Introduction to Probability

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 308

1 Axioms of Probability

- Sample Space and Events
- Axioms of Probability
- Some Simple Propositions
- Sample Spaces With Equally Likely Outcomes
- Probability as a Measure of Belief

Subsection 1

Sample Space and Events

Sample Space

- Consider an experiment whose outcome is not predictable with certainty.
- However, suppose that the set of all possible outcomes is known.
- The set of all possible outcomes of an experiment is known as the **sample space** of the experiment and is denoted by S.

  Examples:

  1. Suppose the outcome of an experiment consists in the determination of the sex of a newborn child. Then $S = \{g, b\}$, where the outcome g means that the child is a girl and b that it is a boy.
  2. Suppose the outcome of an experiment is the order of finish in a race among the 7 horses having post positions 1, 2, 3, 4, 5, 6 and 7. Then

     $$S = \{\text{all } 7! \text{ permutations of } (1, 2, 3, 4, 5, 6, 7)\}.$$

     The outcome $(2, 3, 1, 6, 5, 4, 7)$ means, for instance, that the number 2 horse comes in first, then the number 3 horse, then the number 1 horse, and so on.

Examples (Cont'd)

3. Suppose the experiment consists of flipping two coins. Then the sample space consists of the following four points:

$$S = \{(H, H), (H, T), (T, H), (T, T)\}.$$

The outcome will be (H, H) if both coins are heads, (H, T) if the first coin is heads and the second tails, (T, H) if the first is tails and the second heads, and (T, T) if both coins are tails.

4. Suppose the experiment consists of tossing two dice. Then the sample space consists of the 36 points

$$S = \{(i, j) : i, j = 1, 2, 3, 4, 5, 6\},$$

where the outcome (i, j) is said to occur if i appears on the leftmost die and j on the other die.

5. Suppose the experiment consists of measuring (in hours) the lifetime of a transistor. Then the sample space consists of all nonnegative real numbers; that is, $S = \{x : 0 \leq x < \infty\} = [0, \infty)$.

Events

- Any subset E of the sample space is known as an **event**.
- In other words, an event is a set consisting of possible outcomes of the experiment.
- If the outcome of the experiment is contained in E, then we say that E **has occurred**.

  Examples:

  - In Example 1, let $E = \{g\}$. Then E is the event that the child is a girl.
  - In Example 2, let $E = \{\text{all outcomes in } S \text{ starting with a } 3\}$. Then E is the event that horse 3 wins the race.
  - In Example 3, let $E = \{(H,H),(H,T)\}$. Then E is the event that a head appears on the first coin.
  - In Example 4, let $E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$. Then E is the event that the sum of the dice equals 7.
  - In Example 5, let $E = \{x : 0 \leq x \leq 5\}$. Then E is the event that the transistor lasts no longer than 5 hours.

Union of Two Events

- For any two events E and F of a sample space S, we define the new event $E \cup F$ to consist of all outcomes that are either in E or in F or in both E and F.
- That is, the event $E \cup F$ will occur if either E or F occurs.

  Example: In Example 1, suppose event $E = \{g\}$ and $F = \{b\}$. Then $E \cup F = \{g, b\}$. That is, $E \cup F$ is the whole sample space S.

  In Example 3, suppose $E = \{(H, H), (H, T)\}$ and $F = \{(T, H)\}$. Then
  $$E \cup F = \{(H, H), (H, T), (T, H)\}.$$
  Thus, $E \cup F$ would occur if at least one head appeared.
- The event $E \cup F$ is called the **union** of the event E and the event F.

Intersection of Two Events

- For any two events E and F, we may also define the new event EF, called the **intersection** of E and F, to consist of all outcomes that are both in E and in F.
- That is, the event EF (sometimes written $E \cap F$) will occur only if both E and F occur.

  Example: In Example 3, suppose:

  - $E = \{(H,H),(H,T),(T,H)\}$ is the event that at least 1 head occurs;
  - $F = \{(H,T),(T,H),(T,T)\}$ is the event that at least 1 tail occurs.

  Then $EF = \{(H,T),(T,H)\}$, i.e., it is the event that exactly 1 head and 1 tail occur.

The Null Event and Mutually Exclusive Events

Example: In example 4, suppose:

- $E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$ is the event that the sum of the dice is 7;
- $F = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$ is the event that the sum is 6.

Then the event EF does not contain any outcomes.

Hence EF could never occur.

- To give such an event a name, we shall refer to it as the **null event** and denote it by $\emptyset$.
- That is, $\emptyset$ refers to the event consisting of no outcomes.
- If $EF = \emptyset$, then E and F are said to be **mutually exclusive**.

Union and Intersection of Multiple Events

- We define unions and intersections of more than two events in a way similar to those for two events.
- If $E_1, E_2, \ldots$ are events, then the **union** of these events, denoted by $\bigcup_{n=1}^{\infty} E_n$, is defined to be that event which consists of all outcomes that are in E_n for at least one value of $n = 1, 2, \ldots$:

$$\bigcup_{n=1}^{\infty} E_n = \{x : (\exists n)(x \in E_n)\}.$$

- The **intersection** of the events E_n, denoted by $\bigcap_{n=1}^{\infty} E_n$, is defined to be the event consisting of those outcomes which are in all of the events E_n, $n = 1, 2, \ldots$:

$$\bigcap_{n=1}^{\infty} E_n = \{x : (\forall n)(x \in E_n)\}.$$

The Complement of an Event

- For any event E, we define the new event E^c, referred to as the **complement** of E, to consist of all outcomes in the sample space S that are not in E.
- That is, E^c will occur if and only if E does not occur.

  Example: In Example 4, suppose event

  $$E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}.$$

  Then E^c will occur when the sum of the dice does not equal 7.
- Note that because the experiment must result in some outcome, it follows that $S^c = \emptyset$.

Subsets

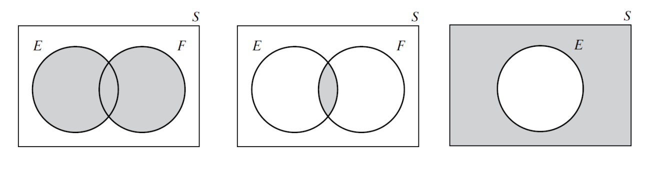
- For any two events E and F, if all of the outcomes in E are also in F, then we say that
  - E is **contained** in F, or
  - E is a **subset** of F, or
  - F is a **superset** of E,

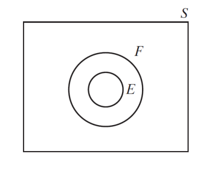
  and write
  $$E \subseteq F \quad \text{or} \quad F \supseteq E.$$
- Thus, if $E \subseteq F$, then the occurrence of E implies the occurrence of F.
- If $E \subseteq F$ and $F \subseteq E$, we say that E and F are **equal** and write
  $$E = F.$$

Venn Diagrams

- The sample space S is represented as consisting of all the outcomes in a large rectangle.
- The events $E, F, G, \ldots$ are represented as consisting of all the outcomes in given circles within the rectangle.

  Example: In the three Venn diagrams shown below, the shaded areas represent, respectively, the events $E \cup F$, EF, and E^c.





- The Venn diagram on the right indicates that $E \subseteq F$.

Set-Theoretic Identities

- The operations of forming unions, intersections and complements of events obey certain rules similar to the rules of algebra:

Commutative Laws:
$E \cup F = F \cup E \qquad E \cap F = F \cap E;$
Associative Laws:
$(E \cup F) \cup G = E \cup (F \cup G) \qquad (E \cap F) \cap G = E \cap (F \cap G);$
Distributive Laws:
$(E \cup F) \cap G = (E \cap G) \cup (F \cap G) \quad (E \cap F) \cup G = (E \cup G) \cap (F \cup G).$

- These relations are verified by showing that any outcome that is contained in the event on the left side of the equality sign is also contained in the event on the right side, and vice versa.

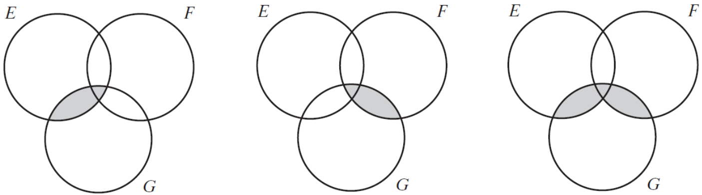
Demonstrations Using Venn Diagrams

- One way of informally "proving" identities is by means of Venn diagrams.

  Example: The distributive law

$$(E \cup F)G = EG \cup FG$$

  may be verified by the sequence of diagrams in the figure:



(a) Shaded region: EG. (b) Shaded region: FG. (c) Shaded region: $(E \cup F)G$.

$(E \cup F)G = EG \cup FG$

De Morgan's Laws

- **De Morgan's Laws:**

$$(\bigcup_{i=1}^{n} E_i)^c = \bigcap_{i=1}^{n} E_i^c, \qquad (\bigcap_{i=1}^{n} E_i)^c = \bigcup_{i=1}^{n} E_i^c.$$

Suppose first that x is an outcome of $(\bigcup_{i=1}^{n} E_i)^c$. Then x is not contained in $\bigcup_{i=1}^{n} E_i$. So x is not contained in any of the events E_i, $i = 1, 2, \ldots, n$. Thus, x is contained in E_i^c for all $i = 1, 2, \ldots, n$. So x is contained in $\bigcap_{i=1}^{n} E_i^c$.

To go the other way, suppose that x is an outcome of $\bigcap_{i=1}^{n} E_i^c$. Then x is contained in E_i^c for all $i = 1, 2, \ldots, n$. This means that x is not contained in E_i for any $i = 1, 2, \ldots, n$. So x is not contained in $\bigcup_{i=1}^{n} E_i$. This implies that x is contained in $(\bigcup_{i=1}^{n} E_i)^c$.

This proves the first of De Morgan's Laws.

De Morgan's Laws (Cont'd)

- To prove the second of De Morgan's laws, we use the first law to obtain
$$\left(\bigcup_{i=1}^{n} E_i^c\right)^c = \bigcap_{i=1}^{n} (E_i^c)^c.$$
Since $(E^c)^c = E$, this is equivalent to
$$\left(\bigcup_{i=1}^{n} E_i^c\right)^c = \bigcap_{i=1}^{n} E_i.$$
Taking complements of both sides of the preceding equation yields the result we seek,
$$\bigcup_{i=1}^{n} E_i^c = \left(\bigcap_{i=1}^{n} E_i\right)^c.$$

Subsection 2

Axioms of Probability

Axioms of Probability

- Consider an experiment whose sample space is S.
- For each event E of the sample space S, we assume that a number $P(E)$ is defined and satisfies the following three axioms:

Axiom 1 $0 \leq P(E) \leq 1$;

Axiom 2 $P(S) = 1$

Axiom 3 For any sequence of mutually exclusive events $E_1, E_2, \ldots$ (that is, events for which $E_i E_j = \emptyset$ when $i \neq j$),

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i).$$

- We refer to $P(E)$ as the **probability of the event** E.

Probability of the Null Event

- Consider a sequence of events $E_1, E_2, \ldots$, where:
  - $E_1 = S$;
  - $E_i = \emptyset$ for $i > 1$.
- These events are mutually exclusive.
- Moreover, $S = \bigcup_{i=1}^{\infty} E_i$.
- Hence, from Axiom 3,

$$P(S) = \sum_{i=1}^{\infty} P(E_i) = P(S) + \sum_{i=2}^{\infty} P(\emptyset).$$

- Taking into account Axiom 1, we get $P(\emptyset) = 0$.
- That is, the null event has probability 0 of occurring.

Finite Sample Spaces and Axiom 3

- Note that it follows that, for any finite sequence of mutually exclusive events $E_1, E_2, \ldots, E_n$,

$$P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i).$$

- This equation follows from Axiom 3 by defining E_i as the null event for all values of i greater than n.
- This equation is equivalent to Axiom 3 when the sample space is finite.
- However, the added generality of Axiom 3 is necessary when the sample space consists of an infinite number of points.

Example

- Suppose our experiment consists of tossing a coin.
  Assume that a head is as likely to appear as a tail.
  Then we would have
  $$P(\{H\}) = P(\{T\}) = \frac{1}{2}.$$
- Suppose, now, the coin were biased.
  Assume that a head were twice as likely to appear as a tail.
  Then we would have
  $$P(\{H\}) = \frac{2}{3}, \quad P(\{T\}) = \frac{1}{3}.$$

Example

- Suppose a die is rolled.

  Assume that all six sides are equally likely to appear.

  Then we would have

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}.$$

  From Axiom 3, it would thus follow that the probability of rolling an even number would equal

$$P(\{2,4,6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{2}.$$

Subsection 3

Some Simple Propositions

Probability of Complement

Proposition

Suppose E is an event in a sample space S. We then have

$$P(E^c) = 1 - P(E).$$

- E and E^c are always mutually exclusive.

  Moreover, $E \cup E^c = S$.

  Hence, by Axioms 2 and 3,

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

  Example: Consider tossing a coin. Suppose that $P(\{H\}) = \frac{3}{8}$.
  It then follows that $P(\{T\}) = 1 - \frac{3}{8} = \frac{5}{8}$.

Probability of Subset

Proposition

Let E, F be events in a sample space S. If $E \subseteq F$, then $P(E) \leq P(F)$.

- Since $E \subseteq F$, it follows that we can express F as
$$F = E \cup E^c F.$$
The events E and $E^c F$ are mutually exclusive. Hence, from Axiom 3,
$$P(F) = P(E) + P(E^c F).$$
This proves the result, since $P(E^c F) \geq 0$.

Example: Consider tossing a die. We have $\{1\} \subseteq \{1, 3, 5\}$. Hence, $P(\{1\}) \leq P(\{1, 3, 5\})$. So the probability of rolling a 1 is less than or equal to the probability of rolling an odd value.

Probability of Unions and Intersections

Proposition

Let E, F be events in a sample space S. Then,

$$P(E \cup F) = P(E) + P(F) - P(EF).$$

- Note that $E \cup F$ can be written as the union of the two disjoint events E and $E^c F$. Thus, from Axiom 3, we obtain

$$P(E \cup F) = P(E \cup E^c F) = P(E) + P(E^c F).$$

Also $F = EF \cup E^c F$. Hence, from Axiom 3, we get $P(F) = P(EF) + P(E^c F)$ or, equivalently,

$$P(E^c F) = P(F) - P(EF).$$

Now we get $P(E \cup F) = P(E) + P(E^c F) = P(E) + P(F) - P(EF)$.

Example

- Antonia is taking two books along on her holiday vacation.
  - With probability 0.5, she will like the first book;
  - With probability 0.4, she will like the second book;
  - With probability 0.3, she will like both books.

  What is the probability that she likes neither book?

  Let B_i denote the event that Antonia likes book i, $i = 1, 2$.

  Then the probability that she likes at least one of the books is

  $$P(B_1 \cup B_2) = P(B_1) + P(B_2) - P(B_1 B_2) = 0.5 + 0.4 - 0.3 = 0.6.$$

  The event that Antonia likes neither book is the complement of the event that she likes at least one of them.

  Hence, we obtain the result

  $$P(B_1^c B_2^c) = P((B_1 \cup B_2)^c) = 1 - P(B_1 \cup B_2) = 0.4.$$

Inclusion-Exclusion for Three Events

- We calculate the probability that any one of the three events E, F and G occurs, namely,
$$P(E \cup F \cup G) = P[(E \cup F) \cup G].$$
- By the proposition, this equals
$$P(E \cup F) + P(G) - P[(E \cup F)G].$$
- Now, it follows from the distributive law that the events $(E \cup F)G$ and $EG \cup FG$ are equivalent.
- Hence, from the preceding equations, we obtain
$$\begin{aligned} P(E \cup F \cup G) &= P(E \cup F) + P(G) - P[(E \cup F)G] \\ &= P(E) + P(F) - P(EF) + P(G) - P(EG \cup FG) \\ &= P(E) + P(F) - P(EF) + P(G) - P(EG) - P(FG) + P(EGFG) \\ &= P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG). \end{aligned}$$

The Inclusion-Exclusion Identity

- The following proposition, can be proved by mathematical induction:

The Inclusion-Exclusion Identity

$$\begin{aligned} P(E_1 \cup E_2 \cup \cdots \cup E_n) &= \textstyle\sum_{i=1}^n P(E_i) - \sum_{i_1 < i_2} P(E_{i_1} E_{i_2}) + \cdots \\ &\quad + (-1)^{r+1} \textstyle\sum_{i_1 < i_2 < \cdots < i_r} P(E_{i_1} E_{i_2} \cdots E_{i_r}) \\ &\quad + \cdots + (-1)^{n+1} P(E_1 E_2 \cdots E_n). \end{aligned}$$

The summation $\sum_{i_1 < i_2 < \cdots < i_r} P(E_{i_1} E_{i_2} \cdots E_{i_r})$ is taken over all of the $\binom{n}{r}$ possible subsets of size r of the set $\{1, 2, \ldots, n\}$.

- In words, the probability of the union of n events equals: the sum of the probabilities of these events taken one at a time, minus the sum of the probabilities of these events taken two at a time, plus the sum of the probabilities of these events taken three at a time, and so on.

Non-Inductive Proof of Inclusion-Exclusion

- Note first that if an outcome of the sample space is not a member of any of the sets E_i, then its probability does not contribute anything to either side of the equality.
- Now, suppose that an outcome is in exactly m of the E_i's, $m > 0$.

  Then, it is in $\cup_i E_i$. So its probability is counted once in $P(\cup_i E_i)$. Moreover, it is contained in $\binom{m}{k}$ subsets of the type $E_{i_1} E_{i_2} \cdots E_{i_k}$. So, its probability is counted $\binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \cdots \pm \binom{m}{m}$ times on the right of the equality sign. Thus, for $m > 0$, we must show that

  $$1 = \binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \cdots \pm \binom{m}{m}.$$

  But $1 = \binom{m}{0}$. Hence, the preceding equation is equivalent to $\sum_{i=0}^{m} \binom{m}{i} (-1)^i = 0$. But, this equation follows from the binomial theorem, since $0 = (-1+1)^m = \sum_{i=0}^{m} \binom{m}{i} (-1)^i (1)^{m-i}$.

Bounds on the Probability of a Union

- In the inclusion-exclusion identity:
  - Going out one term results in an upper bound on the probability of the union;
  $$P(\bigcup_{i=1}^{n} E_i) \leq \sum_{i=1}^{P} (E_1);$$
  - Going out two terms results in a lower bound on the probability;
  $$P(\bigcup_{i=1}^{n} E_i) \geq \sum_{i=1}^{n} P(E_i) - \sum_{j<i} P(E_1 E_j);$$
  - Going out three terms results in an upper bound on the probability;
  $$P(\bigcup_{i=1}^{n} E_i) \leq \sum_{i=1}^{n} P(E_i) - \sum_{j<i} P(E_i E_j) + \sum_{k<j<i} P(E_i E_j E_k);$$
  - Going out four terms results in a lower bound, and so on.

Proof of the Bounds

- Note the identity
$$\bigcup_{i=1}^{n} E_i = E_1 \cup E_1^c E_2 \cup E_1^c E_2^c E_3 \cup \cdots \cup E_1^c \cdots E_{n-1}^c E_n.$$
The right-hand side is the union of disjoint events. Thus, we obtain
$$\begin{array}{rcl} P(\bigcup_{i=1}^n E_i) & = & P(E_1) + P(E_1^c E_2) + P(E_1^c E_2^c E_3) + \cdots \\ & & + P(E_1^c \cdots E_{n-1}^c E_n) \\ & = & P(E_1) + \sum_{i=2}^n P(E_1^c \cdots E_{i-1}^c E_i). \end{array}$$
Let $B_i = E_1^c \cdots E_{i-1}^c = (\bigcup_{j<i} E_j)^c$. But $P(E_i) = P(B_i E_i) + P(B_i^c E_i)$. So $P(E_i) = P(E_1^c \cdots E_{i-1}^c E_i) + P(E_i(E_1^c \cdots E_{i-1}^c)^c) = P(E_1^c \cdots E_{i-1}^c E_i) + P(E_i \bigcup_{j<i} E_j)$. Equivalently, $P(E_1^c \cdots E_{i-1}^c E_i) = P(E_i) - P(\bigcup_{j<i} E_i E_j)$. Substituting this into the previous one,
$$P(\bigcup_{i=1}^{n} E_i) = \sum_i P(E_i) - \sum_i P(\bigcup_{j<i} E_i E_j).$$
Because probabilities are always nonnegative, the inequality follows.

Proof of the Bounds (Cont'd)

- Fix i. Apply the proven inequality to $P(\bigcup_{j<i} E_iE_j)$.

$$P(\cup_{j<i}E_iE_j) \le \sum_{j<i} P(E_iE_j).$$

  But $P(E_1^c \cdots E_{i-1}^c E_i) = P(E_i) - P(\bigcup_{j<i} E_iE_j)$. This yields the second bound.

- Fix i. Apply the second bound to $P(\bigcup_{j<i} E_iE_j)$.

$$\begin{array}{rcl} P(\bigcup_{j<i} E_iE_j) & \ge & \sum_{j<i} P(E_iE_j) - \sum_{k<j<i} P(E_iE_jE_iE_k) \\ & = & \sum_{j<i} P(E_iE_j) - \sum_{k<j<i} P(E_iE_jE_k). \end{array}$$

  Use $P(E_1^c \cdots E_{i-1}^c E_i) = P(E_i) - P(\bigcup_{j<i} E_iE_j)$. This gives the third bound.

- The next inclusion-exclusion inequality is obtained by fixing i and applying the third bound to $P(\bigcup_{j<i} E_iE_j)$, and so on.

Subsection 4

Sample Spaces With Equally Likely Outcomes

Sample Spaces With Equally Likely Outcomes

- Consider an experiment whose sample space S is a finite set, say, $S = \{1, 2, \ldots, N\}$.
- It is often natural to assume that $P(\{1\}) = P(\{2\}) = \cdots = P(\{N\})$. This implies, from Axioms 2 and 3, that

$$P(\{i\}) = \frac{1}{N}, \quad i = 1, 2, \ldots, N.$$

- From this equation, it follows from Axiom 3 that, for any event E,

$$P(E) = \frac{\text{number of outcomes in } E}{\text{number of outcomes in } S}.$$

- In words, if we assume that all outcomes of an experiment are equally likely to occur, then:

  the probability of any event E equals the proportion of outcomes in the sample space that are contained in E.

Example

- If two dice are rolled, what is the probability that the sum of the upturned faces will equal 7?

  We solve this problem under the assumption that all of the 36 possible outcomes are equally likely.

  There are 6 possible outcomes that result in the sum of the dice being equal to 7:

$$(1,6),\ (2,5),\ (3,4),\ (4,3),\ (5,2),\ (6,1).$$

  Hence, the desired probability is $\frac{6}{36} = \frac{1}{6}$.

Example

- If 3 balls are "randomly drawn" from a bowl containing 6 white and 5 black balls, what is the probability that one of the balls is white and the other two black?

  If we regard the order in which the balls are selected as being relevant, then the sample space consists of $11 \cdot 10 \cdot 9 = 990$ outcomes.

  Furthermore, there are three possibilities:
  - The first ball selected is white and the other two are black: $6 \cdot 5 \cdot 4 = 120$ outcomes;
  - The first ball is black, the second is white, and the third is black: $5 \cdot 6 \cdot 4 = 120$ outcomes;
  - The first two balls are black and the third is white: $5 \cdot 4 \cdot 6 = 120$ outcomes.

  Assume that "randomly drawn" means that each outcome in the sample space is equally likely to occur.

  Then the desired probability is $\frac{120+120+120}{990} = \frac{4}{11}$.

Example (Cont'd)

- The same problem could also have been solved by regarding the outcome of the experiment as the unordered set of drawn balls.
- Then the number of outcomes in the sample space is $\binom{11}{3} = \frac{11\cdot 10\cdot 9}{1\cdot 2\cdot 3} = 165$.

  Each set of 3 balls corresponds to 3! outcomes when the order of selection is noted. So, if all outcomes are assumed equally likely when the order of selection is noted, then it follows that they remain equally likely when the outcome is taken to be the unordered set of selected balls.

  Hence, using the latter representation of the experiment, we see that the desired probability is

$$\frac{\binom{6}{1}\binom{5}{2}}{\binom{11}{3}} = \frac{6\cdot 10}{165} = \frac{4}{11}.$$

Random Selection Experiments

- When the experiment consists of a random selection of k items from a set of n items, we have the flexibility of:
  - Either letting the outcome of the experiment be the ordered selection of the k items;
  - Or letting it be the unordered set of items selected.
- In the former case we would assume that each new selection is equally likely to be any of the so far unselected items of the set.
- In the latter case we would assume that all $\binom{n}{k}$ possible subsets of k items are equally likely to be the set selected.

Example

- Suppose 5 people are to be randomly selected from a group of 20 individuals consisting of 10 married couples.

  We want to determine $P(N)$, the probability that the 5 chosen are all unrelated (i.e., no two are married to each other).

  If we regard the sample space as the set of 5 people chosen, then there are $\binom{20}{5}$ equally likely outcomes.

  An outcome that does not contain a married couple can be thought of as being the result of a six-stage experiment:
  - In the first stage, 5 of the 10 couples to have a member in the group are chosen;
  - In the next 5 stages, 1 of the 2 members of each of these couples is selected.

  Thus, the number of possible outcomes in which the 5 members selected are unrelated is $\binom{10}{5}2^5$.

  The desired probability is $P(N) = \frac{\binom{10}{5}2^5}{\binom{20}{5}}$.

Example (Cont'd)

- In contrast, we could let the outcome of the experiment be the ordered selection of the 5 individuals.
- In this setting, the number of equally likely outcomes is $20 \cdot 19 \cdot 18 \cdot 17 \cdot 16$.

  Of these, the number of outcomes resulting in a group of 5 unrelated individuals is $20 \cdot 18 \cdot 16 \cdot 14 \cdot 12$.

  This yields the result

$$P(N) = \frac{20 \cdot 18 \cdot 16 \cdot 14 \cdot 12}{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16}.$$

- We can verify that the two answers are identical.

Example

- A committee of 5 is to be selected from a group of 6 men and 9 women.

  If the selection is made randomly, what is the probability that the committee consists of 3 men and 2 women?

  The sample space consists of $\binom{15}{5}$ equally likely outcomes.

  The number of committees consisting of 3 men and 2 women is $\binom{6}{3}\binom{9}{2}$.

  Hence, the desired probability is

$$\frac{\binom{6}{3}\binom{9}{2}}{\binom{15}{5}} = \frac{240}{1001}.$$

Example

- An urn contains n balls, one of which is special.

  k of these balls are withdrawn one at a time, with each selection being equally likely to be any of the balls that remain at the time.

  What is the probability that the special ball is chosen?

  All of the balls are treated in an identical manner.

  Thus, the number of equally likely outcomes is $\binom{n}{k}$.

  The number of outcomes in which the special ball is included is $\binom{1}{1}\binom{n-1}{k-1}$.

  Therefore,

$$P\{\text{special ball is selected}\} = \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{1 \cdot \frac{(n-1)!}{(k-1)!(n-k)!}}{\frac{n!}{k!(n-k)!}} = \frac{k}{n}.$$

Example (Cont'd)

- We could also have obtained this result by letting A_i denote the event that the special ball is the ith ball to be chosen, $i = 1, \ldots, k$.

  Each one of the n balls is equally likely to be the ith ball chosen.

  Thus, $P(A_i) = \frac{1}{n}$.

  The events A_i are mutually exclusive.

  Hence, we have

$$P\{\text{special ball is selected}\} = P(\bigcup_{i=1}^{k} A_i) = \sum_{i=1}^{k} P(A_i) = \frac{k}{n}.$$

Example (Cont'd)

- We argue that $P(A_i) = \frac{1}{n}$ differently.

  The total number of equally likely outcomes of the experiment is
  $$n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$
  The number of outcomes that result in the special ball being the ith chosen is
  $$(n-1)(n-2)\cdots(n-i+1)(1)(n-i)\cdots(n-k+1) = \frac{(n-1)!}{(n-k)!}.$$
  From this reasoning, it follows that
  $$P(A_i) = \frac{\frac{(n-1)!}{(n-k)!}}{\frac{n!}{(n-k)!}} = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

Example

- Suppose that $n+m$ balls, of which n are red and m are blue, are arranged in a linear order in such a way that all $(n+m)!$ possible orderings are equally likely.

  If we record the result of this experiment by listing only the colors of the successive balls, show that all the possible results remain equally likely.

  Consider any one of the $(n+m)!$ possible orderings.

  Any permutation of the red balls among themselves and of the blue balls among themselves does not change the sequence of colors.

  As a result, every ordering of colorings corresponds to $n!m!$ different orderings of the $n+m$ balls.

  So every ordering of the colors has probability $\frac{n!m!}{(n+m)!}$ of occurring.

Example (Cont'd)

- Suppose that there are 2 red balls, numbered r_1, r_2, and 2 blue balls, numbered b_1, b_2.

  Then, of the 4! possible orderings, there will be 2!2! orderings that result in any specified color combination.

  For instance, the following orderings result in the successive balls alternating in color, with a red ball first:

  $$r_1, b_1, r_2, b_2 \quad r_1, b_2, r_2, b_1 \quad r_2, b_1, r_1, b_2 \quad r_2, b_2, r_1, b_1.$$

  Therefore, each of the possible orderings of the colors has probability $\frac{4}{24} = \frac{1}{6}$ of occurring.

Example

- A poker hand consists of 5 cards.

  If the cards have distinct consecutive values and are not all of the same suit, we say that the hand is a **straight**.

  For instance, a hand consisting of the five of spades, six of spades, seven of spades, eight of spades, and nine of hearts is a straight.

  What is the probability that one is dealt a straight?

  We assume that all $\binom{52}{5}$ possible poker hands are equally likely.

  We want to determine the number of outcomes that are straights.

  We first determine the number of possible outcomes for which the poker hand consists of an ace, two, three, four, and five (the suits being irrelevant).

Example (Cont'd)

- The ace can be any 1 of the 4 possible aces.
- Similarly for the two, three, four, and five.

Thus, the number of outcomes leading to exactly one ace, two, three, four, and five is 4^5.

In 4 of these outcomes all the cards will be of the same suit (such a hand is called a **straight flush**).

Thus, the number of hands that make up a straight of the form ace, two, three, four, and five is $4^5 - 4$.

Similarly, there are $4^5 - 4$ hands that make up a straight of the form ten, jack, queen, king, and ace.

Thus, there are $10(4^5 - 4)$ hands that are straights.

It follows that the desired probability is $\frac{10(4^5-4)}{\binom{52}{5}} \approx 0.0039$.

Example

- A 5-card poker hand is said to be a **full house** if it consists of:
  - 3 cards of the same denomination;
  - 2 cards of the same denomination (necessarily different from the first).

  Thus, one kind of full house is three of a kind plus a pair.

  What is the probability that one is dealt a full house?

  We assume that all $\binom{52}{5}$ possible hands are equally likely.

  We determine the number of possible full houses.

  The number of different hands of, say, 2 tens and 3 jacks is $\binom{4}{2}\binom{4}{3}$.

  For the pairs and the 3 cards:
  - There are 13 different choices for the kind of pair;
  - After a pair has been chosen, there are 12 other choices for the denomination of the remaining 3 cards.

  Thus, the probability of a full house is

$$\frac{13 \cdot 12 \cdot \binom{4}{2}\binom{4}{3}}{\binom{52}{5}} \approx 0.0014.$$

Example

- In the game of bridge, the entire deck of 52 cards is dealt out to 4 players. What is the probability that:
  - (a) One of the players receives all 13 spades?
  - (b) Each player receives 1 ace?

(a) Let E_i be the event that hand i has all 13 spades.

Then

$$P(E_i) = \frac{1}{\binom{52}{13}}, \quad i = 1, 2, 3, 4.$$

The events E_i, $i = 1, 2, 3, 4$, are mutually exclusive.

Hence, the probability that one of the hands is dealt all 13 spades is

$$P(\bigcup_{i=1}^{4} E_i) = \sum_{i=1}^{4} P(E_i) = \frac{4}{\binom{52}{13}} \approx 6.3 \times 10^{-12}.$$

Example (Part(b))

(b) We determine the number of outcomes in which each of the distinct players receives exactly 1 ace.

Put aside the aces.

The number of divisions of the other 48 cards when each player is to receive 12 is $\binom{48}{12,12,12,12}$.

The number of ways of dividing the 4 aces so that each player receives 1 is 4!.

So the number of possible outcomes in which each player receives exactly 1 ace is $4!\binom{48}{12,12,12,12}$.

The number of possible hands is $\binom{52}{13,13,13,13}$.

So the desired probability is

$$\frac{4!\binom{48}{12,12,12,12}}{\binom{52}{13,13,13,13}} \approx 0.1055.$$

Example

- If n people are present in a room, what is the probability that no two of them celebrate their birthday on the same day of the year?

  How large need n be so that this probability is less than $\frac{1}{2}$?

  Each person can celebrate his or her birthday on any one of 365 days (ignoring the possibility of a birthday on February 29).

  So the number of possible outcomes is $(365)^n$.

  The number of outcomes in which no two share the same birthday is $(365)(364)(363)\cdots(365-n+1)$.

  Thus, the desired probability is $\frac{(365)(364)(363)\cdots(365-n+1)}{365^n}$.

  This probability is less than $\frac{1}{2}$ when $n \geq 23$.

  That is, if there are 23 or more people in a room, then the probability that at least two of them have the same birthday exceeds $\frac{1}{2}$.

Example

- A deck of 52 playing cards is shuffled, and the cards are turned up one at a time until the first ace appears.

  Is the next card - that is, the card following the first ace - more likely to be the ace of spades or the two of clubs?

  We calculate how many of the 52! possible orderings of the cards have the ace of spades immediately following the first ace.

  Each ordering of the 52 cards can be obtained by:
  - First ordering the 51 cards different from the ace of spades;
  - Then inserting the ace of spades into that ordering.

  Furthermore, for each of the 51! orderings of the other cards, there is only one place where the ace of spades can be placed so that it follows the first ace.

  Therefore, there are 51! orderings that result in the ace of spades following the first ace.

Example (Cont'd)

- It follows that

$$P\{\text{the ace of spades follows the first ace}\} = \frac{51!}{52!} = \frac{1}{52}.$$

By the same argument, the probability that the two of clubs (or any other specified card) follows the first ace is also $\frac{1}{52}$.
In other words, each of the 52 cards of the deck is equally likely to be the one that follows the first ace!

Example

- A football team consists of 20 offensive and 20 defensive players.

  The players are to be paired in groups of 2 for the purpose of determining roommates.

  (a) If the pairing is done at random, what is the probability that there are no offensive-defensive roommate pairs?

  (b) What is the probability that there are $2i$ offensive-defensive roommate pairs, $i = 1, 2, \ldots, 10$?

  The number of ways of dividing the 40 players into 20 ordered pairs of two each is $\binom{40}{2,2,\ldots,2} = \frac{40!}{(2!)^{20}}$.

  I.e., there are $\frac{40!}{2^{20}}$ ways of dividing the players into a first pair, a second pair, and so on.

  Hence, the number of ways of dividing the players into (unordered) pairs of 2 each is $\frac{40!}{2^{20}20!}$.

Example (Part (a))

(a) A division will result in no offensive-defensive pairs if the offensive (and defensive) players are paired among themselves.

The number of ways of pairing each group of 20 among themselves is $\frac{20!}{2^{10}10!}$.

Hence, the number of divisions resulting in no offensive-defensive pairs is $\left(\frac{20!}{2^{10}10!}\right)^2$.

The probability of no offensive-defensive roommate pairs, call it P_0, is given by

$$P_0 = \frac{\left(\frac{20!}{2^{10}10!}\right)^2}{\frac{40!}{2^{20}20!}} = \frac{(20!)^3}{(10!)^2 40!}.$$

Example (Part (b))

(b) Now we determine P_{2i}, the probability that there are $2i$ offensive - defensive pairs.

The number of ways of selecting the $2i$ offensive (or defensive) players out of the 20 is $\binom{20}{2i}$.

Thus, the number of ways of selecting the $2i$ offensive and the $2i$ defensive players who are to be in the pairs is $\binom{20}{2i}^2$.

After the selection, the pairing can be done by:

- Pairing the first offensive player selected with any of the $2i$ defensive players selected;
- Pairing the second offensive player selected with any of the remaining $2i - 1$ defensive players, selected;

  $\vdots$

Hence, the $4i$ selected players can be paired up into $(2i)!$ possible offensive-defensive pairs.

Example (Part (b) Cont'd)

- The remaining $20-2i$ offensive (and defensive) players must be paired among themselves.

  This can be done, for each remaining group, in $\frac{(20-2i)!}{2^{10-i}(10-i)!}$ ways.

  It follows that the number of divisions which lead to $2i$ offensive-defensive pairs is

$$\binom{20}{2i}^2 (2i)! \left[\frac{(20-2i)!}{2^{10-i}(10-i)!}\right]^2.$$

  Hence,

$$P_{2i} = \frac{\binom{20}{2i}^2 (2i)! \left[\frac{(20-2i)!}{2^{10-i}(10-i)!}\right]^2}{\frac{(40)!}{2^{20}(20)!}}, \quad i = 0, 1, \ldots, 10.$$

Example

- The registry of a club shows that:
  - 36 members play tennis;
  - 28 play squash;
  - 18 play badminton;
  - 22 play both tennis and squash;
  - 12 play both tennis and badminton;
  - 9 play both squash and badminton;
  - 4 play all three sports.

  How many members play at least one of three sports?

  Let N denote the number of members of the club.

  Introduce a probability by assuming that a member of the club is randomly selected.

  If, for any subset C of members, we let $P(C)$ denote the probability that the selected member is contained in C, then

$$P(C) = \frac{\text{number of members in } C}{N}.$$

Example (Cont'd)

- Now let:
  - T be the set of members that plays tennis;
  - S be the set that plays squash;
  - B be the set that plays badminton.

  We get, from Inclusion-Exclusion,

$$\begin{aligned} P(T \cup S \cup B) &= P(T) + P(S) + P(B) \\ &\quad - P(TS) - P(TB) - P(SB) \\ &\quad + P(TSB) \\ &= \frac{36 + 28 + 18 - 22 - 12 - 9 + 4}{N} \\ &= \frac{43}{N}. \end{aligned}$$

  Hence, we can conclude that 43 members play at least one of the sports.

Example: The Matching Problem

- Each of N men at a party throws his hat into the center of the room.
  - First, the hats are mixed up;
  - Then, each man randomly selects a hat.

What is the probability that none of the men selects his own hat?

We first calculate the complementary probability of at least one man's selecting his own hat.

Let E_i, $i = 1, \ldots, N$, be the event that the Man i selects his own hat.

Then, the probability that at least one of the men selects his own hat is $P(\bigcup_{i=1}^N E_i)$. By Inclusion-Exclusion,

$$
\begin{aligned}
P(\textstyle\bigcup_{i=1}^N E_i) \quad = \quad & \textstyle\sum_{i=1}^N P(E_i) - \sum_{i_1<i_2} P(E_{i_1} E_{i_2}) + \cdots \\
& + (-1)^{n+1} \textstyle\sum_{i_1<i_2<\cdots<i_n} P(E_{i_1} E_{i_2} \cdots E_{i_n}) \\
& + \cdots + (-1)^{N+1} P(E_1 E_2 \cdots E_N).
\end{aligned}
$$

Example: The Matching Problem (Cont'd)

- Suppose we represent the outcome as a vector of N numbers, where the ith element is the number of the hat drawn by Man i.

  E.g., $(1,2,3,\ldots,N)$ means that each man selects his own hat.

  Then the number of possible outcomes is $N!$.

  The event that each of the n men $i_1, i_2, \ldots, i_n$ selects his own hat is $E_{i_1}E_{i_2}\ldots E_{i_n}$. The number of ways this can occur is

  $$(N-n)(N-n-1)\cdots 3\cdot 2\cdot 1=(N-n)!.$$

  Hence, assuming that all $N!$ possible outcomes are equally likely, we see that

  $$P(E_{i_1}E_{i_2}\cdots E_{i_n})=\frac{(N-n)!}{N!}.$$

Example: The Matching Problem (Conclusion)

- The number of terms in $\sum_{i_1<i_2<\cdots<i_n} P(E_{i_1}E_{i_2}\cdots E_{i_n})$ is $\binom{N}{n}$. Hence

$$\sum_{i_1<i_2<\cdots<i_n} P(E_{i_1}E_{i_2}\cdots E_{i_n}) = \frac{N!(N-n)!}{(N-n)!n!N!} = \frac{1}{n!}.$$

Thus,

$$P(\bigcup_{i=1}^{N} E_i) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{N+1}\frac{1}{N!}.$$

So, the probability that none of the men selects his own hat is

$$1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^N}{N!} \overset{n\to\infty}{\approx} \frac{1}{e}.$$

Example

- Suppose 10 married couples are seated at random at a round table. What is the probability that no wife sits next to her husband?

  Let E_i, $i = 1, 2, \ldots, 10$, be the event that the ith couple sit next to each other.

  Then the desired probability is $1 - P(\bigcup_{i=1}^{10} E_i)$.

  From Inclusion-Exclusion,

$$\begin{array}{rcl} P(\bigcup_1^{10} E_i) & = & \sum_1^{10} P(E_i) \\ & & - \cdots + (-1)^{n+1} \sum_{i_1 < i_2 < \cdots < i_n} P(E_{1_1} E_{i_2} \cdots E_{i_n}) \\ & & + \cdots - P(E_1 E_2 \cdots E_{10}). \end{array}$$

Example (Cont'd)

- We compute $P(E_{i_1}E_{i_2}\cdots E_{i_n})$.

  The number of cyclic arrangements of 20 people is 19!.

  The arrangements that result in a specified set of n men sitting next to their wives can be accomplished by:

  - Sitting the n married couples as if they were single entities: $(20-n-1)!$ ways;
  - Allowing that each of the n married couples can be sat next to each other in one of two possible ways.

  Thus, the number of arrangements that result in a specified set of n men each sitting next to their wives is $2^n(20-n-1)!$.

  Therefore,
  $$P(E_{i_1}E_{i_2}\cdots E_{i_n}) = \frac{2^n(19-n)!}{19!}.$$

Example (Cont'd)

- Using Inclusion-Exclusion, we obtain the probability that at least one married couple sits together:

$$\binom{10}{1}2^1\frac{18!}{19!} - \binom{10}{2}2^2\frac{17!}{19!} + \cdots - \binom{10}{10}2^{10}\frac{9!}{19!}$$
$$\approx 0.6605.$$

So the desired probability is approximately

$$0.3395.$$

Subsection 5

Probability as a Measure of Belief

Personal or Subjective View of Probability

- Thus far we have interpreted the probability of an event of a given experiment as being a measure of how frequently the event will occur when the experiment is continually repeated.
- Consider the statements:
  - "It is 90 percent probable that Shakespeare actually wrote Hamlet";
  - "The probability that Oswald acted alone in assassinating Kennedy is 0.8."
- The simplest and most natural interpretation of such statements is that the probabilities referred to are measures of the individual's degree of belief in the statements that he or she is making.
- This interpretation of probability as being a measure of the degree of one's belief is often referred to as the **personal** or **subjective** view of probability.
- Whether we interpret probability as a measure of belief or as a long-run frequency of occurrence, its mathematical properties remain unchanged.

Example

- Consider a 7-horse race.

  Suppose a bookie feels that:
  - Each of the first 2 horses has a 20 percent chance of winning;
  - Horses 3 and 4 each have a 15 percent chance;
  - The remaining 3 horses have a 10 percent chance each.

  Which of the following two options is better for the bookie to wager at even money?
  1. The winner will be one of the first three horses.
  2. The winner will be one of the horses 1, 5, 6, and 7?

  On the basis of his personal probabilities concerning the outcome, his probability of winning:
  - The first bet is $0.2 + 0.2 + 0.15 = 0.55$;
  - The second bet is $0.2 + 0.1 + 0.1 + 0.1 = 0.5$.

  Hence, the first wager is more attractive.