03/26/2025: Conditional Probability & Independence

CSCI 246: Discrete Structures

Textbook reference: Sec 32, Scheinerman

Graded Quiz Pickup

Quizzes are in the front of the room, grouped into four bins (A-G, H-L, M-R, S-Z) by last name. The quizzes are upside down with your last name on the back. Come find yours before, during, or after class.

Alert. Some of you weren't here the Friday before spring break. These stacks also include uncollected reading quizzes from Wed. Mar 12 – Intro to Probability (Part 1).

Poll

How many of you would like me to post the slide deck immediately after class so that you can continue working on the group exercises before the solutions are posted?

Today's Agenda

- Reading quiz (5 mins)
- Mini-lecture ($\approx$ 20 mins)
- Group exercises ($\approx$ 20 mins)

Feedback on Monday’s Quiz

Reading Quiz

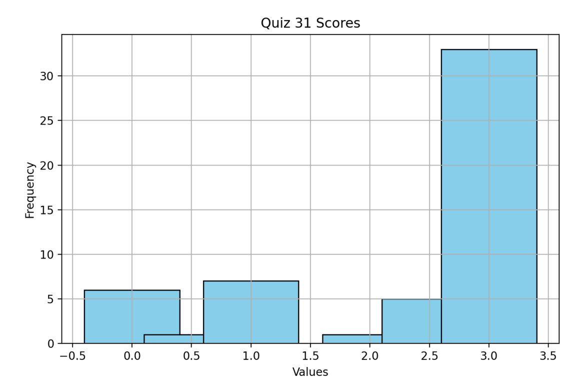


Figure 1: Median Score = 3/3 (100%)

Rubric.

1. 1 point. Correctly stating *False*.
2. 2 points. Correctly providing fix to formula.

Reading quiz

Reading Quiz (Conditional Probability and Independence)

1. Let A and B be events in a probability space (S, P). Suppose $P(B) = \frac{1}{2}$ and $P(A \cap B) = \frac{1}{8}$. Compute $P(A \mid B)$.
2. Let A and B be events in a probability space (S, P). Under what condition does

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

Review solutions to Monday’s group exercises

Thoughts on conditional probability and independence

Conditional probability

Definition

Let A and B be events in a probability space (S, P), and suppose $P(B) \neq 0$. Then the **conditional probability** $P(A \mid B)$, the probability of A given B, is

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

Example

Let A be the event of a student missing the school bus, and B be the event that the student's alarm clock malfunctions.

Both events are rare. But the conditional probability $P(A \mid B)$ is high. It can be interpreted as follows: if we shrink the underlying sample space from S to the event B, what proportion of the little black square $P(B)$ is covered by the shaded area $P(A \cap B)$?

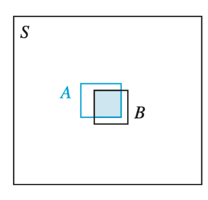


Figure 2: Probability of events A, B, and $A \cap B$, given as areas. Note that the sample space S must have area 1.

Another interpretation of $P(A|B)$

If you sample an outcome from the sample space S and you **know** it's in B, then what's the probability that it's **also** in A?

Remark: Misunderstandings about conditional probability underly some of the most fundamental and important mistakes in human reasoning!

Base rate fallacy

Example: Infectious disease test. Let $\mathcal{I}$ be the event that a person is infected with a disease, and $\mathcal{P}$ be the event that the person tests positive for the disease.

Suppose we collect the following data:

Number of people	Infected	Uninfected	Total
Test positive	20	450	470
Test negative	0	9530	9530
Total	20	9980	10000

The test is very accurate:

$$P(\mathcal{P} \mid \mathcal{I}) = \frac{P(\mathcal{I} \cap \mathcal{P})}{P(\mathcal{I})} = \frac{20}{20} = 1.0, \qquad P(\overline{\mathcal{P}} \mid \overline{\mathcal{I}}) = \frac{P(\overline{\mathcal{I}} \cap \overline{\mathcal{P}})}{P(\overline{\mathcal{I}})} = \frac{450}{9980} = 0.955$$

But the probability of having the disease if you test positive is very low!

$$P(\mathcal{I} \mid \mathcal{P}) = \frac{P(\mathcal{I} \cap \mathcal{P})}{P(\mathcal{P})} = \frac{20}{470} \approx .042$$

Base rate fallacy

Example: Infectious disease test. Let $\mathcal{I}$ be the event that a person is infected with a disease, and $\mathcal{P}$ be the event that the person tests positive for the disease.

Suppose we collect the following data:

Number of people	Infected	Uninfected	Total
Test positive	20	450	470
Test negative	0	9530	9530
Total	20	9980	10000

The test is very accurate:

$$P(\mathcal{P} \mid \mathcal{I}) = \frac{P(\mathcal{I} \cap \mathcal{P})}{P(\mathcal{I})} = \frac{20}{20} = 1.0, \qquad P(\overline{\mathcal{P}} \mid \overline{\mathcal{I}}) = \frac{P(\overline{\mathcal{I}} \cap \overline{\mathcal{P}})}{P(\overline{\mathcal{I}})} = \frac{450}{9980} = 0.955$$

But the probability of having the disease if you test positive is very low!

$$P(\mathcal{I} \mid \mathcal{P}) = \frac{P(\mathcal{I} \cap \mathcal{P})}{P(\mathcal{P})} = \frac{20}{470} \approx .042$$

Other examples. Drunk driving, terrorist identification, *why most published research findings are false.*

Monte Hall Problem

There are three doors on the set for a game show. Behind one door is a car and behind the other two doors are goats.

You get to pick a door to open. The host of the show then opens one of the other doors and reveals a goat. What should you do if you want to maximize your chance of winning the car: stay with your original door or switch – or would the likelihood of winning be the same either way?

Monte Hall Problem

There are three doors on the set for a game show. Behind one door is a car and behind the other two doors are goats.

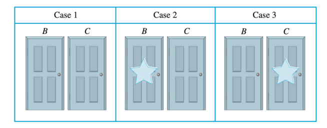
You get to pick a door to open. The host of the show then opens one of the other doors and reveals a goat. What should you do if you want to maximize your chance of winning the car: stay with your original door or switch – or would the likelihood of winning be the same either way?

Try it! `https://www.mathwarehouse.com/monty-hall-simulation-online/`

Solution to Monte Hall Problem

A common (and incorrect) analysis. Once one door has been revealed, the probability of a car being behind either other door is $\frac{1}{2}$, so it doesn't matter whether you switch or not.

The correct analysis.



Just before the host opens one of the closed doors, there is no information about the location of the prize. Call the door you pick A. There are three equally likely possibilities for what lies behind the doors:

- (Case 1) The prize is behind A. Here the host could open either door. You would **win by staying** with your original choice: door A.
- (Case 2) The prize is behind B. Here, the host must open door C, and so you would **win by switching** to door B.
- (Case 3) The prize is behind C. Here, the host must open door B, and so you would **win by switching** to door C.

Thus, in two of the three equally likely cases, you would win by switching.

A quick conversation with a hypothetical skeptic

Question. I find both lines of analysis equally compelling! How can I know for sure that the second analysis is the correct one?

A quick conversation with a hypothetical skeptic

Question. I find both lines of analysis equally compelling! How can I know for sure that the second analysis is the correct one?

Answers.

1. **Mathematical argument.** We can provide a formal mathematical argument using conditional probabilities. For instance, see pp. 227 of Scheinerman.

A quick conversation with a hypothetical skeptic

Question. I find both lines of analysis equally compelling! How can I know for sure that the second analysis is the correct one?

Answers.

1. **Mathematical argument.** We can provide a formal mathematical argument using conditional probabilities. For instance, see pp. 227 of Scheinerman.

2. **Empirical argument.** We can perform repeated experiments (on a computer or in vivo), and find that the number of wins when switching tends towards $\frac{2}{3}$, and the number of wins when staying tends towards $\frac{1}{3}$.

A quick conversation with a hypothetical skeptic

Question. I find both lines of analysis equally compelling! How can I know for sure that the second analysis is the correct one?

Answers.

1. **Mathematical argument.** We can provide a formal mathematical argument using conditional probabilities. For instance, see pp. 227 of Scheinerman.

2. **Empirical argument.** We can perform repeated experiments (on a computer or in vivo), and find that the number of wins when switching tends towards $\frac{2}{3}$, and the number of wins when staying tends towards $\frac{1}{3}$.

Try it! `https://www.mathwarehouse.com/monty-hall-simulation-online/`

Group exercises

aaron.loomis: 10
adam.wyszynski: 15
alexander.goetz: 17
alexander.knutson: 7
anthony.mann: 20
blake.leone: 10
bridger.voss: 5
caitlin.hermanson: 12
cameron.wittrock: 2
carsten.brooks: 13
carver.wambold: 1
colter.huber: 16
conner.reed1: 2
connor.mizner: 15
connor.yetter: 19
derek.price4: 16
devon.maurer: 7
emmeri.grooms: 6
erik.moore3: 17
ethan.johnson18: 11
evan.barth: 8
evan.schoening: 7
griffin.short: 11
jack.fry: 20
jacob.ketola: 9
jacob.ruiz1: 18
jacob.shepherd1: 9
jada.zorn: 8
jakob.kominsky: 16
james.brubaker: 9
jeremiah.mackey: 20
jett.girard: 14
john.fotheringham: 21
jonas.zeiler: 6
joseph.mergenthaler: 13
joseph.triem: 12
julia.larsen: 5
justice.mosso: 19
kaden.price: 6
lucas.jones6: 3
luka.derry: 19
luke.donaldson1: 1
lynsey.read: 13
mason.barnocky: 10
matthew.nagel: 14
micaylyn.parker: 11
michael.oswald: 17
nolan.scott1: 18
owen.obrien: 8
pendleton.johnston: 12
peter.buckley1: 4
reid.pickert: 21
ryan.barrett2: 3
samuel.hemmen: 18
samuel.mosier: 14
samuel.rollins: 4
sarah.periolat: 4
timothy.true: 2
tristan.nogacki: 3
tyler.broesel: 5
william.elder1: 1
yebin.wallace: 15
zeke.baumann: 21

Group exercises

1. Let (S, P) be the sample space with $S = \{1, 2, \ldots, 10\}$ and $P(x) = \frac{1}{10}$ for all $x \in S$. Let A be the event "is even" and B be the event "is prime". Please calculate the following:

a.) $P(A)$	c.) $P(A \mid B)$	e.) $P(\overline{B} \mid A)$	g.) $P(\overline{B} \mid \overline{A})$.
b.) $P(B)$	d.) $P(B \mid A)$	f.) $P(B \mid \overline{A})$	

   Are the events A and B independent?

2. A card is drawn from a well-shuffled deck of 52-cards.
   a. What is the probability that it is a spade (♠)?
   b. What is the probability that it is a king?
   c. What is the probability that it is the king of spades?
   d. Are the events in parts (a) and (b) independent?

3. Prove Scheinerman Prop. 32.4.

4. Let A and B be events in a sample space with $P(A \cap B) \neq 0$. Prove that $P(A \mid B) = P(B \mid A)$ if and only if $P(A) = P(B)$.

Solution to group exercise #1a,b,c,d

Problem. Let (S, P) be the sample space with $S = \{1, 2, \ldots, 10\}$ and $P(x) = \frac{1}{10}$ for all $x \in S$. Let A be the event "is even" and B be the event "is prime". Please calculate the following: (a) $P(A)$, (b) $P(B)$, (c) $P(A \mid B)$, (d) $P(B \mid A)$.

Solution. We have

$$\begin{aligned} S &= \{1, 2, \ldots, 10\} & |S| &= 10 \\ A &= \{2, 4, 6, 8, 10\} & |A| &= 5 \\ B &= \{2, 3, 5, 7\} & |B| &= 4 \\ A \cap B &= \{2\} & |A \cap B| &= 1 \end{aligned}$$

By the equally likely probability formula, we have

$$P(A) = \frac{|A|}{|S|} = 5/10, \qquad P(B) = \frac{|B|}{|S|} = 4/10, \qquad P(A \cap B) = \frac{|A \cap B|}{|S|} = 1/10.$$

Hence,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/10}{4/10} = 1/4, \qquad P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{1/10}{5/10} = 1/5.$$

Interpretation: $P(A \mid B)$ gives the probability of randomly drawing an even number from the prime numbers in S. It makes sense that $P(A \mid B) = 1/4$: looking at set B, there are 4 prime numbers in S, only one of them is also even, and all of these numbers are equally likely to be drawn. Similarly, $P(B \mid A)$ gives the probability of randomly drawing a prime number from the even numbers in S. Note there are 5 even numbers, and only one of them is prime.

Solution to group exercise #1e,f,g

Problem. Let (S, P) be the sample space with $S = \{1, 2, \ldots, 10\}$ and $P(x) = \frac{1}{10}$ for all $x \in S$. Let A be the event "is even" and B be the event "is prime". Please calculate the following: (e) $P(\overline{B} \mid A)$, (f) $P(B \mid \overline{A})$, (g) $P(\overline{B} \mid \overline{A})$.

Solution. We have

$$\begin{aligned} S &= \{1, 2, \ldots, 10\} & |S| &= 10 \\ \overline{A} &= \{1, 3, 5, 7, 9\} & |\overline{A}| &= 5 \\ \overline{B} &= \{1, 4, 6, 8, 9, 10\} & |\overline{B}| &= 6 \\ A \cap \overline{B} &= \{4, 6, 8, 10\} & |A \cap \overline{B}| &= 4 \\ \overline{A} \cap B &= \{3, 5, 7\} & |\overline{A} \cap B| &= 3 \\ \overline{A} \cap \overline{B} &= \{1, 9\} & |\overline{A} \cap \overline{B}| &= 2 \end{aligned}$$

We have

$$P(\overline{B} \mid A) = \frac{\overline{B} \cap A}{P(A)} = \frac{4/10}{5/10} = 4/5.$$

$$P(B \mid \overline{A}) = \frac{B \cap \overline{A}}{P(\overline{A})} = \frac{3/10}{5/10} = 3/5$$

$$P(\overline{B} \mid \overline{A}) = \frac{\overline{B} \cap \overline{A}}{P(\overline{A})} = \frac{2/10}{5/10} = 2/5.$$

Example Interpretation: $P(B \mid \overline{A})$ gives the probability of randomly drawing a prime number from the odd (non-even) numbers in S. Looking at $\overline{A}$, we see that there are 5 odd numbers (all of them equally likely to be drawn), and 3 of them are prime.

Solution to group exercise #2

Problem. A card is drawn from a well-shuffled deck of 52-cards.

a. What is the probability that it is a spade (♠)?
b. What is the probability that it is a king?
c. What is the probability that it is the king of spades?
d. Are the events in parts (a) and (b) independent?

Solution. Let A be the event "is a spade" and B be the event "is king". By the equally likely probability formula,

a. $P(A) = \frac{|A|}{|S|} = \frac{13}{52} = \frac{1}{4}$.
b. $P(B) = \frac{|B|}{|S|} = \frac{4}{52} = \frac{1}{13}$.
c. $P(A \cap B) = \frac{|A \cap B|}{|S|} = \frac{1}{52} = \frac{1}{52}$.
d. Yes. Note that

$$\frac{1}{52} = P(A \cap B) = P(A) \times P(B) = \frac{1}{4} \times \frac{1}{13}$$

so the definition of independent events is verified. Intuitively, the suit of a card (e.g. spade) provides no information about its rank (e.g. king).

Solution to group exercise #3

Problem. Prove Scheinerman Prop. 32.4.

Solution.

- $\boxed{(3) \implies (1).}$

$$P(A \mid B) \overset{\textit{def.}}{=} \frac{P(A \cap B)}{P(B)} \overset{\textit{(3)}}{=} \frac{P(A)\cancel{P(B)}}{\cancel{P(B)}} = P(A).$$

- $\boxed{(3) \implies (2).}$

$$P(A \mid B) \overset{\textit{def.}}{=} \frac{P(A \cap B)}{P(A)} \overset{\textit{(3)}}{=} \frac{\cancel{P(A)}P(B)}{\cancel{P(A)}} = P(B).$$

- $\boxed{(1) \implies (3).}$

$$P(A \cap B) \overset{\textit{def.}}{=} P(A \mid B)P(B) \overset{\textit{(1)}}{=} P(A)P(B)$$

- $\boxed{(2) \implies (3).}$

$$P(A \cap B) \overset{\textit{def.}}{=} P(B \mid A)P(A) \overset{\textit{(2)}}{=} P(B)P(A)$$

In the above, "*def.*" refers to the definition of conditional probability.

Solution to group exercise #4

Problem. Let A and B be events in a sample space with $P(A \cap B) \neq 0$. Prove that $P(A \mid B) = P(B \mid A)$ if and only if $P(A) = P(B)$.

Remark. Recall that to prove an "if and only if" statement, we must prove that if-then propositions hold in both directions.

Solution.

- $\boxed{\implies}$. We assume $P(A \mid B) = P(B \mid A)$ and show $P(A) = P(B)$. Since $P(A \mid B) = P(B \mid A)$, by the definition of conditional probability, we have

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

  Cross multiplying, we obtain

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

  Since $P(A \cap B) \neq 0$, we can divide both sides by it, leaving $P(A) = P(B)$.

- $\boxed{\impliedby}$. We assume $P(A) = P(B)$ and show $P(A \mid B) = P(B \mid A)$. We have

  $$P(A \mid B) \overset{\text{def.}}{=} \frac{P(A \cap B)}{P(B)} \overset{\text{assumpt.}}{=} \frac{P(A \cap B)}{P(A)} \overset{\text{def.}}{=} P(B \mid A).$$