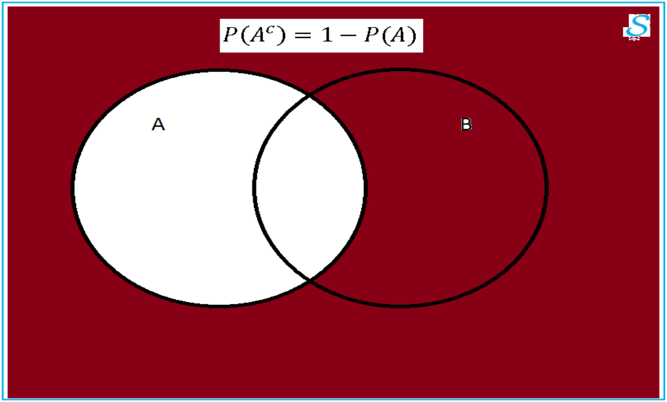
The set of all outcomes that are not in A.



Combinations of Events

Complement of A, A^c

The set of all outcomes that are not in A.



Example: Risk of Mortality during the West Africa Ebola Virus Outbreak of 2014†

	Total
Yes	11310
No	17306
Total	28616

P(Yes)=11310/28616≈0.40.

What if we have additional information such as location (country) of the case?

What is the probability of mortality given the location of the case (i.e. $P(Yes|location)$)?

- $P(Yes|Guinea) = 2544/3814 \approx 0.67$.
- $P(Yes|SierraLeone) = 3956/14124 \approx 0.28$.
- $P(Yes|Liberia) = 4810/10678 \approx 0.45$.

† http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/case-counts.html

What is a Conditional Probability?

Conditional Probability of Events A and B

If A and B are events in S, and $P(B) > 0$, then the conditional probability of A given B is

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

and

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

Note: This gives us the general formulas for computing $P(A \cap B)$.

- $P(A \cap B) = P(A)P(B|A)$.
- $P(A \cap B) = P(B)P(A|B)$.
- Note: $P(A)P(B|A) = P(B)P(A|B)$

Independence

Independence

Two events A and B are said to be independent if the occurence of one event in no way influences the probability of occurence of the other event. More formally this means

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

or equivalently

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

So, for independent events A and B, $P(A \cap B) = P(A)P(B)$.

Let's contrast independent events with mutually exclusive events.

	Independent	Mutually Exclusive
$P(A \cap B)$		
$P(A|B)$		
$P(B|A)$		

Practicum 1

See Practicum 1 Handout