Ch1: Axioms of Probability

1.2 Sample space and events

Experiment (eg. Tossing a die)
Outcome (sample point)
S (sample space) ={all possible outcomes}
Event: subset of sample space

- Ex1.1: tossing a coin once
  S = {H, T}
- Ex1.2: flipping a coin and tossing a die if T
  or flipping a coin again if H
  S={ }

Sample space and events

Ex 1.5: A bus with a capacity of 34 passengers stops at a station some time between 11:00 A.M. and 11:40 A.M. every day.

- sample space of the experiment, consisting of counting the number of passengers on the bus and measuring the arrival time of the bus
- $S = \{(i,t) : \qquad\qquad\qquad\qquad \}$
- event that the bus arrives between 11:20 A.M. and 11:40 A.M. with 27 passengers.
- $F = \{(27,t) : \qquad\qquad\qquad \}$

Sample space and events

- Event E has occurred in an experiment: If the outcome of an experiment belongs to E.

Relations between different events

- Subset: $E \subseteq F$
- Equality: if $E \subseteq F$ and $F \subseteq E$, hence $E = F$
- Intersection: EF or $E \cap F$
- Union: $E \cup F$
- Complement: E^C
- Difference: $E - F$
  - $E^C = S - E$ $\qquad E - F =$

Relations between different events

- ***Certainty***: An event is called **certain** if its occurrence is inevitable.
  - The sample space is a certain event.
- ***Impossibility***: An event is called **impossible** if there is certainty in its nonoccurrence.
  - The empty set ∅, which is S^c, is an impossible event.

Relations between different events

- ***Mutually Exclusiveness*: If the joint occurrence of two events *E* and *F* is impossible, we say that *E* and *F* are **mutually exclusive**.
  - *E* and *F* are mutually exclusive if .
- A set of events $\{E_1, E_2, \ldots\}$ is called **mutually exclusive** if the joint occurrence of any two of them is impossible, that is, if $\forall i \neq j$, $E_i E_j =$.

Relations between different events

- If $\{E_1, E_2, \ldots, E_n\}$ is a set of events, by $\cup_{i=1}^{n} E_i$ we mean the event in which at least one of the events Ei, $1 \le i \le n$, occurs.
- By $\cap_{i=1}^{n} E_i$, we mean an event that occurs only when all of the events Ei, $1 \le i \le n$, occur.

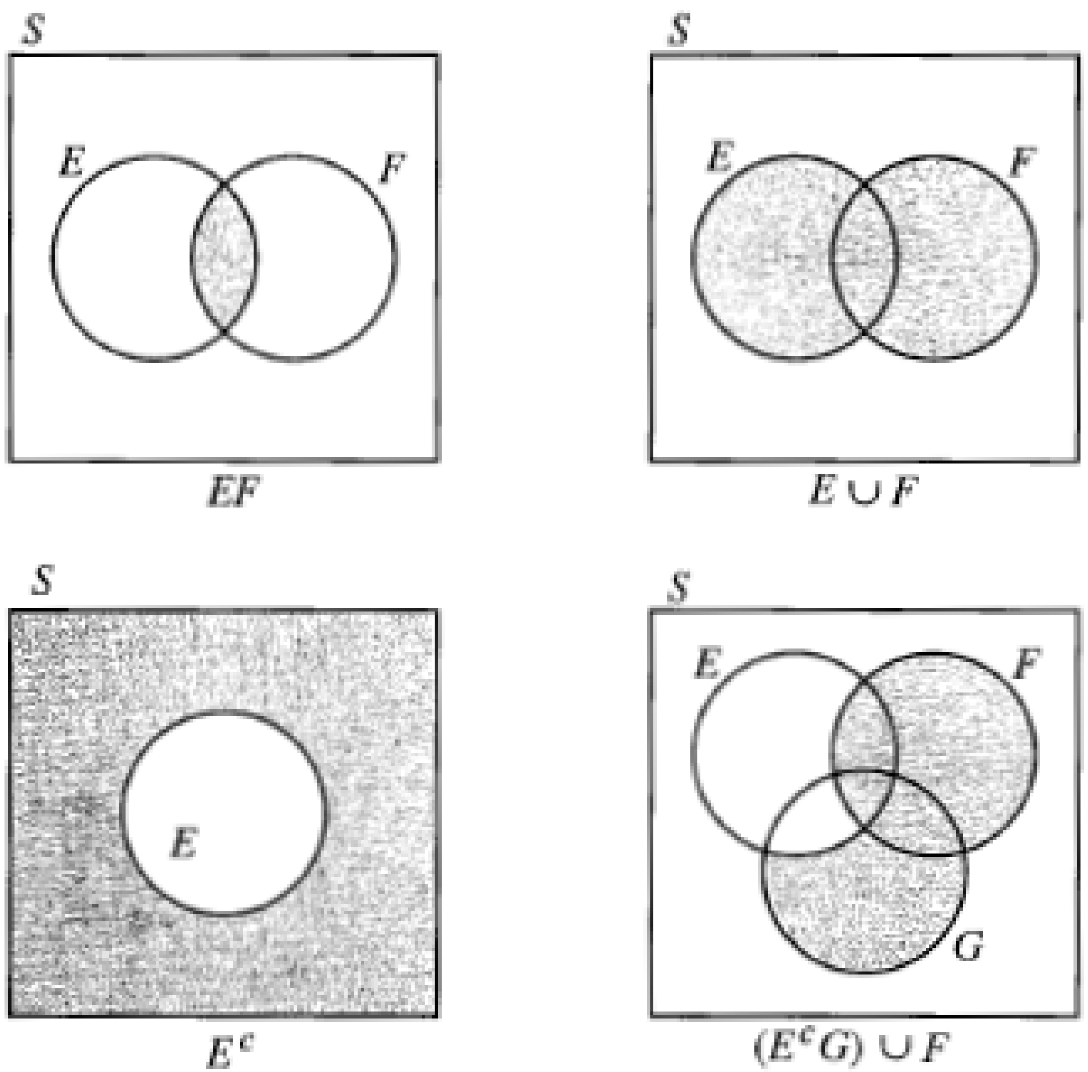

Figure 1.1 Venn diagrams of the events specified.

Sample space and events

- Commutative laws $E \cup F = F \cup E, EF = FE$
- Associative laws
$$E \cup (F \cup G) = (E \cup F) \cup G, E(FG) = (EF)G$$
- Distributive laws
$$(EF) \cup H = (E \cup H)(F \cup H), (E \cup F)H = EH \cup FH$$
- De Morgan's 1st law:
$$(E \cup F)^C = E^C F^C, \left(\bigcup_{i=1}^{n} E_i\right)^C = \bigcap_{i=1}^{n} E_i^C, \left(\bigcup_{i=1}^{\infty} E_i\right)^C = \bigcap_{i=1}^{\infty} E_i^C$$
- De Morgan's 2nd law:
$$(EF)^C = E^C \cup F^C, \left(\bigcap_{i=1}^{n} E_i\right)^C = \bigcup_{i=1}^{n} E_i^C, \left(\bigcap_{i=1}^{\infty} E_i\right)^C = \bigcup_{i=1}^{\infty} E_i^C$$
- E = ES =

Prove De Morgan's first law:

- First show that $(E \cup F)^C \subseteq E^C F^C$

  Then $E^C F^C \subseteq (E \cup F)^C$

1.3 Axioms of probability

Definition (Probability Axioms):

- ***S***: *the sample space of a random phenomenon*
- ***A***: *an event of S*
- ***P***: *a real-valued function for each event A, i.e., P:* $2^S \rightarrow R$
- *If P satisfies the following axioms, then it is called a* ***probability*** *and the number P(A) is said to be the* ***probability of A***.

1.3 Axioms of probability

Axiom 1 $P(A) \geq 0$

Axiom 2 $P(\varsigma) = 1$

Axiom 3 If {A1, A2, A3, …} is a sequence of mutually exclusive events (i.e. the joint occurrence of every pair of them is impossible: $A_i A_j = \phi$ when $i \neq j$), then

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

Equally likely

- Let S be the sample space of an experiment. Let A and B be events of S.
  - A and B are **equally likely** if

- Let ω_1 and ω_2 be sample points of S.
  - ω_1 and ω_2 are **equally likely** if

Theorem 1.1

- *The probability of the empty set ∅ is 0. That is, P(∅) =   .*
- *Proof:* Let $A_1 = \varsigma$ and $A_i = \phi$ for $i \geq 2$; then $A_1, A_2, A_3, \ldots$ is a sequence of mutually exclusive events. Thus, by Axiom 3,

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

implying that $\sum_{i=2}^{\infty} P(\phi) = 0$. This is only possible only if $P(\phi) = 0$.

Theorem 1.2

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

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

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

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

$$= \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(A_i) = \sum_{i=1}^{n} P(A_i) + \qquad = \sum_{i=1}^{n} P(A_i)$$

- If a sample space contains N points that are equally likely to occur, then the probability of each outcome (sample point) is $1/N$.

1. $S = \{S_1, S_2, \ldots, S_N\}$
2. $P(\{S_1\}) = P(\{S_2\}) = \ldots = P(\{S_N\})$
3. $\because P(S) = 1$, and events are mutually exclusive

   $\therefore 1 = P(S) = P(\{S_1, S_2, \ldots, S_N\})$

   $=$

4. $P(\{S_1\}) = \dfrac{1}{N}$, Thus $P(\{S_i\}) = \dfrac{1}{N}$

Theorem 1.3

- *If the sample space S of an experiment has N points that are all equally likely to occur, then*
  - *for any event A of S, $P(A) =$*

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

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?

1. $S =$ $\{a_2b_2, a_2b_3, a_2b_4, a_3b_2, a_3b_3, a_3b_4, a_4b_2, a_4b_3, a_4b_4\}$
2. A = $\{a_2b_3, a_2b_4, a_3b_2, a_3b_4, a_4b_2, a_4b_3\}$
3. N = , N(A) = , -> N(A)/N = 2/3

1.4 Basic Theorems

- Theorem 1.4: *For any event A,*
  $P(A^c) = 1 - P(A)$

1. $\because AA^C = \phi, \therefore\ A$ and A^C are mutually exclusive, Thus

   $$P\left(A \cup A^C\right) = P(A) + P\left(A^C\right)$$

2. $A \cup A^C =$ and $P(S) = 1$

   So $1 = P(S) = P\left(A \cup A^C\right) =$

3. $$P\left(A^C\right) = 1 - P(A)$$

Theorem 1.5

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

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

1. From Fig1.2, $A \subseteq B$, $B = (B - A) \cup A$
2. <- mutually exclusive
3. $P(B) = P((B - A) \cup A) = P(B - A) + P(A)$

   -> $P(B - A) = P(B) - P(A)$

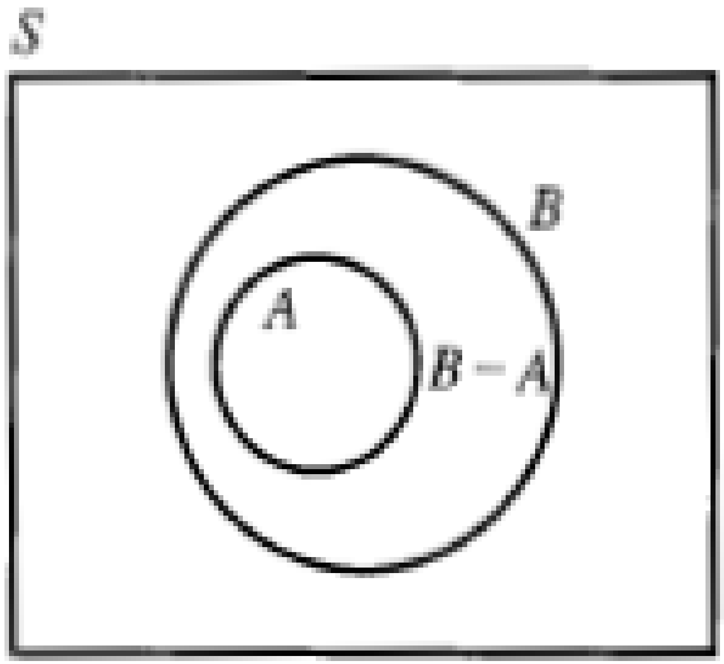

Figure 1.2 $A \subseteq B$ implies that $B = (B - A) \cup A$.

Corollary

- *If $A \subseteq B$, then $P(A) \leq P(B)$.*

1. By Theorem 1.5 $P(B-A)=P(B)-P(A)$
2. $P(B-A)\geq 0$ →
3. $P(B)\geq P(A)$

Theorem 1.6

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

1. $A \cup B = A \cup (B - AB) (Fig.1.3)$
2. $A(B - AB) = \phi$ → mutually exclusive
3. $P(A \cup B) = P(\qquad\qquad) = P(\quad) + P(\qquad\quad)$
4. $AB \subseteq B$, Theorem implies that

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

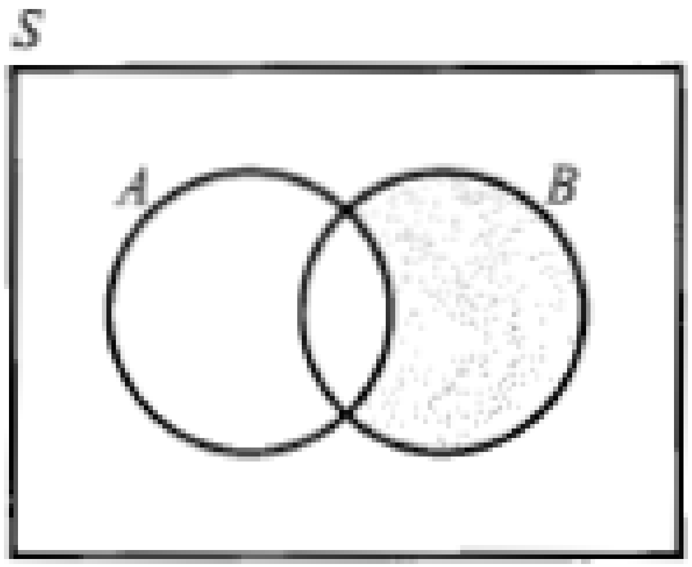

Figure 1.3 The shaded region is $B - AB$. Thus $A \cup B = A \cup (B - AB)$.

Inclusion-Exclusion Principle

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

$$= \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} P(A_i A_j) + \sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n} P(A_i A_j A_k)$$

$$- \ldots + (-1)^{n-1} P(A_1 A_2 A_3 \ldots A_n)$$

Example 1.17

- Suppose that 25% of the population of a city read newspaper A, 20% read newspaper B, 13% read C, 10% read both A and B, 8% read both A and C, 5% read B and C, and 4% read all three. If a person from this city is selected at random, what is the probability that he or she does not read any of these newspapers?

1. E, F, G : events that the person reads A, B, and C.
2. $P(E \cap F \cap G)$

$$=$$

$$= 0.25 + 0.2 + 0.13 - 0.1 - 0.08 - 0.05 + 0.04$$

$$= 0.39$$

Theorem 1.7

- $P(A) = P(AB) + P(AB^c)$.

1. $A = AS = A(B \cup B^c) = AB \cup AB^c$
2. AB and AB^c are mutually exclusive
3. $P(A) = P(AB \cup AB^c) = P(AB) + P(AB^c)$

Example 1.19

- In a community, 32% of the population are male smokers; 27% are female smokers. What percentage of the population of this community smoke?

1. A: event that a randomly selected person from this community smokes.

   B: event that the person is male.

2. $P(A) = \qquad\qquad\qquad = 0.32 + 0.27 = 0.59$

1.5 Continuity of probability function

- f : **R** $\rightarrow$ **R** is called continuous at a point $c \in$ **R** if
  - **R** denotes the set of all real numbers.
- f : **R** $\rightarrow$ **R** *is continuous on* **R** *if and only if, for every convergent sequence* $\{x_n\}_{n=1}^{\infty}$ *in* **R***,* $\lim_{n \to \infty} f(x_n) =$

Increasing

- A sequence $\{E_n, n \geq 1\}$ of events of a sample space is called **increasing** if
  $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots \subseteq E_n \subseteq E_{n+1} \cdots;$
- For an increasing sequence of events $\{En, n \geq 1\}$, by $\lim_{n\to\infty} E_n$ we mean the event that **at least** one Ei , $1 \leq i < \infty$ occurs. Therefore,

$$\lim_{n\to\infty} E_n =$$

Decreasing

- A sequence $\{E_n, n \geq 1\}$ of events of a sample space is called **decreasing** if

  $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots \supseteq E_n \supseteq E_{n+1} \supseteq \cdots$.
- For a decreasing sequence of events $\{E_n, n \geq 1\}$, by $\lim_{n\to\infty} E_n$ we mean the event that **every** E_i occurs. Therefore,

$$\lim_{n\to\infty} E_n =$$

Theorem 1.8

- **(Continuity of Probability Function)** *For any increasing or decreasing sequence of events, {En, n ≥ 1},*

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

1. Let $F_1=E_1$, $F_2=E_2-E_1$, …$F_n=E_n-E_{n-1}$
2. $\{F_i, i \geq 1\}$ is a mutually exclusive set of events that satifies

$$(1)\bigcup_{i=1}^{n} F_i = \bigcup_{i=1}^{n} E_i = E_n \text{ ,n=1,2,…}$$

$$(2)\bigcup_{i=1}^{\infty} F_i = \bigcup_{i=1}^{\infty} E_i$$

Theorem 1.8

3. If $\{E_n, n \geq 1\}$ is increasing

$$P(\lim_{n\to\infty} E_n) = \qquad = \qquad = \sum_{i=1}^{\infty} P(F_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(F_i)$$

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

4. If $\{E_n, n \geq 1\}$ is decreasing

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

$$= 1 - P(\lim_{n\to\infty} E_n^c) = 1 - \lim_{n\to\infty} P(E_n^c)$$

$$= 1 - \lim_{n\to\infty} [1 - P(E_n)] = 1 - 1 + \lim_{n\to\infty} P(E_n) = \lim_{n\to\infty} P(E_n)$$

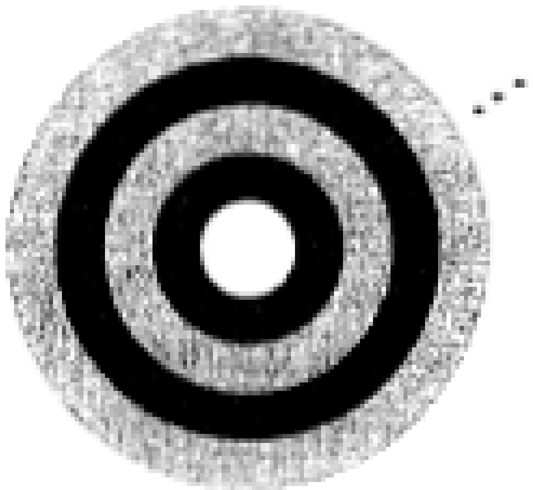

Figure 1.5 The circular disks are the E_i's and the shaded circular annuli are the F_i's, except for F_1, which equals E_1.

Example 1.20

- Suppose that some individuals in a population produce offspring of the same kind.
- If with prob. $exp[-(2n^2+7)/(6n^2)]$ the entire population completely dies out by the nth generation before producing any offspring.
- What is the prob. that such a population survives forever?

Example 1.20

- Ans: Let En denote the event of extinction of the entire population by the nth generation
- Then $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots \subseteq E_n \subseteq E_{n+1} \cdots;$

  because if En occurs then En+1 occurs.

  Hence by Theorem 1.8

  p[survives forever] = 1 – p[eventually dies out]

  $= 1 - p(\bigcup_{i=1}^{\infty} E_i)$

  $= 1 - \lim_{n\to\infty} P(E_n)$

  $= 1 - \lim_{n\to\infty} (exp[-(2n^2+7)/(6n^2)]) = 1 - \exp(-1/3)$

1.6 Probabilities 0 and 1

- If E and F are events with probabilities 1 and 0, respectively,
  - it is not correct to say that E is the sample space S and F is the empty set $\emptyset$.

Example

- Suppose that an experiment consists of selecting a random point from the interval(0, 1).
- Since every point in (0, 1) has a decimal representation such as 0.529387043219721 · · · , the experiment is equivalent to picking an endless decimal from (0, 1) at random (note that if a decimal terminates, all of its digits from some point on are 0).
- In such an experiment we want to compute the probability of selecting the point 1/3.
- In other words, we want to compute the probability of choosing in a random selection of an endless decimal.

1. A_n: event that the selected decimal has 3 as its first *n* digits. Then $A_1 \supset A_2 \supset \ldots \supset A_n \supset A_{n+1} \supset \ldots$

2. $P(A_n) =$

3. $\bigcap_{n=1}^{\infty} A_n = \qquad (Theorem 1.8)$

4. P($\frac{1}{3}$ is selected)

$$= P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} (\frac{1}{10})^n = 0$$

1.7 Random selection of points from intervals

- The probability that a random point selected from (a, b) falls into the interval (a, (a+b)/2) is 1/2. The probability that it falls into [(a+b)/2, b) is also 1/2.

  1. p1 : events that point belong to (a, (a+b)/2)

     p2 : events that point belong to [(a+b)/2, b)

  2. $(\qquad \cup \qquad) = (a,b)$

  3. p1 + p2 = 1 and p1 = p2. Therefore

     p1 = p2 = 1/2

- The probability that a random point selected from (a, b) falls into the interval (a, (2a + b)/3] is 1/3. The probability that it falls into ((2a + b)/3, (a+2b)/3] is 1/3, and the probability that it falls into ((a + 2b)/3, b) is 1/3.

1. p1 : events that point belong to (a, (2a + b)/3]
   p2 : events that point belong to ((2a + b)/3, (a+2b)/3]
   p3 : events that point belong to ((a + 2b)/3, b)

2. $(a,\frac{2a+b}{3}]\cup(\frac{2a+b}{3},\frac{a+2b}{3}]\cup(\frac{a+2b}{3},b)=(a,b)$

3. p1 + p2 + p3 = 1 and p1 = p2 = p3. Therefore

Definition

- A point is said to be **randomly selected from an interval** (a, b) if
  - any two subintervals of (a, b) that have the same length are equally likely to include the point.
- The probability associated with the event that the subinterval (α, β) contains the point is defined to be ( )/( ).