Independence and Counting

Yi-Fen Liu

Department of IECS, FCU

References:
- Dimitri P. Bertsekas and John N. Tsitsiklis, *Introduction to Probability*, Sections 1.5-1.6
- Slides are credited from Prof. Berlin Chen, NTNU.

Independence (1/2)

- Recall that conditional probability $\mathbf{P}(A|B)$ captures the partial information that event B provides about event A

- A special case arises when the occurrence of event B provides no such information and does not alter the probability A has occurred

$$\mathbf{P}(A|B) = \mathbf{P}(A)$$

  - A is independent of B (B also is independent of A)

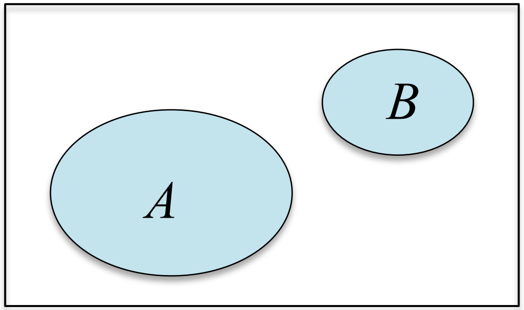
$$\Rightarrow \mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \mathbf{P}(A)$$

$$\Rightarrow \mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$$

Independence (2/2)

- A and B are independent => A and B are disjoint?
  - No! Why?
    - A and B are disjoint then $\mathbf{P}(A \cap B) = 0$
    - However, if $\mathbf{P}(A) > 0$ and $\mathbf{P}(B) > 0$

$$\Rightarrow \mathbf{P}(A \cap B) \neq \mathbf{P}(A)\mathbf{P}(B)$$



- Two disjoint events A and B with $\mathbf{P}(A) > 0$ and $\mathbf{P}(B) > 0$ are never independent

Independence: An Example (1/3)

- **Example 1.19.** Consider an experiment involving two successive rolls of a 4-sided die in which all 16 possible outcomes are equally likely and have probability 1/16

Using Discrete Uniform Probability Law here

(a) Are the events,

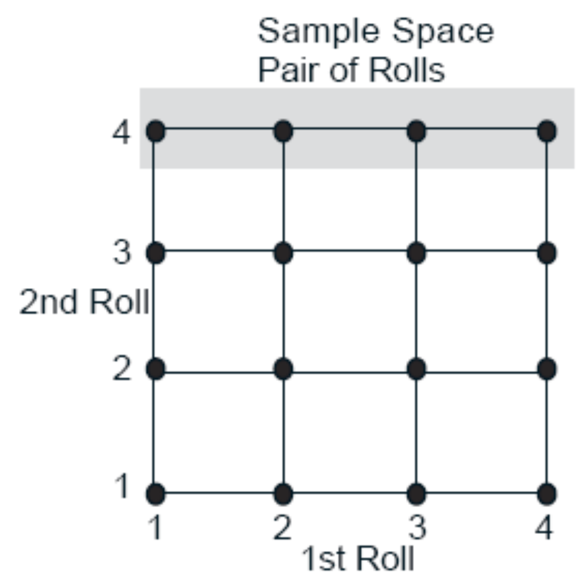
A_i = {1st roll results in i },
B_j = {2nd roll results in j }, independent

$$\mathbf{P}\left(A_i \cap B_j\right) = \frac{1}{16}$$

$$\mathbf{P}\left(A_i\right) = \frac{4}{16},\ \mathbf{P}\left(B_j\right) = \frac{4}{16}$$

$$\Rightarrow \mathbf{P}\left(A_i \cap B_j\right) = \mathbf{P}\left(A_i\right)\mathbf{P}\left(B_j\right)$$

$$\Rightarrow\ A_i \text{ and } B_j \text{ are independent!}$$



Independence: An Example (2/3)

(b) Are the events,

A = {1st roll results is a 1},

B = {sum of the two rolls is a 5 }, independent?

$$\mathbf{P}(A) = \frac{4}{16} \quad \text{(the results of two rolls are } (1,1),(1,2),(1,3),(1,4))$$

$$\mathbf{P}(B) = \frac{4}{16} \quad \text{(the results of two rolls are } (1,4),(2,3),(3,2),(4,1))$$

$$\mathbf{P}(A \cap B) = \frac{1}{16} \quad \text{(the only one result of two rolls is } (1,4))$$

$$\Rightarrow \mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$$

$$\Rightarrow \ A \text{ and } B \text{ are independent!}$$

Independence: An Example (3/3)

(c) Are the events,

A = {maximum of the two rolls is 2},
B = {minimum of the two rolls is 2 }, independent?

$$\mathbf{P}(A) = \frac{3}{16} \quad \text{(the results of two rolls are (1,2),(2,1),(2,2))}$$

$$\mathbf{P}(B) = \frac{5}{16} \quad \text{(the results of two rolls are (2,2),(2,3),(2,4),(3,2),(4,2))}$$

$$\mathbf{P}(A \cap B) = \frac{1}{16} \quad \text{(the only one result of two rolls is (2,2))}$$

$$\Rightarrow \mathbf{P}(A \cap B) \neq \mathbf{P}(A)\mathbf{P}(B)$$

$$\Rightarrow A \text{ and } B \text{ are dependent!}$$

Conditional Independence (1/2)

- Given an event C, the event A and B called conditionally independent if

1

$$\mathbf{P}(A \cap B \mid C) = \boxed{\mathbf{P}(A \mid C)\mathbf{P}(B \mid C)}$$

  - We also know that

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

multiplication rule

2

$$= \frac{\cancel{\mathbf{P}(C)}\boxed{\mathbf{P}(B \mid C)\mathbf{P}(A \mid B \cap C)}}{\cancel{\mathbf{P}(C)}}$$

  - If $\mathbf{P}(B \mid C) > 0$, we have an alternative way to express conditional independence

3

$$\mathbf{P}(A \mid B \cap C) = \mathbf{P}(A \mid C)$$

Conditional Independence (2/2)

- Notice that independence of two events A and B with respect to the unconditionally probability law does not imply conditional independence, and vice versa

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B) \not\Leftrightarrow \mathbf{P}(A \cap B | C) = \mathbf{P}(A|C)\mathbf{P}(B|C)$$

- If A and B are independent, the same holds for
  (i) A and B^c
  (ii) A^c and B^c
  – How can we verify it? (See Problem 43)

Conditional Independence: Examples (1/2)

- **Example 1.20.** Consider two independent fair coin tosses, in which all four possible outcomes are equally likely. Let

Using Discrete Uniform Probability Law here

$H_1 = \{$ 1st toss is a head $\}$, $\quad (H, T), (H, H)$

$H_2 = \{$ 2nd toss is a head $\}$, $\quad (T, H), (H, H)$

$D = \{$ the two tosses have different results $\}$. $\quad (T, H), (H, T)$

$$\mathbf{P}(H_1 | D) = \frac{1}{2} \qquad (H, T)$$

$$\mathbf{P}(H_2 | D) = \frac{1}{2} \qquad (T, H)$$

$$\Rightarrow \mathbf{P}(H_1 \cap H_2 | D) = \frac{\mathbf{P}(H_1 \cap H_2 \cap D)}{\mathbf{P}(D)} = 0 \neq \mathbf{P}(H_1 | D)\mathbf{P}(H_2 | D)$$

$\Rightarrow$ H_1 and H_2 are conditionally dependent!

Conditional Independence: Examples (2/2)

- **Example 1.21.** There are two coins, a blue and a red one
  - We choose one of the two at random, each being chosen with probability 1/2, and proceed with two independent tosses
  - The coins are biased: with the blue coin, the probability of heads in any given toss is 0.99, whereas for the red coin it is 0.01
  - Let B be the event that the blue coin was selected. Let also H_i be the event that the *i*-th toss resulted in heads

conditional case: $$\mathbf{P}(H_1 \cap H_2 | B) = \mathbf{P}(H_1|B)\mathbf{P}(H_2|B)$$

Given the choice of a coin, the events H_1 and H_2 are independent

unconditional case: $$\mathbf{P}(H_1 \cap H_2) \overset{?}{=} \mathbf{P}(H_1)\mathbf{P}(H_2)$$

$$\mathbf{P}(H_1) = \mathbf{P}(B)\mathbf{P}(H_1|B) + \mathbf{P}(B^C)\mathbf{P}(H_1|B^C) = \frac{1}{2}\cdot 0.99 + \frac{1}{2}\cdot 0.01 = \frac{1}{2}$$

$$\mathbf{P}(H_2) = \mathbf{P}(B)\mathbf{P}(H_2|B) + \mathbf{P}(B^C)\mathbf{P}(H_2|B^C) = \frac{1}{2}\cdot 0.99 + \frac{1}{2}\cdot 0.01 = \frac{1}{2}$$

$$\mathbf{P}(H_1 \cap H_2) = \mathbf{P}(B)\mathbf{P}(H_1 \cap H_2|B) + \mathbf{P}(B^C)\mathbf{P}(H_1 \cap H_2|B^C)$$

$$= \frac{1}{2}\cdot 0.99 \cdot 0.99 + \frac{1}{2}\cdot 0.01 \cdot 0.01 \neq \frac{1}{4}$$

Independence of a Collection of Events (1/2)

- We say that the events $A_1, A_2, \ldots, A_n$ are independent if

$$\mathbf{P}\left(\bigcap_{i\in S} A_i\right) = \prod_{i\in S}\left(A_i\right), \text{ for every subset } S \text{ of } \{1,2,\ldots,n\}$$

- For example, the independence of three events A_1, A_2, A_3 amounts to satisfying the four conditions

$$\mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2)$$
$$\mathbf{P}(A_1 \cap A_3) = \mathbf{P}(A_1)\mathbf{P}(A_3)$$
$$\mathbf{P}(A_2 \cap A_3) = \mathbf{P}(A_2)\mathbf{P}(A_3)$$
$$\mathbf{P}(A_1 \cap A_2 \cap A_3) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3)$$

2^n-n-1

Independence of a Collection of Events (2/2)

- Independence means that the occurrence or nonoccurrence of any number of the events from that collection carries no information on the remaining events or their complements

$$\mathbf{P}\left(A_1 \cup A_2 \middle| A_3 \cap A_4\right) = \mathbf{P}\left(A_1 \cup A_2\right)$$

$$\mathbf{P}\left(A_1 \cup A_2^c \middle| A_3^c \cap A_4\right) = \mathbf{P}\left(A_1 \cup A_2^c\right)$$

(see the end-of-chapter problems)

Independence of a Collection Events: Examples (1/4)

- **Example 1.22. Pairwise independence does not imply independence**
  - Consider two independent fair coin tosses, and the following events:

$H_1 = \{\text{ 1st toss is a head }\}$, $(H, T), (H, H)$

$H_2 = \{\text{ 2nd toss is a head }\}$, $(T, H), (H, H)$

$D = \{\text{ the two tosses have different results }\}$. $(T, H), (H, T)$

$$\mathbf{P}(H_1 \cap H_2) = \mathbf{P}(H_1)\mathbf{P}(H_2)$$

$$\mathbf{P}(H_1 \cap D) = \mathbf{P}(H_1)\mathbf{P}(D)$$

$$\mathbf{P}(H_2 \cap D) = \mathbf{P}(H_2)\mathbf{P}(D)$$

However, $\mathbf{P}(H_1 \cap H_2 \cap D) = 0 \neq \mathbf{P}(H_1)\mathbf{P}(H_2)\mathbf{P}(D)$

Independence of a Collection Events: Examples (2/4)

- **Example 1.23. The equality**

$$\mathbf{P}(A_1 \cap A_2 \cap A_3) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3)$$

is not enough for independence.

  – Consider two independent fair coin tosses, and the following events:

$$A = \{\text{ 1st roll is 1, 2, or 3}\},$$
$$B = \{\text{ 1st roll is 3, 4, or 5}\},$$
$$C = \{\text{ the sum of the two rolls is 9 }\}.$$

$$\mathbf{P}(A \cap B \cap C) = \frac{1}{36} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{4}{36} = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)$$

However,

$$\mathbf{P}(A \cap B) = \frac{1}{6} \neq \frac{1}{2} \cdot \frac{1}{2} = \mathbf{P}(A)\mathbf{P}(B)$$

$$\mathbf{P}(A \cap C) = \frac{1}{36} \neq \frac{1}{2} \cdot \frac{4}{36} = \mathbf{P}(A)\mathbf{P}(C)$$

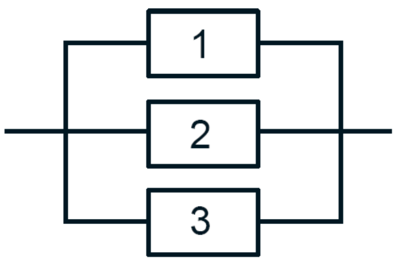
$$\mathbf{P}(B \cap C) = \frac{1}{12} \neq \frac{1}{2} \cdot \frac{4}{36} = \mathbf{P}(B)\mathbf{P}(C)$$

Independence of a Collection Events: Examples (3/4)

- **Example 1.24. Network connectivity.** A computer network connects two nodes *A* and *B* through intermediate nodes *C, D, E, F* (See next slide)
  - For every pair of directly connected nodes, say *i* and *j*, there is a given probability p_{ij} that the link from *i* to *j* is up. We assume that link failures are independent of each other
  - What is the probability that there is a path connecting *A* and *B* in which all links are up?

Series Connection

$$\mathbf{P}(\text{series subsystem succeeds}) = p_1 p_2 \cdots p_n$$

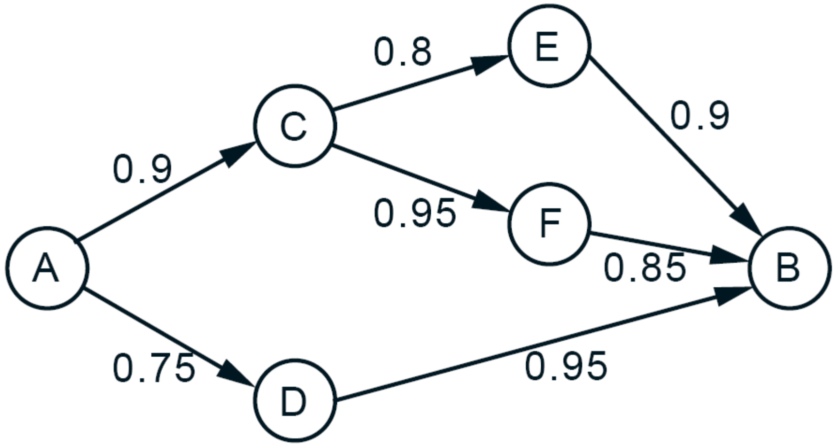


Parallel Connection

$$\begin{aligned}\mathbf{P}(\text{parallel subsystem succeeds}) &= 1 - \mathbf{P}(\text{parallel subsystem fails}) \\ &= 1 - (1-p_1)(1-p_2)\cdots(1-p_n)\end{aligned}$$

Independence of a Collection Events: Examples (4/4)

- **Example 1.24.** (cont.)



$$\mathbf{P}(\mathrm{C} \to \mathrm{B}) = 1 - \big(1 - \mathbf{P}(\mathrm{C} \to \mathrm{E} \to \mathrm{B})\big)\big(1 - \mathbf{P}(\mathrm{C} \to \mathrm{F} \to \mathrm{B})\big)$$

$$= 1 - (1 - 0.8 \cdot 0.9)(1 - 0.95 \cdot 85)$$

$$= 0.946$$

$$\mathbf{P}(\mathrm{A} \to \mathrm{C} \to \mathrm{B}) = \mathbf{P}(\mathrm{A} \to \mathrm{C})\mathbf{P}(\mathrm{C} \to \mathrm{B}) = 0.9 \cdot 0.946 = 0.851$$

$$\mathbf{P}(\mathrm{A} \to \mathrm{D} \to \mathrm{B}) = \mathbf{P}(\mathrm{A} \to \mathrm{D})\mathbf{P}(\mathrm{D} \to \mathrm{B}) = 0.75 \cdot 0.95 = 0.712$$

$$\mathbf{P}(\mathrm{A} \to B) = 1 - \big(1 - \mathbf{P}(\mathrm{A} \to \mathrm{C} \to \mathrm{B})\big)\big(1 - \mathbf{P}(\mathrm{A} \to \mathrm{D} \to \mathrm{B})\big)$$

$$= 1 - (1 - 0.851) \cdot (1 - 0.712) = 0.957$$

Recall: Counting in Probability Calculation

- Two applications of the discrete uniform probability law
  - when the sample space Ω has a finite number of equally likely outcomes, the probability of any event A is given by

$$\mathbf{P}(A) = \frac{\text{number of elements of A}}{\text{number of elements of } \Omega}$$

  - when we want to calculate the probability of an event A with a finite number of equally likely outcomes, each of which has an already known probability p. Then the probability of A is given by

$$\mathbf{P}(A) = p \cdot (\text{number of elements of } A)$$

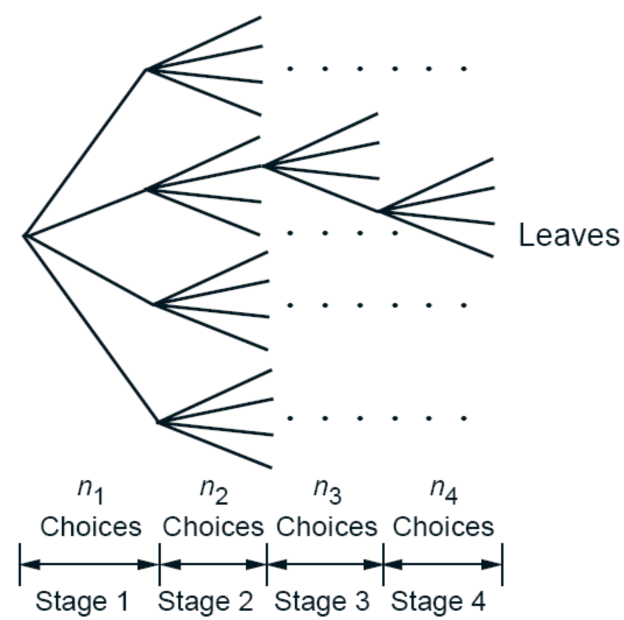
    - E.g., the calculation of *k* heads in *n* coin tosses

The Counting Principle

- Consider a process that consists of *r* stages. Suppose that
  a) There are n_1 possible results for the first stage
  b) For every possible result of the first stage, there are n_2 possible results at the second stage
  c) More generally, for all possible results of the first *i* -1 stages, there are n_i possible results at the *i*-th stage

  Then, the total number of possible results of the *r*-stage process is

$$n_1 \; n_2 \cdots \; n_r$$



Common Types of Counting

- Permutations of *n* objects

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

- *k*-permutations of *n* objects

$$\frac{n!}{(n-k)!}$$

- Combinations of *k* out of *n* objects

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- Partitions of n objects into *r* groups with the *i*-th group having n_i objects

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}$$

Summary of Chapter 1 (1/2)

- A probability problem can usually be broken down into a few basic steps:
  1. The description of the sample space, i.e., the set of possible outcomes of a given experiment
  2. The (possibly indirect) specification of the probability law (the probability of each event)
  3. The calculation of probabilities and conditional probabilities of various events of interest

Summary of Chapter 1 (2/2)

- Three common methods for calculating probabilities
  - The counting method: if the number of outcome is finite and all outcome are equally likely

$$\mathbf{P}(A) = \frac{\text{number of elements of A}}{\text{number of elements of } \Omega}$$

  - The sequential method: the use of the multiplication (chain) rule

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

  - The divide-and-conquer method: the probability of an event is obtained based on a set of conditional probabilities

$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(A_1 \cap B) + \cdots + \mathbf{P}(A_n \cap B) \\ &= \mathbf{P}(A_1)\mathbf{P}(B | A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B | A_n)\end{aligned}$$

    - A_1, … A_n are disjoints that form a partition of sample space

Recitation

- SECTION 1.5 Independence
  - Problems 37, 38, 39, 40