Chapter 1

Axioms of Probability

Outline of Chapter 1

- 1.1 Introduction
- 1.2 Sample space and events
- 1.3 Axioms of probability
- 1.4 Basic theorems
- 1..5 Continuity of probability function
- 1.6 Probabilities of 0 and 1
- 1.7 Random selection of points from intervals
- 1.8 What is simulation?

Section 1.1

Introduction

1.1 Introduction

- In search of natural laws that govern a phenomenon, science often faces "**events**" that may or may not occur.
- In any experiment, an event that may or may not occur is called **random**.
- If the occurrence of an event is inevitable, it is called **certain**, and if it can never occur, it is called **impossible**.

1.1 Introduction

Luca Paccioli(1445-1514) Italian
(studies of chances of events)
Niccolo Tartaglia(1499-1557)
Girolamo Cardano(1501-1576)
Galileo Galielei(1564-1642)

Blaise Pascal(1623-1662) French
Pierre de Fermat(1601-1665)
1655 Christian Huygens(1629-1695) Dutch
first book "On Calculations in Games of Chance)
James Bernoulli(1654-1705)
Abraham de Moivre(1667-1754)

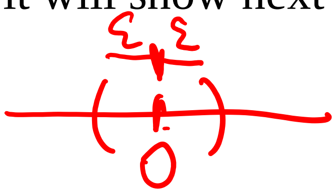
Pierre-Simon Laplace(1749-1827)
Simeon Denis Poisson(1781-1840)
Karl Friedrich Gauss(1777-1855)

Pafnuty Chebyshev(1821-1894) Russian
Andrei Markov(1856-1922)
Aleksandr Lyapunov(1857-1918)

Relative frequency and its problem

- In practice, $\lim_{n\to\infty} n(A)/n$ cannot be computed since it is impossible to repeat an experiment infinitely many times. Moreover, if for a large n, $n(A)/n$ is taken as an approximation for the probability of A, there is no way to analyze the error
- There is no reason to believe that the limit of $n(A)/n$, as $n \to \infty$, exists. Also, if the existence of this limit is accepted as an axiom, many dilemmas arise that cannot be solved
- By this definition, probabilities that are based on our personal belief and knowledge are not justifiable
  - The probability that the price of oil will be raised in the next six months is 60%
  - The probability that it will snow next Christmas is 30%.

$\lim_{n\to\infty} \left(\frac{1}{3}\right)^n = 0$

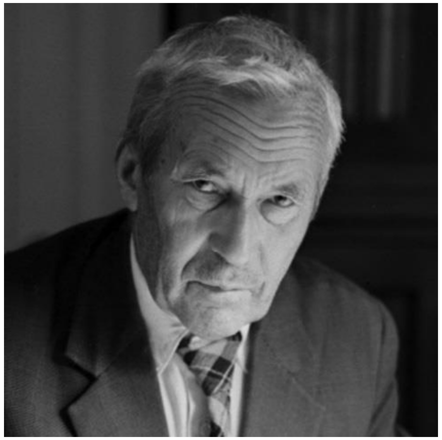


$\forall \varepsilon > 0, \exists N > 0$, s.t. $n > N$, $\left|\left(\frac{1}{3}\right)^n - 0\right| < \varepsilon$.

1900 David Hilbert(1862-1943) pointed out the urgent need for an axiomatic treatment of the theory of probability

Emile Borel(1871-1956)
Serge Bernstein(1880-1968)
Richard von Mises(1883-1953)

*1933 Andrei Kolmogorov(1903-1987) Russian
successfully axiomatized the theory of probability



Section 1.2

Sample Space and Events

1.2 Sample space and events

- Experiment (eg. Tossing a die)
- Outcome(sample point)
- Sample space={all outcomes}
- Event: subsets of a sample space
  - If the outcome of an experiment belongs to an event E, we say that the event E has **occurred**.

- Ex1.1 tossing a coin once
  - sample space S = {H, T}
- Ex1.2 flipping a coin and tossing a die if T or flipping a coin again if H
  - S={T1,T2,T3,T4,T5,T6,HT,HH}
  - E={T3,T4,T5,T6} is an event that flipping a coin results in T, and tossing a die results in a number greater than or equal to 3
  - If one flips a coin, gets T, tosses a die and get 5, we say that event E has occurred

Sample space and events

- Ex1.3 measuring the lifetime of a light bulb

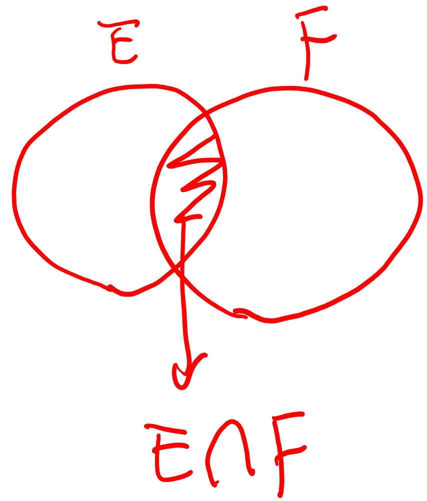
$S = \{x: x \geq 0\}$
$E = \{x: x \geq 100\}$ is the event that the light bulb lasts at least 100 hours

- Ex1.4 all families with 1, 2, or 3 children (genders specified in descending order of their ages)
  - S= {b,g,bg,gb,bb,gg,bbb,bgb,bbg,bgg,ggg,gbg,ggb,gbb}
  - F={b, bg, bb, bbb, bgb, bbg, bgg} represents families with male oldest child
  - G={gg, bgg, gbg, ggb} represents families with exactly two girls

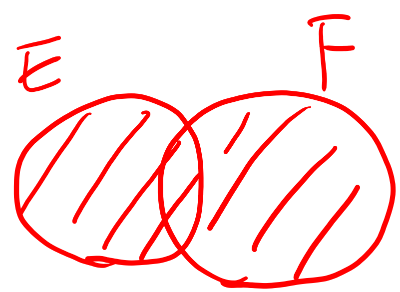
- If the outcome of an experiment belongs to an event E, then we say that the event E has **occurred**

Relations between events

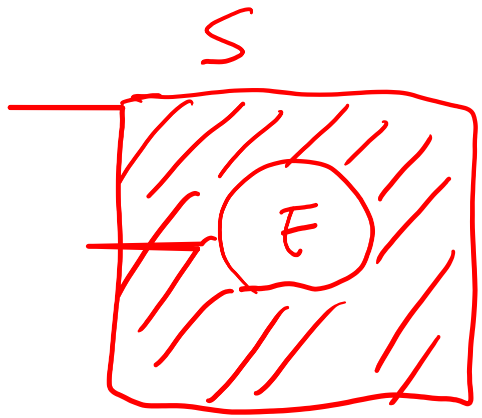
- Subset: $E \subseteq F$
- Equality: $E \subseteq F$ and $F \subseteq E$
- Intersection: EF or $E \cap F$



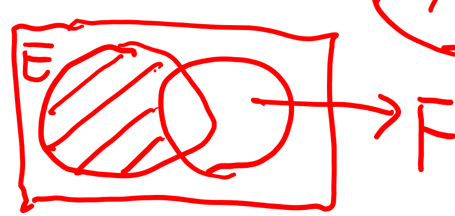
- Union: $E \cup F$



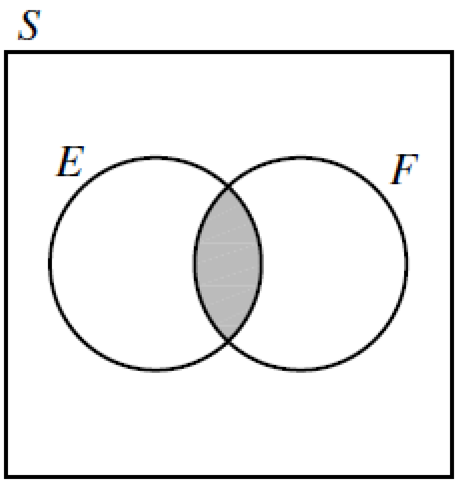
- Complement: E^c



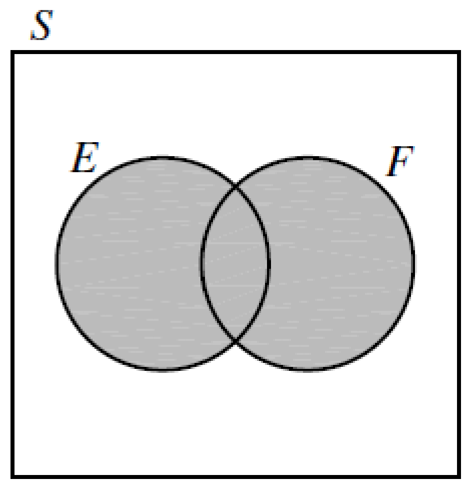
- Difference: $E - F = E \cap F^c$



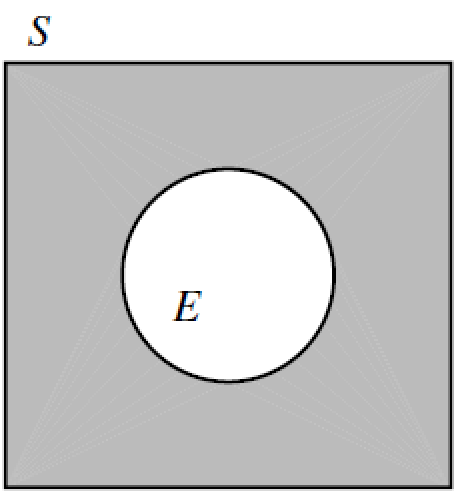
Venn diagrams



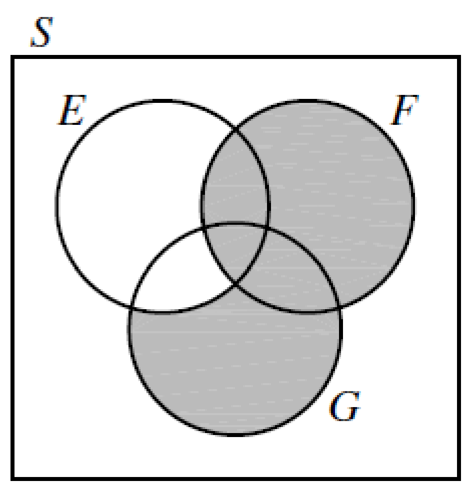
EF



$E \cup F$



E^c



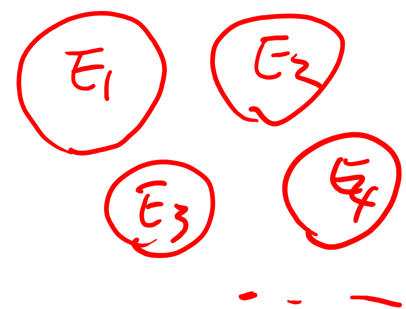
$(E^c G) \cup F$

- Certainty
  - An event is called certain if its probability is one
- Impossibility
  - An event is called impossible if its probability is zero
- Mutually exclusiveness
  - A set of events are called mutually exclusive if the intersection of any two of them is the empty set

disjoint

- Similarly, we define

$$\bigcup_{i=1}^{n} E_i \,, \bigcap_{i=1}^{n} E_i \,, \bigcup_{i=1}^{\infty} E_i \,, \text{and } \bigcap_{i=1}^{\infty} E_i$$



$E_i \cap E_j = \phi$

$\forall i, j, i \neq j$

Useful identities between events

- $(E^c)^c = E, E \cup E^c = S$, and $EE^c = \emptyset$
- Commutative law:

$$E \cup F = F \cup E, \quad EF = FE$$

- Associative laws:

$$E \cup (F \cup G) = (E \cup F) \cup G, \quad E(FG) = (EF)G$$

- Distributive laws:

$$(EF) \cup H = (E \cup H)(F \cup H), \quad (E \cup F)H = (EH) \cup (FH)$$

- $(E^c)^c = E, E \cup E^c = S$, and $EE^c = \emptyset$
- Commutative law:

$$E \cup F = F \cup E, \quad EF = FE$$

- Associative laws:

$$E \cup (F \cup G) = (E \cup F) \cup G, \quad E(FG) = (EF)G$$

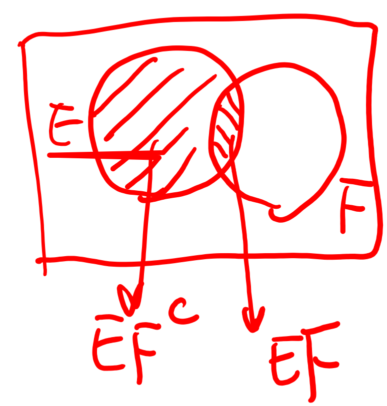
- Distributive laws:

$$(EF) \cup H = (E \cup H)(F \cup H), \quad (E \cup F)H = (EH) \cup (FH)$$

- Useful identities

$$E = EF \cup EF^c$$

$$E = ES = E(F \cup F^c) = EF \cup EF^c$$



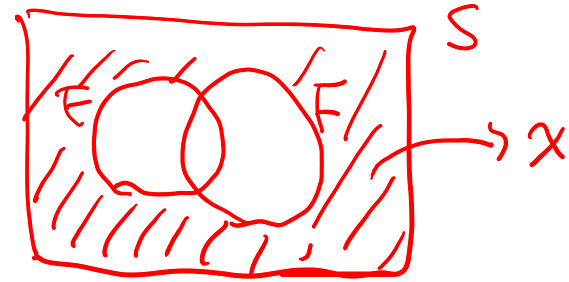
De Morgan's Laws

- De Morgan's first laws:

$$(E \cup F)^c = E^c F^c, \qquad \left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c, \qquad \left(\bigcup_{i=1}^{\infty} E_i\right)^c = \bigcap_{i=1}^{\infty} E_i^c$$

- De Morgan's second laws:

$$(EF)^c = E^c \cup F^c, \qquad \left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^c, \qquad \left(\bigcap_{i=1}^{\infty} E_i\right)^c = \bigcup_{i=1}^{\infty} E_i^c$$



Example 1.9

- Prove De Morgan's first law: For E and F, two events of a sample space S, $(E \cup F)^c = E^c F^c$
- We prove $(E \cup F)^c \subseteq E^c F^c$ and $(E \cup F)^c \supseteq E^c F^c$
- To prove the first inclusion, suppose that x is an outcome that belongs to $(E \cup F)^c$
- Then x does not belong to $E \cup F$, meaning that <u>x is neither in E</u> <u>nor in F</u>

$x \notin E$, $x \notin F$

- So x belongs to both E^c and F^c and hence to $E^c F^c$

$x \in E^c$, $x \in F^c$

- The other inclusion can be proved similarly

Section 1.3

Axioms of Probability

What is an axiom?

Assumption vs axiom

$$A \to B$$
$$A_1 \to A$$
$$A_2 \to A_1$$
$$\vdots$$
$$A_n \to A_{n-1}$$

1.3 Axioms of probability

$f: R \to R$

P : Event $\to [0,1]$

- S: the sample space
- A: an event
- P is called a probability
- $P(A)$ is called the probability of A if the following axioms are satisfied

$\forall i \neq j \quad A_i \cap A_j = \phi$

Axiom 1 $P(A) \geq 0.$

Axiom 2 $P(S) = 1.$

Axiom 3 *If $\{A_1, A_2, A_3, \ldots\}$ is a sequence of mutually exclusive events (i.e., the joint occurrence of every pair of them is impossible: $A_i A_j = \emptyset$ when $i \neq j$), then*

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Equally likely

- Events A and B are equally likely if $P(A) = P(B)$
- Sample points ω and μ are equally likely if events $\{\omega\}$ and $\{\mu\}$ are equally likely, i.e. $P(\{\omega\}) = P(\{\mu\})$

- **Theorem 1.1** The probability of the empty set $\emptyset$ is 0. That is, $P(\emptyset) = 0$
- **Proof**.

$A_i A_j = \phi \quad \forall i \neq j$

- Let $A_1 = S$ and $A_i = \emptyset$ for $i \geq 2$
- Then, $A_1, A_2, A_3, \ldots$ is a sequence of mutually exclusive events
- By Axiom 3,

$$P(S) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = P(S) + \sum_{i=2}^{\infty} P(\emptyset)$$

implying that $\sum_{i=2}^{\infty} P(\emptyset) = 0$

- This is only possible if $P(\emptyset) = 0$

Theorem 1.2

- **Theorem 1.2** *Let* $\{A_1, A_2, \ldots, A_n\}$ *be a mutually exclusive set of events. Then*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

$A \cup \phi = A$

- **Proof.** For $i > n$, let $A_i = \emptyset$.
- Then $A_1, A_2, A_3, \ldots$ is a sequence of mutually exclusive events.
- Thus, by Axiom 3 and Theorem 1.1, we get

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n} A_i\right) &= P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \\
&= \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(A_i) = \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(\emptyset) \\
&= \sum_{i=1}^{n} P(A_i). \quad \blacklozenge
\end{aligned}$$

- Axiom 3 is stated for a *countably infinite* collection of mutually exclusive events.
- For this reason, it is also called the *axiom of countable additivity*.
- Theorem 1.1 states that the same property holds for a *finite* collection of mutually exclusive events as well.
- That is, P also satisfies *finite additivity*.

- One might want ask if one can replace axiom 3 by its finite version, i.e.

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

- The answer is no! We would then not be able to handle the countable additivity.
- The countable union and intersection of events occur often.
- Consider tossing a die repeatedly.
- Event A_n represents that the **first** six occurs in the n-th toss.
- Then, $\{A_n, n = 1, 2, \cdots\}$ are mutually exclusive.
- $\bigcup_{i=1}^{\infty} A_i$ is the event that a six **eventually** occurs

$i \neq j$

$A_i A_j = \phi$

1 2 ... n

2 3 1 ... 3 1 5 ... 6

An important implication of Theorem 1.2

- An important special case of Theorem 1.2

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

- An important implication of Theorem 1.2

$$0 \leq P(A) \leq 1$$

- To see this, note that from Theorem 1.2

$$P(\underbrace{A \cup A^c}_{S}) = P(A) + P(A^c)$$

- By Axiom 2,

$$P(A \cup A^c) = P(S) = 1$$

- Therefore, $P(A) + P(A^c) = 1$. This and Axiom 1 imply that $P(A) \leq 1$

Sample Spaces with Equally Likely Outcomes

- Suppose that a sample space contains N points that are equally likely to occur
  - The probability of each occurrence of a sample point is $1/N$

$$
\begin{aligned}
1 = P(S) &= P\big(\{s_1, s_2, \ldots, s_N\}\big) \\
&= P\big(\{s_1\}\big) + P\big(\{s_2\}\big) + \cdots + P\big(\{s_N\}\big) = N P\big(\{s_1\}\big)
\end{aligned}
$$

- Before Kolmogorov introduced the three axioms in 1933, this was taken as the definition of probability.
- It is now called the classical definition of probability

Theorem 1.3 Classical Definition of Probability

- **Theorem 1.3** *Let S be the sample space of an experiment. If S has N points that are all equally likely to occur, then for any event A of S,*

$$P(A) = \frac{N(A)}{N}.$$

where $N(A)$ is the number of points of A.

- **Proof.** Let $S = \{s_1, s_2, ..., s_N\}$, where each s_i is an outcome (a sample point) of the experiment.

$A = \{s_1, s_2, \cdots, s_{N(A)}\}$

- Since the outcomes are equiprobable,

$$P(\{s_i\}) = 1/N \quad \text{for all } i, 1 \le i \le N.$$

- Now let $A = \{s_{i_1}, s_{i_2}, \ldots, s_{i_{N(A)}}\}$, where $s_{i_j} \in S$ for all i_j.

Proof of Theorem 1.3 - Continuation

- Since $\{s_{i_1}\}, \{s_{i_2}\}, \ldots, \{s_{i_{N(A)}}\}$ are mutually exclusive, Axiom 3 implies that

$$\begin{aligned} P(A) &= P(\{s_{i_1}, s_{i_2}, \ldots, s_{i_{N(A)}}\}) \\ &= P(\{s_{i_1}\}) + P(\{s_{i_2}\}) + \ldots + P(\{s_{i_{N(A)}}\}) \\ &= \frac{1}{N} + \frac{1}{N} + \ldots + \frac{1}{N} \\ &= \frac{N(A)}{N} \end{aligned}$$

Example 1.12

- An elevator with two passengers stops at the second, third, and fourth floors.
- If it is equally likely that a passenger gets off at any of the three floors,
  - what is the probability that the passengers get off at different floors?

Solution of Example 1.13

- Let a and b denote the two passengers and a_2b_4 mean that a gets off at the second floor and b gets off at the fourth floor, with similar representations for other cases.
- Let A be the event that the passengers get off at different floors.
- Then

$$S = \{a_2b_2, a_2b_3, a_2b_4, a_3b_2, a_3b_3, a_3b_4, a_4b_2, a_4b_3, a_4b_4\}$$
$$A = \{a_2b_3, a_2b_4, a_3b_2, a_3b_4, a_4b_2, a_4b_3\}$$

- So $N = 9$ and $N(A) = 6$.
- Therefore, the desired probability is $N(A)/N = 6/9 = 2/3$.

Section 1.4

Basic Theorems

Theorem 1.4

- **Theorem 1.4** *For any event* A, $P(A^c) = 1 - P(A)$.
- **Proof.** A and A^c are mutually exclusive.
- Thus,

$$P(A) + P(A^c) = 1$$

- But $A \cup A^c = S$ and $P(S) = 1$, so

$$1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$$

- Therefore, $P(A^c) = 1 - P(A)$.

Theorem 1.5

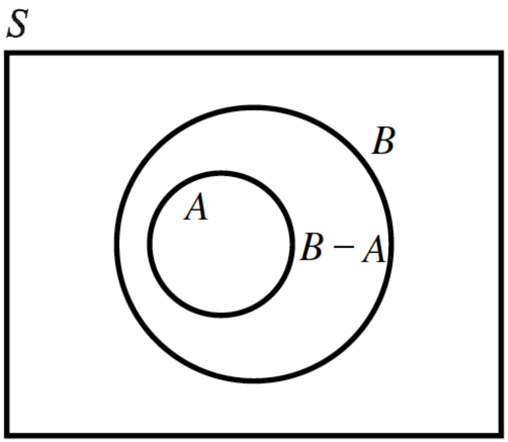
- **Theorem 1.5** *If $A \subseteq B$, then*

$$P(B - A) = P(BA^c) = P(B) - P(A).$$

- **Proof.** $A \subseteq B$ implies that $B = (B - A) \cup A$ (see Figure 1.2).
- But $(B - A)A = \emptyset$.
- So the events $B - A$ and A are mutually exclusive, and

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

- This gives $P(B - A) = P(B) - P(A)$.



Corollary of Theorem 1.5

- **Corollary. If $A \subseteq B$, then $P(A) \leq P(B)$.**
- **Proof.**
- By Theorem 1.5, $P(B - A) = P(B) - P(A)$.
- Since $P(B - A) \geq 0$, we have that $P(B) - P(A) \geq 0$.
- Hence $P(B) \geq P(A)$.

Theorem 1.6

- **Theorem 1.6** $P(A \cup B) = P(A) + P(B) - P(AB)$
- **Proof.** Since $A \cup B = A \cup (B - AB)$ (see Figure 1.3) and $A(B - AB) = \emptyset$,
- So A and $B - AB$ are mutually exclusive events and

$$P(A \cup B) = P(A \cup (B - AB)) = P(A) + P(B - AB). \qquad (1.3)$$

- Now since $AB \subseteq B$, Theorem 1.5 implies that

$$P(B - AB) = P(B) - P(AB).$$

- Therefore, (1.3) gives

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

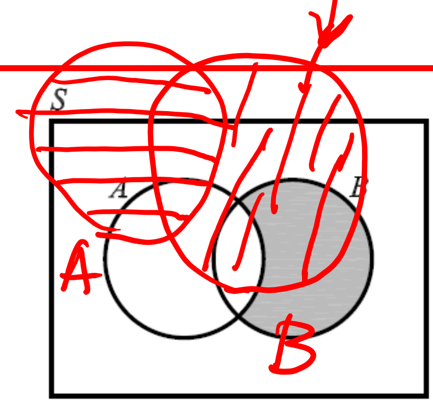
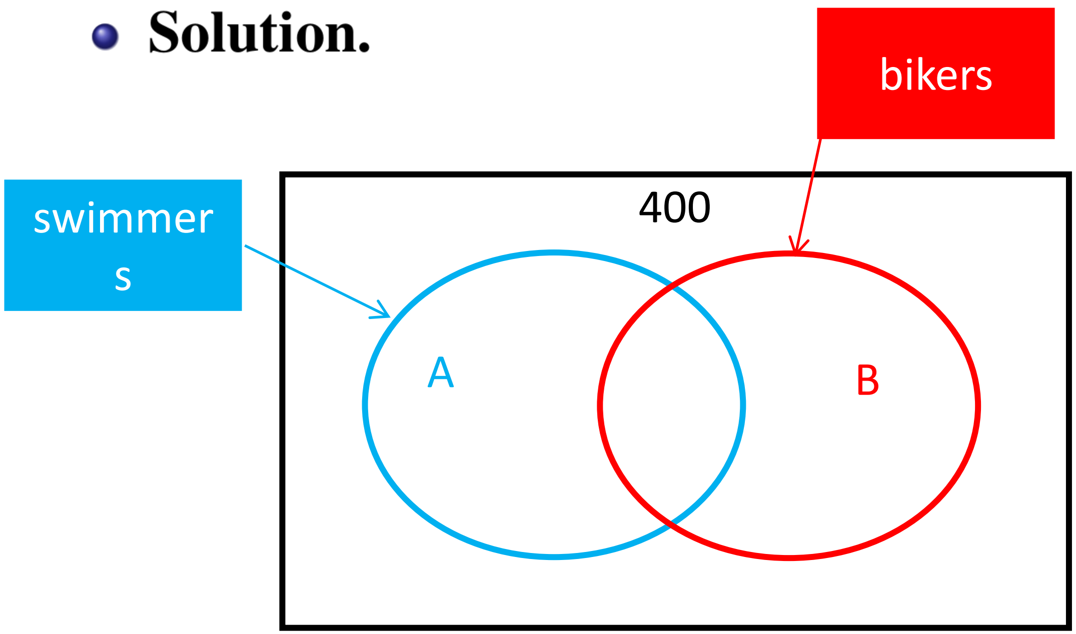


Figure 1.3 The shaded region is $B - AB$. Thus $A \cup B = A \cup (B - AB)$.

Example 1.16

- Suppose that in a community of 400 adults, 300 bike or swim or do both, 160 swim, and 120 swim and bike.
- What is the probability that an adult, selected at random from this community, bikes?
- **Solution.**



- $|S| = 400$
- $|A \cup B| = 300$
- $|A| = 160$
- $|A \cap B| = 120$

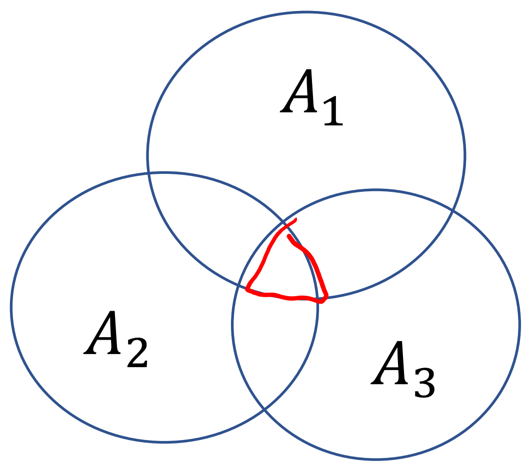
- $|B| = ?$

- Let A be the event that the person swims and B be the event that he or she bikes.
- Then $P(A \cup B) = 300/400$, $P(A) = 160/400$, and $P(AB) = 120/400$.
- Hence the relation $P(A \cup B) = P(A) + P(B) - P(AB)$ implies that

$$\begin{aligned} P(B) &= P(A \cup B) + P(AB) - P(A) \\ &= \frac{300}{400} + \frac{120}{400} - \frac{160}{400} \\ &= \frac{260}{600} \end{aligned}$$

Generalization of theorem 1.6

$$P(A_1 \cup A_2 \cup A_3)$$
$$= P(A_1) + P(A_2) + P(A_3) - P(A_1\, A_2) - P(A_2A_3)$$
$$- P(A_1A_3) + P(A_1A_2A_3)$$



Inclusion-exclusion principle

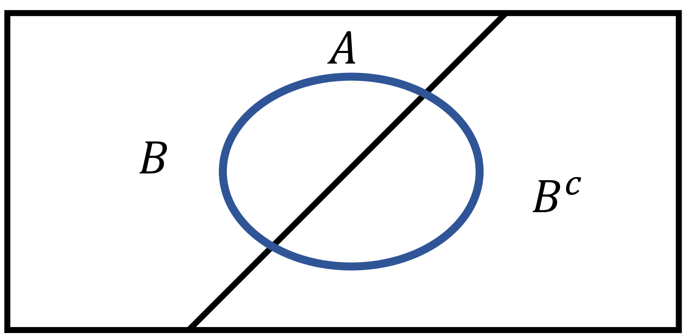
- Add the probabilities of those intersections that are formed of an odd number of events
- Subtract the probabilities of those formed of an even number of events.
- This formula can be proven by induction

$$
\begin{aligned}
P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(A_i A_j) \\
&\quad + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(A_i A_j A_k) \\
&\quad - \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n)
\end{aligned}
$$

Theorem 1.7

- **Theorem 1.7** $P(A) = P(AB) + P(AB^c)$.
- **Proof.** Clearly, $A = AS = A(B \cup B^c) = AB \cup AB^c$.
- Since AB and AB^c are mutually exclusive,

$$P(A) = P(AB \cup AB^c) = P(AB) + P(AB^c).$$



Example 1.20

- In a community,
  - 32% of the population are male smokers;
  - 27% are female smokers.
- What percentage of the population of this community smoke?
- **Solution.**
- Let A be the event that a randomly selected person from this community is a smoker.
- Let B be the event that the person is male.
- By Theorem 1.7,

$$P(A) = P(AB) + P(AB^c) = 0.32 + 0.27 = 0.59$$

Homework 1

$$\lim_{n\to\infty} a_n = c$$

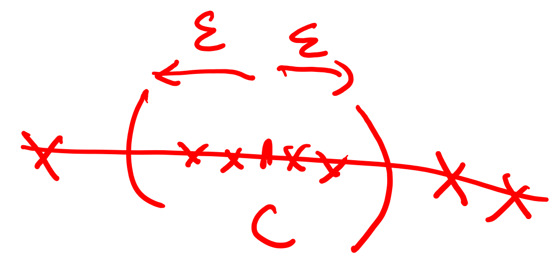
$\forall \varepsilon > 0, \exists N > 0$, s.t.

$|a_n - c| < \varepsilon$

$n \geq N$.

- Section 1.2: B.29, B.30, B.33, B.34
- Section 1.4: A.24, B.40, B.42, B.46

- Due date: 4 pm, Friday, March 8, 2024
- No late homework is accepted.
- EECS building room 845



- Hint for problem B.33 of Section 1.2.
  - By the event that infinitely many of the A_i's occur, the author means the event that **all but finitely many** of the A_i's occur.

$\bigcap \bigcup A_i \qquad \bigcup \bigcap A_i$

Section 1.5

Continuity of Probability Functions

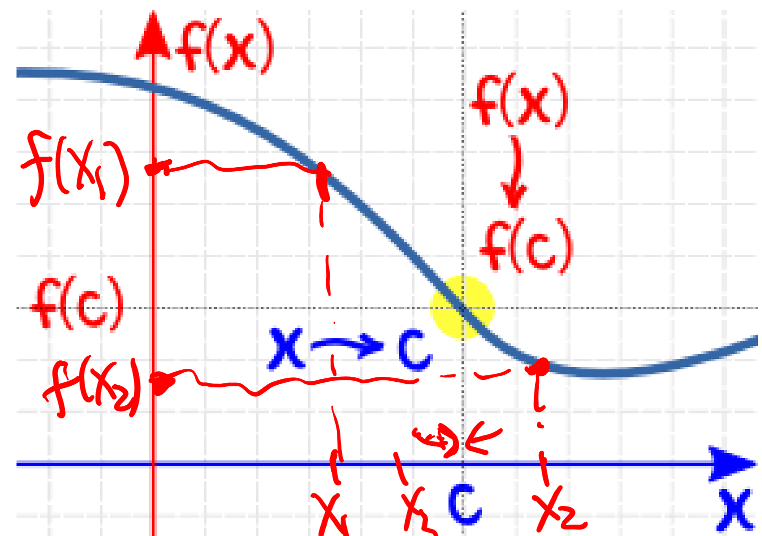
Continuous Real Functions

- Let f be a real valued function
- That is, $f : R \longrightarrow R$.
- f is said to be continuous at c, if sequence $f(x_n)$ converges to $f(c)$ for any sequence $\{x_n\}$ that converges to c
- $f(x_n) \longrightarrow f(c)$ whenever $x_n \longrightarrow c$, or

$x_1, x_2, x_3, \ldots \rightarrow c$

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right)$$

- Equivalently, $\lim_{x\to c} f(x) = f(c)$
- f is continuous on the real line if it is continuous at all real points



- Probability functions map from a set to a real number
- We now show that probability functions are continuous
- A sequence of events of a sample space is called increasing if

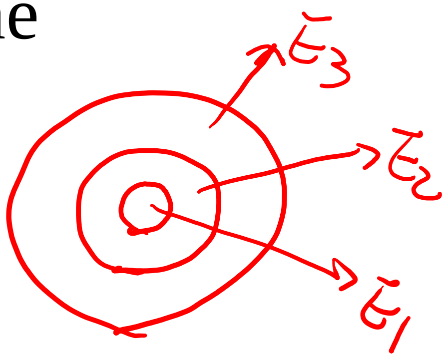
$$E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots \subseteq E_n \subseteq E_{n+1} \cdots$$

- It is called decreasing if

$$E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots \supseteq E_n \supseteq E_{n+1} \supseteq \cdots$$

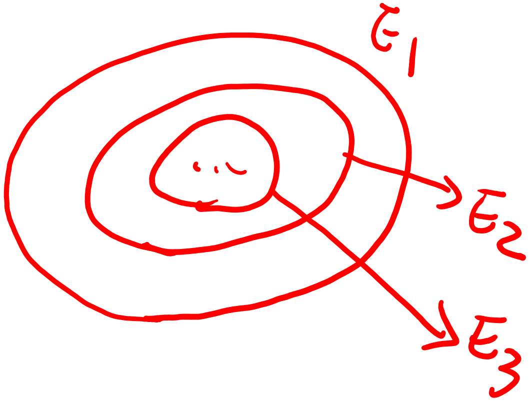
- For an increasing sequence of events, define

$$\lim_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} E_n.$$



- For a decreasing sequence of events, define

$$\lim_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n$$



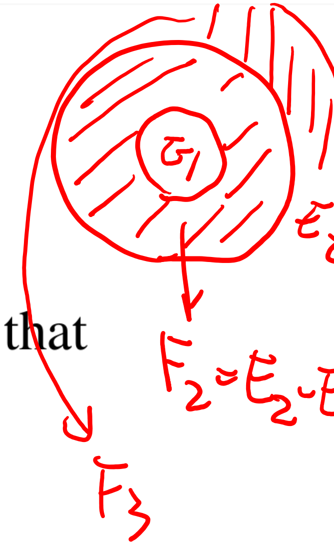
Theorem 1.8 Continuity of Probability Function

- **Theorem 1.18** *For any increasing or decreasing sequence of events,* $\{E_n, n \geq 1\}$,

$$\lim_{n\to\infty} P(E_n) = P\left(\lim_{n\to\infty} E_n\right).$$

- **Proof of Theorem 1.8.**
- For increasing sequence $\{E_n, n \geq 1\}$, define

$$F_1 = E_1, \quad F_n = E_n - E_{n-1}, \quad n = 2, 3, \ldots$$



- Clearly, $\{F_i, i \geq 1\}$ is a mutually exclusive set of events that satisfies the following relations:

$$\bigcup_{i=1}^{n} F_i = \bigcup_{i=1}^{n} E_i = E_n, \qquad n = 1, 2, \ldots$$

$$\bigcup_{i=1}^{\infty} F_i = \bigcup_{i=1}^{\infty} E_i$$

axiom3

- Hence,

$$P\left(\lim_{n\to\infty} E_n\right) = P\left(\bigcup_{i=1}^{\infty} E_i\right) = P\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} P(F_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(F_i)$$

$$= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} F_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} E_i\right) = \lim_{n\to\infty} P(E_n)$$

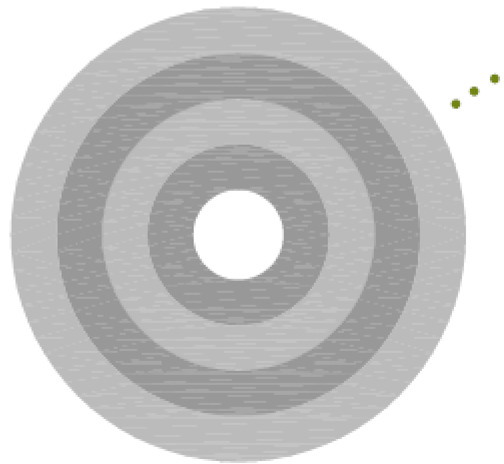


Figure 1.5 The circular disks are the E_i's and the shaded circular annuli are the F_i's, except for F_1, which equals E_1.

- Now consider decreasing sequence $\{E_n, n \geq 1\}$.
- Sequence $\{E_n^c, n \geq 1\}$ is an increasing sequence.
- It follows from the previous analysis that

$$
\begin{aligned}
P\left(\lim_{n\to\infty} E_n\right) &= P\left(\bigcap_{i=1}^{\infty} E_i\right) = 1 - P\left[\left(\bigcap_{i=1}^{\infty} E_i\right)^c\right] = 1 - P\left(\bigcup_{i=1}^{\infty} E_i^c\right) \\
&= 1 - P\left(\lim_{n\to\infty} E_n^c\right) = 1 - \lim_{n\to\infty} P(E_n^c) \\
&= 1 - \lim_{n\to\infty} [1 - P(E_n)] \\
&= 1 - 1 + \lim_{n\to\infty} P(E_n) \\
&= \lim_{n\to\infty} P(E_n).
\end{aligned}
$$

$P(A^c) = 1 - P(A)$

Example 1.21

- Suppose that some individuals in a population produce offspring of the same kind.
- The offspring of the initial population are called *second generation*
- The offspring of the second generation are called *third generation*, and so on.
- If with probability

$$\exp\left(-\frac{2n^2+7}{6n^2}\right)$$

the entire population completely dies out by the nth generation generation before producing any offspring, what is the probability that such a population survives forever?

- **Solution.**
- Let E_n denote the event of extinction of the entire population by the nth generation; then

$$E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n \subseteq \cdots .$$

- If E_n occurs, then E_{n+1} also occurs.
- Hence, by Theorem 1.8,

$$\begin{aligned}
&P(\{\text{population survives forever}\})\\
&= 1 - P(\{\text{population eventually dies out}\})\\
&= 1 - P\left(\bigcup_{i=1}^{\infty} E_i\right) = 1 - \lim_{n\to\infty} P(E_n)\\
&= 1 - \lim_{n\to\infty} \exp\left(-\frac{2n^2+7}{6n^2}\right)\\
&= 1 - e^{-1/3}
\end{aligned}$$

Section 1.6

Probabilities 0 and 1

1.6 Probabilities 0 and 1

- It is **correct** that
  - $P(\emptyset) = 0$;
  - $P(S) = 1$.
- However, it is **incorrect** to say that
  - the empty set is the only event that has probability 0;
  - the sample space is only event that has probability 1.
- There are infinitely many events that
  - have probability 0;
  - have probability 1.

An Example

- Randomly select a point from the interval (0,1).
- Each point in (0,1) has a decimal representation

$$0.529387043219721\cdots$$

- Let A_n be the event that the selected decimal has 3 as its first n digits; then

$$A_1 \supset A_2 \supset A_3 \supset A_4 \supset \cdots \supset A_n \supset A_{n+1} \supset \cdots$$

$$P(A_n) = (1/10)^n$$

$$P\left(\frac{1}{3} \text{ is selected}\right) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} \left(\frac{1}{10}\right)^n = 0.$$

- $P\left((0,1) - \left\{\frac{1}{3}\right\}\right) = 1$