STA 130A: Mathematical Statistics: Brief Course

Lecture 4: Independence and Counting Principle

Dogyoon Song

Fall 2026, UC Davis

Announcements

Homework 1 is due tomorrow (Thu, Oct 1, 11:59 PM)

- Please submit on time and follow the submission instructions
- Instructor office hours: Wed, 2–3 PM (MSB 1143)
- TA office hours: Thu, 9–11 AM (MSB 1117)
- Feel free to post questions on Piazza (the earlier, the better)

Agenda

Last time:

- Conditional probability
- Bayes' rule

Today:

- Independence
- Modeling with independent components and repeated trials
- Counting principle and simple applications

Independence

Sometimes knowing that B occurred does not change the probability of A

- When $P(B) > 0$, this means $P(A \mid B) = P(A)$

Definition (Independence)

The event A is **independent** of B if

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

- Independence is symmetric: A is independent of $B \implies B$ is independent of A
  - We can unambiguously say "A and B are independent events"
- Events from distinct, non-interacting experiments are often modeled as independent
- Independence is **not the same as disjointness**
  - If $A \cap B = \emptyset$ and $P(A), P(B) > 0$, then A and B cannot be independent (e.g., A and A^c)
- A and B are independent $\implies A^c$ and B are independent
  - Because $P(A^c \cap B) = P(B) - P(A \cap B) = P(B) - P(A)P(B) = P(A^c)P(B)$

Checking independence: Two rolls of a fair die (1/3)

Example (2 successive die rolls)

Model two independent rolls of a fair die; all $6 \times 6 = 36$ outcomes are equally likely

Question: Are the following two events independent?

$$A_1 = \{\text{1st roll results in } 6\}, \qquad B_1 = \{\text{2nd roll results in } 6\}$$

Answer: Yes, because

$$P(A_1 \cap B_1) = P\big(\text{the outcome of the two rolls is } (6,6)\big) = \frac{1}{36}$$

$$P(A_1) = \frac{P\big(\{(6,1),(6,2),\ldots,(6,6)\}\big)}{P\big(\{(1,1),(1,2),\ldots,(6,6)\}\big)} = \frac{6}{36} = \frac{1}{6}$$

$$P(B_1) = \frac{P\big(\{(1,6),(2,6),\ldots,(6,6)\}\big)}{P\big(\{(1,1),(1,2),\ldots,(6,6)\}\big)} = \frac{6}{36} = \frac{1}{6}$$

Observe $P(A_1 \cap B_1) = P(A_1)\,P(B_1)$, which verifies the independence of A_1 and B_1)

Checking independence: Two rolls of a fair die (2/3)

Example (2 successive die rolls)

Model two independent rolls of a fair die; all $6 \times 6 = 36$ outcomes are equally likely

Question: Are the following two events independent?

$$A_2 = \{\text{1st roll is 6}\}, \qquad B_2 = \{\text{the sum of the two rolls is 7}\}$$

Answer: Yes, because

$$P(A_2 \cap B_2) = P\big(\{(6,1)\}\big) = \frac{1}{36}$$

$$P(A_2) = \frac{P\big(\{(6,1),(6,2),\dots,(6,6)\}\big)}{P\big(\{(1,1),(1,2),\dots,(6,6)\}\big)} = \frac{6}{36} = \frac{1}{6}$$

$$P(B_2) = \frac{P\big(\{(6,1),(5,2),\dots,(1,6)\}\big)}{P\big(\{(1,1),(1,2),\dots,(6,6)\}\big)} = \frac{6}{36} = \frac{1}{6}$$

Observe $P(A_2 \cap B_2) = P(A_2)\,P(B_2)$, which verifies the independence of A_2 and B_2)

Follow-up: What about $B_2' = \{\text{the sum is 5}\}$? No, becauase $P(A_2 \cap B_2') = 0$ but $P(B_2') > 0$

Checking independence: Two rolls of a fair die (3/3)

Example (2 successive die rolls)

Model two independent rolls of a fair die; all $6 \times 6 = 36$ outcomes are equally likely

Question: Are the following two events independent?

$$A_3 = \{\text{maximum of the two rolls is } 3\}, \qquad B_3 = \{\text{minimum of the two rolls is } 3\}$$

Answer: No, because

$$P(A_3 \cap B_3) = P\big(\{(3,3)\}\big) = \frac{1}{36}$$

$$P(A_3) = \frac{P\big(\{(1,3),(2,3),(3,3),(3,2),(3,1)\}\big)}{P\big(\{(1,1),(1,2),\dots,(6,6)\}\big)} = \frac{5}{36}$$

$$P(B_3) = \frac{P\big(\{(3,6),(3,5),(3,4),(3,3),(4,3),(5,3),(6,3)\}\big)}{P\big(\{(1,1),(1,2),\dots,(6,6)\}\big)} = \frac{7}{36}$$

Observe $P(A_3 \cap B_3) = \frac{1}{36} \neq \frac{35}{36^2} = P(A_3)\,P(B_3)$, thus A_3 and B_3) are not independent

Conditional independence

Recall: for any event C with $P(C) > 0$, the map $A \mapsto P(A \mid C)$ is itself a probability law

- So we can ask whether A and B are independent under this conditional law

Definition (Conditional independence)

The events A and B are **conditionally independent** given an event C if

$$P(A \cap B \mid C) = P(A \mid C)\, P(B \mid C)$$

If $P(B \cap C) > 0$, this is equivalent to $P(A \mid B \cap C) = P(A \mid C)$ because

$$\begin{aligned} P(A \cap B \mid C) &= \frac{P(A \cap B \cap C)}{P(C)} && \text{by definition} \\ &= \frac{P(C)\, P(B \mid C)\, P(A \mid B \cap C)}{P(C)} && \text{by multiplicative rule} \\ &= P(B \mid C)\, P(A \mid B \cap C) \end{aligned}$$

- Once C is known to have occurred, knowing B gives no additional information about A
- Independence *can change when we condition on additional information*

Conditioning can change independence

Example (2 fair coin tosses)

Consider two independent fair coin tosses, and let

$$A = \{\text{1st toss is a head}\}$$
$$B = \{\text{2nd toss is a head}\}$$

- These events are independent. Now condition on

$$C = \{\text{the two tosses have different results}\}$$

- The events A and B are not conditionally independent given C because

$$P(A \mid C) = \frac{1}{2}, \qquad P(B \mid C) = \frac{1}{2}, \qquad P(A \cap B \mid C) = 0$$

Why? Given C, the only possible outcomes are HT and TH, so if one toss is Head, the other must be Tail

The independence of A and B does not imply conditional independence

(Optional) Conditional independence $\not\Rightarrow$ Independence

Example (2 different biased coin tosses)

Suppose there are two coins, fair and biased (say, landing on head with prob 0.9), and we choose one of the two at random with probability $1/2$ each, and toss the coin twice:

$$A_i = \{\text{the i-th toss is a head}\}$$
$$B = \{\text{the coin is biased}\}$$

- The events A_1 and A_2 are independent, given the choice of a coin (=conditioned on B)
- However, without conditioning on B, A_1 and A_2 are not independent:

$$P(A_i) = P(B)\,P(A_i \mid B) + P(B^c)\,P(A_i \mid B^c) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{9}{10} = \frac{7}{10}, \qquad i \in \{1,2\},$$
$$P(A_1 \cap A_2) = P(B)\,P(A_1 \cap A_2 \mid B) + P(B^c)\,P(A_1 \cap A_2 \mid B^c) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{9}{10}\cdot\frac{9}{10} = \frac{53}{100}$$

Why? The hidden variable "which coin was chosen" creates dependence between the two tosses

The conditional independence of A and B does not imply unconditional independence

Pop-up quiz

Suppose two independent fair coin tosses. Let

$$A = \{\text{1st toss is H}\}, \qquad B = \{\text{2nd toss is H}\}, \qquad C = \{\text{the two tosses are different}\}.$$

Which statement is correct?

a) A and B are not independent, but are conditionally independent given C.

b) A and B are independent, and are conditionally independent given C.

c) A and B are independent, but are not conditionally independent given C.

d) A and B are neither independent nor conditionally independent given C.

Answer: C.

$$P(A) = P(B) = \frac{1}{2}, \qquad P(A \cap B) = \frac{1}{4},$$

so A and B are independent. But given C, the only outcomes are HT and TH, so

$$P(A \mid C) = P(B \mid C) = \frac{1}{2}, \qquad P(A \cap B \mid C) = 0.$$

Follow-up: Would the answer change if we condition on C^c instead of C?

Independence of several events

We can extend the definition of the events for more than two events

Definition (Independence (mutual independence))

The events $A_1, A_2, \ldots, A_n$ are **independent** if

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

For the case of three events, A_1, A_2, A_3 are independent if and only if

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1)\,P(A_2), \\ P(A_1 \cap A_3) &= P(A_1)\,P(A_3), \\ P(A_2 \cap A_3) &= P(A_2)\,P(A_3), \\ P(A_1 \cap A_2 \cap A_3) &= P(A_1)\,P(A_2)\,P(A_3) \end{aligned}$$

The first three (pairwise independence) do not imply the fourth, and vice versa

Pairwise independence is not enough for mutual independence

Example (2 fair coin tosses)

Consider two independent fair coin tosses, and let

$$\begin{aligned} A_1 &= \{\text{1st toss is a head}\} \\ A_2 &= \{\text{2nd toss is a head}\} \\ B &= \{\text{the two tosses have different results}\} \end{aligned}$$

- These three events are pairwise independent:
  - A_1 and A_2 are independent by definition
  - A_1 and B are independent because

$$P(B \mid A_1) = \frac{P(A_1 \cap B)}{P(A_1)} = \frac{P\big(\{(H,T)\}\big)}{P\big(\{(H,T),(H,H)\}\big)} = \frac{1}{2} = P(B)$$

- However,

$$P(A_1 \cap A_2 \cap B) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(A_1)\, P(A_2)\, P(B)$$

  and therefore, these three events are not independent

(Optional) Three-way factorization $\not\Longrightarrow$ mutual independence

Example (2 fair die rolls)

Consider two independent rolls of a fair six-sided die, and let

$$\begin{aligned} A &= \{\text{1st roll is 1, 2, or 3}\} \\ B &= \{\text{1st roll is 3, 4, or 5}\} \\ C &= \{\text{the sum of the two rolls is 9}\} = \{(3,6),(4,5),(5,4),(6,3)\} \end{aligned}$$

- Observe that

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

- Nevertheless, any pair of these three events are NOT independent:

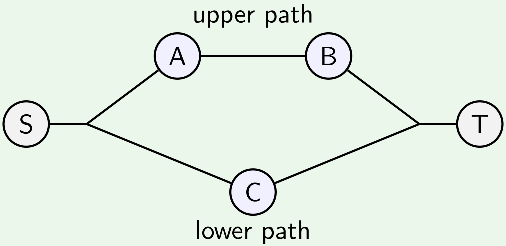
$$\begin{aligned} P(A \cap B) = P\big(\{\text{1st roll is 3}\}\big) &= \frac{1}{6} \neq P(A)\,P(B) \\ P(A \cap C) = P\big(\{(3,6)\}\big) &= \frac{1}{36} \neq P(A)\,P(C) \\ P(B \cap C) = P\big(\{(3,6),(4,5),(5,4)\}\big) &= \frac{1}{12} \neq P(B)\,P(C) \end{aligned}$$

  and thus, these three events are not independent

Example: Using independence for modeling (1/2)

Example (Network connectivity)

Consider a network connecting two nodes S and T through intermediate nodes A, B, C as follows:



Suppose the intermediate nodes A, B, C fail independently, with probabilities

$$p_a = 0.1, \qquad p_b = 0.05, \qquad p_c = 0.2.$$

Question: What is the probability that there is a path connecting S and T without failure?

Example: Using independence for modeling (2/2)

Example (Network connectivity)

Answer: Define two events

$$U = \{\text{the upper path works}\}, \qquad L = \{\text{the lower path works}\}.$$

Then, since A, B are independent,

$$P(U) = (1 - p_a)(1 - p_b), \qquad P(L) = 1 - p_c.$$

Since U depends only on A, B and L depends only on C, the events U and L are independent. Thus,

$$\begin{aligned} P(\text{there is a working path}) &= P(U \cup L) \\ &= 1 - P(U^c \cap L^c) \\ &= 1 - (1 - P(U))(1 - P(L)) \end{aligned}$$

With $p_a = 0.1$, $p_b = 0.05$, $p_c = 0.2$,

$$P(U) = 0.9 \cdot 0.95 = 0.855, \qquad P(L) = 0.8,$$

and consequently,

$$P(\text{there is a working path}) = 1 - (1 - 0.855)(1 - 0.8) = 0.971$$

The counting principle

For a finite sample space Ω with equally likely outcomes,

$$P(A) = \frac{|A|}{|\Omega|}$$

Question: How do we compute the number of elements in A and Ω?

The counting principle (product rule)

Suppose a procedure consists of r successive stages, and the following hold:

(a) There are n_1 possible results at the first stage

(b) For any sequence of possible results until stage $(i-1)$, there are n_i possible results at the i-th stage

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

$$n_1 n_2 \cdots n_r.$$

- **Divide-and-conquer:** Break a complicated counting problem into simple stages
- Multiply the number of choices across stages
- Example: 3 Bernoulli trials have $2 \cdot 2 \cdot 2 = 2^3$ possible outcomes

The counting principle: Examples

Example (The number of subsets)

Question: How many subsets does a set $A = \{a_1, a_2, \ldots, a_n\}$ have?

Answer: We can consider a sequential process that examines one element at a time to decide whether to include it in the subset or not. Therefore, the number of subsets is

$$\underbrace{2 \times 2 \times \cdots \times 2}_{n \text{ times}} = 2^n$$

Example (The number of telephone numbers)

A local telephone number consists of a 7-digit sequence whose first digit cannot be 0 or 1

Question: How many distinct telephone numbers are there within the local area?

Answer: We have a total of 7 stages, and a choice of one out of 10 elements at each stage, except for the first where we only have 8 choices. Therefore, the answer is

$$8 \times \underbrace{10 \times 10 \times \cdots \times 10}_{6 \text{ times}} = 8 \times 10^6$$

Example: Independent trials

Example (3 independent Bernoulli trials)

Consider a sequence of 3 independent Bernoulli trials (e.g., coin tosses):

$$P(\{\text{Head}\}) = p, \qquad P(\{\text{Tail}\}) = 1 - p$$

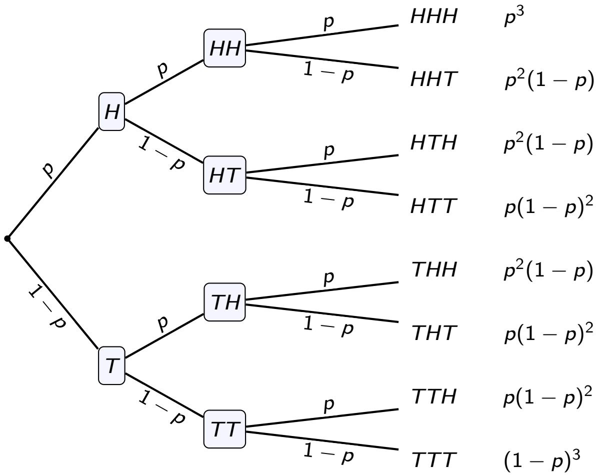
Question: Compute $p(k) = P(k \text{ heads})$

Answer: We can visualize independent Bernoulli trials using a sequential description (next slide)

$$p(k) = \begin{cases} p^3, & k = 3, \\ 3p^2(1-p), & k = 2, \\ 3p(1-p)^2, & k = 1, \\ (1-p)^3, & k = 0. \end{cases}$$

To compute $p(k)$, we need to count how many outcomes produce exactly k Heads
$\longrightarrow$ This leads naturally to the basic principles of counting

Illustration of 3 independent Bernoulli trials



Question: How can we count outcomes systematically, *without listing every outcome*?

Wrap-up

Independence

- A and B are independent if and only if $P(A \cap B) = P(A)P(B)$
- Independence means that learning one event does not change the probability of the other
- Independence is not the same as disjointness
- Independence of several events require factorization of probabilities over every subcollection
- Conditioning can change independence

Modeling with independence

- Combine independent components and trials using products, unions, and complements

Counting

- Multiply numbers of choices across stages; probabilistic independence is not required
- For a finite equally likely model, $P(A) = |A|/|\Omega|$

Suggested reading: [BT08, Ch. 1.5]

Next lecture: Permutations, combinations, and discrete random variables.

References

Dimitri Bertsekas and John N Tsitsiklis.
Introduction to probability, volume 1.
Athena Scientific, 2nd edition, 2008.