Lecture 2: Probability, conditional probability, and independence

Theorem 1.2.6.

Let $S = \{s_1, s_2, ...\}$ and $\mathscr{F}$ be all subsets of S.
Let $p_1, p_2, ...$ be nonnegative numbers that sum to 1.
The following defines a probability on $\mathscr{F}$:

$$P(A) = \sum_{i: s_i \in A} p_i, \quad A \in \mathscr{F}$$

(sum over the empty set is defined to be 0).

Theorem 1.2.8-9.

If P is a probability and A and B are events, then

a. $P(A) \leq 1$;
b. $P(A^c) = 1 - P(A)$;
c. $P(B \cap A^c) = P(B) - P(A \cap B)$;
d. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
e. If $A \subset B$, then $P(A) \leq P(B)$.

Birth date problem

There are n balls in N boxes, $n \leq N$.
What is the probability that there are n boxes having exactly one ball?

$$P = \frac{N(N-1)\cdots(N-n+1)}{N^n} = \frac{N!}{N^n(N-n)!}$$

When there are n persons in a party, what is the probability p_n that there are at least two persons having the same birth date?
n persons $=$ n balls
$N = 365$ days

$$P(\text{no one has the same birth date with others}) = \frac{365!}{365^n(365-n)!}$$

$$p_n = 1 - \frac{365!}{365^n(365-n)!}$$

$p_{20} = 0.4058$, $p_{30} = 0.6963$, $p_{50} = 0.9651$, $p_{60} = 0.9922$

Reading assignments: Sections 1.2.3 and 1.2.4

Union of three events

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

$$- P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Example:
$P(A) = P(B) = P(C) = 1/4$, A and B are disjoint,
$P(B \cap C) = P(A \cap C) = 1/6$
$P(A \cup B \cup C) = ?$

$$P(A \cup B \cup C) = 1/4 + 1/4 + 1/4 - 0 - 1/6 - 1/6 + 0 = 7/12$$

General addition formula

For events $A_1, ..., A_n$,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

Theorem 1.2.11.

If P is a probability, A, $A_1, A_2, \ldots$, $C_1, C_2, \ldots$ are events, and $C_i \cap C_j = \emptyset$ for any $i \neq j$ and $\cup_{i=1}^{\infty} C_i = S$, then

a. $P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$;

b. $P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$.

- Property a is useful when $P(A \cap C_i)$ is easy to calculate.
- Property b is called Boole's inequality.

Bonferroni's inequality

For n events $A_1, .., A_n$,

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

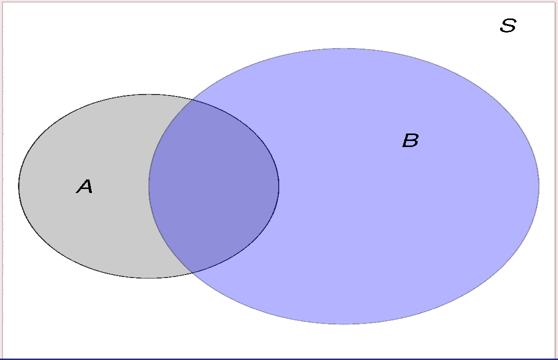
Then

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

which is called Bonferroni's inequality.

Conditional probability

Sometimes we want to know the probability of event A given that another event B has occurred.



Definition 1.3.2.

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

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

- The conditional probability can be viewed as:
  - The sample space is reduced to B.
  - Then A occurs iff $A \cap B$ occurs.
  - To have a probability, we need to divide $P(A \cap B)$ by $P(B)$.
- If $A \cap B = \emptyset$, then $P(A \cap B) = 0$ and $P(A|B) = 0$.
- For convenience, we define $P(A|B) = 0$ when $P(B) = 0$.
- $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$.
- $P(\cdot|B)$ is a probability:
  - $P(\emptyset|B) = 0$, $P(S|B) = P(B|B) = 1$;
  - $P(A^c|B) = 1 - P(A|B)$;
  - $P(A \cup C|B) = P(A|B) + P(C|B) - P(A \cap C|B)$.

Example 1.3.4 (Three prisoners)

Three prisoners, A, B, and C, are on death row.
The warden knows that one of the three prisoners will be pardoned.
Prisoner A tries to get the warden to tell him who had been pardoned, but the warden refuses.
Prisoner A then asks which of B or C will be executed.
The warden thinks for a while, then tells A that B is to be executed.

Prisoner A's thinking:

- Let A, B, and C denote the events that A, B, and C is pardoned, respectively. Then $P(A) = P(B) = P(C) = 1/3$.
- Given that B will be executed, then either A or C will be pardoned; my chance of being pardoned has risen from 0.33 to 0.5.
- Prisoner A's calculation: $P(A|B^c) = \frac{1}{2}$.

Warden's reasoning: Each prisoner has a 1/3 chance of being pardoned; either B or C must be executed, so I have given A no information about whether A will be pardoned.

Prisoner pardoned	Warden tells A	Probability
A	B dies	1/6
A	C dies	1/6
B	C dies	1/3
C	B dies	1/3

$$P(A|\text{warden says B dies}) = \frac{P(\text{warden says B dies} \cap A)}{P(\text{wanden says B dies})} = \frac{1/6}{1/6 + 1/3} = \frac{1}{3}$$

Reason for the difference: $\{\text{warden says B dies}\} \subset B^c$ but the two are not the same; e.g., the warden will not tell A whether A dies or not.

Useful formulas

- For events A_1, A_2,...,

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\cdots P\left(A_n \middle| \bigcap_{i=1}^{n-1} A_i\right)$$

- For disjoint B_1, B_2,..., with $\cup_{i=1}^{\infty} B_i = S$,

$$P(A) = \sum_{i=1}^{\infty} P(B_i)P(A|B_i)$$

In particular, $P(A) = P(B)P(A|B) + P(B^c)P(A|B^c)$.

Example

There are 90 red balls and 10 white balls. We select randomly one ball at a time without replacement.

What is the probability that the 1st white ball appears at the 3rd draw?

Define $A_i = \{\text{the } i\text{th selection is a white ball}\}$

$B_i = \{\text{the } i\text{th selection is a red ball}\}$

$$P(B_1 \cap B_2 \cap A_3) = P(B_1)P(B_2|B_1)P(A_3|B_1 \cap B_2) = \frac{90}{100} \times \frac{89}{99} \times \frac{10}{98}$$

Sensitive questions

People may not give you a true answer if he/she is asked “did you cheat in the last exam"?
If we want to know the probability that a person cheats in a given exam, we may use the following procedure.
We ask a person to flip a fair coin.
If the result is “heads", then the question is “did you cheat"?
If the result is “tails", then the question is “is your birthday afetr July 1"?
No one, except the person being asked, knows which question is asked.
The person will give the truthful answer.

$A =$ the event that the person answers “yes".
$B =$ the event that the person is asked the 1st question.

$$P(A) = P(B)P(A|B) + P(B^c)P(A|B^c) = 0.5 \times P(A|B) + 0.5 \times 0.5$$

$$P(A|B) = 2P(A) - 0.5$$

If we estimate $P(A)$ by asking 10,000 people, then we can estimate $P(A|B)$.

Theorem 1.3.5. Bayes formula

Let B, $A_1, A_2, \ldots$ be events with $A_i \cap A_j = \emptyset$, $i \neq j$, and $\cup_{i=1}^{\infty} A_i = S$. Then,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}.$$

In particular,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Example

$A =$ the event that a person has a type of disease, $P(A) = 0.0004$.
$B =$ a medical test indicates that a tested person has the disease,
$P(B|A) = 0.99$, $P(B|A^c) = 0.05$.
$P(A|B) = ?$
By Bayes formula,

$$P(A|B) = \frac{0.0004 \times 0.99}{0.0004 \times 0.99 + 0.9996 \times 0.05} = 0.007808$$

Definition 1.3.7.

Two events A and B are (statistically) independent iff

$$P(A \cap B) = P(A)P(B).$$

- This definition is still good if $P(A) = 0$ or $P(B) = 0$.
- A and B are independent iff $P(A|B) = P(A)$ (or $P(B|A) = P(B)$). That is, the probability of the occurrence of A is not affected by the occurrence of B.
- If A and B are independent, then the following pairs are also independent: A and B^c, A^c and B, and A^c and B^c; e.g.,

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)P(B^c)$$

Example 1.3.11 (Letters)

S consists of the 3! permutations of the letters a, b, and c, along with the three triplets of each letter:

$$S = \left\{ \begin{array}{ccc} aaa & bbb & ccc \\ abc & bca & cba \\ acb & bac & cab \end{array} \right\}$$

Suppose that each element of S has probability 1/9 to occur.

Define $A_i = \{i\text{th place in the triple is occupied by a}\}$, $i = 1, 2, 3$.

By looking at S, we obtain that

$$P(A_1) = P(A_2) = P(A_3) = 1/3$$

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = 1/9$$

Thus,

$$P(A_i \cap A_j) = P(A_i)P(A_j) \quad i \neq j$$

and A_i and A_j are independent for any pair of (i, j), $i \neq j$.

It is easy to find events that are not independent, e.g., $C = \{abc, bca\}$,

$$P(C) = 2/9, \qquad P(A_1 \cap C) = 1/9 \neq P(C)P(A_1) = 2/27$$

If A and B has some relationship, usually they are not independent.

- If $A \subset B$, then $P(A) = P(A \cap B) = P(A)P(B)$ iff $P(A) = 0$ or $P(B) = 1$.
- If $A \cap B = \emptyset$, then $P(A \cap B) = P(A)P(B) = 0$ iff $P(A) = 0$ or $P(B) = 0$.

Definition 1.3.12.

A collection of events $A_1, ..., A_n$ are (mutually) independent iff for any subcollection $A_{i_1}, ..., A_{i_k}$,

$$P\left(A_{i_1} \cap \cdots \cap A_{i_k}\right) = P(A_{i_1}) \cdots P(A_{i_k})$$

- If we know $P(A_i)$, $i = 1, ..., n$, we may not know $P(\cup_{i=1}^n A_i)$. But if $A_1, ..., A_n$ are independent, then

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

$$+ \sum_{i<j<k} P(A_i)P(A_j)P(A_k) + \cdots + (-1)^{n-1} P(A_1)P(A_2)\cdots P(A_n)$$

- If events A, B, C are independent, then so are $A \cup B$ and C, $A \cap B$ and C, and $A - B$ and C, e.g.,

$$\begin{aligned} P((A \cup B) \cap C) &= P((A \cap C) \cup (B \cap C)) \\ &= P(A \cap C) + P(B \cap C) - P(A \cap C \cap B) \\ &= P(A)P(C) + P(B)P(C) - P(A)P(C)P(B) \\ &= P(C)[P(A) + P(B) - P(A)P(B)] = P(C)P(A \cup B) \end{aligned}$$

Pairwise independence

Events $A_1, ..., A_n$ are pairwise independent iff A_i and A_j are independent for any pair (i, j).

Mutual independence is stronger than pairwise independence.

Example 1.3.11 (continued)

A_1, A_2, A_3 are pairwise independent.

They are not mutually independent, because

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{9} \neq \frac{1}{27} = P(A_1)P(A_2)P(A_3)$$

Example 1.3.10 (Tossing two dice)

An experiment consists of tossing two dice.

The sample space is

$$S = \left\{ \begin{array}{ccc} (1,1) & \cdots & (1,6) \\ (2,1) & \cdots & (2,6) \\ \cdots & \cdots & \cdots \\ (6,1) & \cdots & (6,6) \end{array} \right\}$$

Consider

$$\begin{aligned} A &= \{\text{the same digit appear twice}\} \qquad P(A) = 1/6 \\ B &= \{\text{the sum is between 7 and 10}\} \qquad P(B) = 1/2 \\ C &= \{\text{the sum is 2, or 7, or 8}\} \qquad P(C) = 1/3 \end{aligned}$$

Then

$$P(A \cap B \cap C) = P(\{(4,4)\}) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{2} \times \frac{1}{3} = P(A)P(B)P(C)$$

However, A, B, and C are not independent, because

$$P(B \cap C) = P(\{\text{the sum is 7 or 8}\}) = \frac{11}{36} \neq \frac{1}{6} = P(B)P(C)$$

Independent experiments

- Two experiments are independent iff any event in experiment 1 and any event in experiment 2 are independent.
- n experiments are independent iff $A_1, ..., A_n$ are independent, where A_j is any event from experiment j, $j = 1, ..., n$.
- Independence of experiments is an important concept in statistics.

Conditional independence

Events A and B are conditionally independent given event C iff

$$P(A \cap B|C) = P(A|C)P(B|C)$$

Conditional independence is an important concept in statistics.

Independence does not imply conditional independence

In Example 1.3.11,

$$A_i = \{i\text{th place in the triple is occupied by a}\}, \qquad i = 1, 2, 3.$$

A_1 and A_2 are independent, but they are not conditionally independent given A_3, because

$$P(A_1 \cap A_2|A_3) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_3)} = \frac{1/9}{1/3} = \frac{1}{3}$$

$$P(A_1|A_3)P(A_2|A_3) = \frac{P(A_1 \cap A_3)}{P(A_3)} \frac{P(A_2 \cap A_3)}{P(A_3)} = P(A_1)P(A_2) = \frac{1}{9}$$

Conditional independence does not imply independence

$$S = \left\{ \begin{array}{cccc} aaa & bbb & ccc & abd \\ abc & bca & cba & acd \\ acb & bac & cab & bcd \end{array} \right\}$$

$$A_i = \{i\text{th place in the triple is occupied by a}\}, \qquad i = 1, 2, 3.$$

$$B = \{\text{d does not appear}\}$$

A_1 and A_2 are conditionally independent given B, because

$$P(A_1 \cap A_2 | B) = \frac{1}{9} = \frac{1}{3} \times \frac{1}{3} = P(A_1|B)P(A_2|B)$$

However, A_1 and A_2 are not independent, because

$$P(A_1 \cap A_2) = \frac{1}{9} \neq \frac{5}{12} \times \frac{3}{12} = P(A_1)P(A_2)$$

Mutual independence implies conditional independence

If A, B, C are mutually independent, then A and B are conditionally independent given C, because

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