

AP® Statistics

2A

Probability

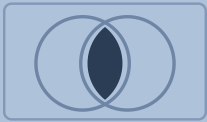
WHAT ARE THE CHANCES?



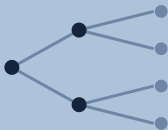
TWO-WAY TABLES



SAMPLE SPACE



VENN DIAGRAMS



TREE DIAGRAMS

Guided Notes Packet

DSAHS 2026-2027 · MR. LEVY

Two Categorical Variables

Two-way tables · marginal, joint & conditional relative frequencies · bar graphs
· association

Two-way table A table of _____ that summarizes data on the relationship between _____ categorical variables for some group of individuals. Also called a _____ table — the AP® exam uses both names.

Yellowstone National Park surveyed 1526 winter visitors about environmental club membership and snowmobile use.

SNOWMOBILE USE	CLUB MEMBER: NO	CLUB MEMBER: YES	TOTAL
Never used	445	212	657
Snowmobile renter	497	77	574
Snowmobile owner	279	16	295
Total	1221	305	1526

Sanity check. Before computing anything, make sure the table adds up — every path to the grand total must agree:

- a) All six cells: $445 + 212 + 497 + 77 + 279 + 16 = \underline{\hspace{2cm}}$
- b) Row totals: $657 + 574 + 295 = \underline{\hspace{2cm}}$
- c) Column totals: $1221 + 305 = \underline{\hspace{2cm}}$

Marginal relative frequency The percent or proportion of individuals that have a specific value for _____ categorical variable. Found in the _____ of the table.

- a) Find the marginal relative frequencies for environmental club membership.
- b) Find the marginal relative frequencies for snowmobile use.

Joint relative frequency The percent or proportion of individuals that have a specific value for one categorical variable **and** a specific value for another categorical variable.

- a) What percent of people in the sample are environmental club members *and* own snowmobiles?
- b) What proportion of people in the sample are not environmental club members *and* never use snowmobiles?

AP® EXAM TIP

Why "joint"? The joints in your body are where two bones come together. A joint relative frequency is found where a **row** and a **column** come together.

Who wrote the hits? The Beatles have more #1 songs on the Billboard Hot 100 than any other artist. All 205 Beatles songs released in the U.S.:

	HARRISON	LENNON	MCCARTNEY	STARR	MULTIPLE	TOTAL
Reached #1	0	5	12	0	3	20
Did not reach #1	25	69	63	2	26	185
Total	25	74	75	2	29	205

- What percentage of Beatles songs reached number one?
- What proportion of Beatles songs were written by Paul McCartney alone and reached number one?

Which one is marginal and which is joint? Label each.

Conditional relative frequency The percent or proportion of individuals that have a specific value for one categorical variable **among** individuals who share the same value of another categorical variable (the _____).

The denominator follows the "of." Both questions below are conditional — but they condition in opposite directions.

- What percentage of songs written by Paul McCartney alone reached number one?
- What proportion of number one songs were written by Paul McCartney alone?

AP® EXAM TIP

When translating words into proportions, the value in the **denominator** is found next to the word "of." Percentage OF McCartney songs? Divide by _____. Proportion OF number one songs? Divide by _____.

Conditional distribution The distribution of one categorical variable among individuals who share the same value of another variable — the _____ column or row at once.

SNOWMOBILE USE	CLUB MEMBER: NO	CLUB MEMBER: YES	TOTAL
Never used	445	212	657
Snowmobile renter	497	77	574
Snowmobile owner	279	16	295
Total	1221	305	1526

- Find the conditional distribution of snowmobile use among environmental club members.
- Find the conditional distribution of snowmobile use among those who are *not* club members.

AP® EXAM TIP

When comparing groups of **different sizes**, use relative frequencies (percents or proportions), not counts. Avoid “More club members never use snowmobiles” when you mean “A greater *percentage* of club members never use snowmobiles.”

Three graphs for two categorical variables.

Side-by-side bar graph Displays the distribution of a categorical variable for each value of another. Bars are _____ together and placed side by side.

Segmented bar graph Displays the distribution as segments of a rectangle, with the area of each segment proportional to the _____ of individuals in that category. Every bar reaches _____ %.

Mosaic plot A modified segmented bar graph in which the _____ of each rectangle is proportional to the _____ of individuals in that category.

Which of the three shows that far more respondents are *NOT* environmental club members — and how can you tell?

Association There is an association between two variables if knowing the value of one variable helps us _____ the value of the other. If knowing the value of one variable does **not** help us predict the other, there is _____ association.

AN ASSOCIATION LOOKS LIKE...

Segment heights _____ between the groups.

NO ASSOCIATION LOOKS LIKE...

Every group's segments sit at the _____ heights.

CAUTION: association does not necessarily imply _____ !

Practice — Titanic. 1317 passengers by class and survival:

	SURVIVED	DID NOT SURVIVE	TOTAL
First Class	201	123	324
Second Class	118	166	284
Third Class	181	528	709
Total	500	817	1317

1. Calculate the **marginal** relative frequency for 3rd Class.
2. Calculate the **joint** relative frequency for 2nd Class survival.
3. Calculate the **conditional** relative frequency for survival given the passenger was in 1st Class.
4. Is there an association between class and survival? *Justify your claim.*

TOPIC 2.3 · UNIT 2A

Estimating Probabilities Using Simulation

The idea of probability · the law of large numbers · outcomes & events · running a simulation

Probability The probability of any outcome of a chance process is a number between _____ and _____ that describes the _____ of times the outcome would occur in a very _____ series of repetitions.

Law of large numbers If we observe more and more repetitions of any chance process, the proportion of times a specific outcome occurs gets _____ and _____ to its probability.

Randomness is predictable in the _____, but not in the _____.

Law of averages The *misguided* belief that the results of a chance process have to _____ in the short run. This is NOT a real law.

Understanding check. Two games with a fair coin. Game 1: you win if you throw between 45% and 55% heads. Game 2: you win if you throw more than 60% heads. For each game, would you rather flip **20** times or **200** times?

Four words, in order. Random process → trial → outcome → event.

Random process Generates outcomes that are determined purely by _____.

Trial One _____ of a random process.

Outcome The result of _____ trial of a random process.

Event Any _____ of outcomes from some chance process.
We use capital letters (A, B, C) to name events.

One flip of a coin is a _____; landing heads is an _____; "at least one head in three flips" is an _____.

Where does a probability live? Every probability is a number from **0** to **1**.
Place each event on the number line below.

MARK 0, 0.5 AND 1 — LABEL IMPOSSIBLE / 50-50 / CERTAIN

- a) rolling a 7 on one standard die
- b) a fair coin lands heads
- c) the sun rises tomorrow
- d) rolling doubles with two dice

Interpret this. Gallup reports that 76% of U.S. Gen Z members believe they have a great future ahead of them, so the probability a randomly selected U.S. Gen Zer believes this is 0.76.

- a) Interpret this probability.
- b) If a researcher randomly selects 100 U.S. Gen Zers, will exactly 76 of them believe they have a great future? Explain.

Quick check — short-run superstitions.

- a) I just flipped a fair coin 6 times and got tails every time. Is heads more likely on the next flip?

b) Which is more likely in 6 flips of a fair coin: TTTTTT or THHTTH?

Simulation The imitation of _____ behavior, based on a model that accurately reflects the situation.

The simulation process:

1. Describe how to use a chance device to imitate _____ trial. Tell what you will _____ at the end of each trial.
2. Perform _____ trials of the simulation.
3. Use the results to _____ the probability and answer the question of interest.

A simulation **estimates** a probability. Use $\approx$, never $=$.

The 23-box problem. A cereal company puts one of 5 NASCAR driver cards in each box — Logano, Harvick, Elliott, Patrick, Johnson — and claims all 5 are equally likely. A fan buys boxes until she has all 5. It takes her **23 boxes**. Does this provide convincing evidence that the 5 cards are not equally likely?

- a) Describe how to use a random number generator to perform one trial of the simulation.

The dotplot in class shows the number of boxes needed in 50 trials.

- b) Explain what the dot at 20 represents.
- c) Estimate the probability that it takes 23 or more boxes. Does this give convincing evidence the cards are not equally likely?

Who gets tested? A company randomly selects 4 employees per month for a drug test. There are **25** employees on the factory floor and **7** in the front office. This month all 4 selected are from the floor, and the floor workers suspect the selection was not fair.

- a) Describe how you would use a random number generator to carry out this simulation.
- b) Perform 3 trials. Record the numbers and what they represent.

- c) In 39 of 100 trials, all 4 were from the floor. Is this convincing evidence the process was unfair?

AP® EXAM TIP

If asked to describe a simulation, your written process alone must let the reader reproduce your results. When sampling **without** replacement, say that repeated numbers are ignored.

Practice — the \$15 million violin. 10 world-class soloists played 12 violins (6 new, 6 old, including a Stradivarius) wearing modified welder's goggles so they could not tell which was which. **9 of the 10** chose a new violin.

- a) Assuming soloists are equally likely to choose any violin, describe how to use a random number generator for one trial.
- b) Explain what the dot at 9 represents.
- c) Estimate the probability that 9 or more soloists would choose a new violin.
- d) Is there convincing evidence that new violins are more likely to be chosen?

Two ways to find a probability.

EMPIRICAL

Estimated from _____ or simulation
— observe many trials and use the proportion that occurred. Always an estimate: use _____.

THEORETICAL

Calculated from a _____ of the chance process. A fair coin has two equally likely outcomes, so $P(\text{heads}) = 0.5$ exactly. Use _____.

Introduction to Probability

Probability models · sample space · the basic rules · the complement rule

Probability model A description of some chance process that consists of two parts: (1) a list of all _____, and (2) the _____ for each outcome.

Sample space The list of all possible _____ of a chance process. A sample space can be very simple or very complex.

Identify the sample space for one trial of each process:

- a) Flipping a coin $S =$ _____
- b) Rolling a standard die $S =$ _____
- c) Tossing 3 coins and recording the number of tails $S =$ _____

Rolling two dice. The sample space has _____ outcomes. Because the dice are fair, each outcome is equally likely with probability _____.

SKETCH THE 6 × 6 SAMPLE SPACE

- a) Let A = getting double 1's. Find $P(A)$.
- b) Let B = rolling doubles. Find $P(B)$.
- c) Let C = getting a sum of 5. Find $P(C)$.

Two facts that fall out of the sample space:

- The sum of the probabilities of all outcomes in a sample space is _____. $P(S) =$ _____
- The outcomes in a sample space are _____ — they do not overlap.

Equally likely outcomes. If all outcomes in the sample space are equally likely, then

$$P(A) = \frac{\text{number of outcomes in event } A}{\text{total number of outcomes in } S}$$

Prime time — build the model yourself. Roll a fair **four-sided** die and a fair **six-sided** die. The table shows every possible sum.

4 - SIDE + 6 - SIDE	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10

- Give a probability model for this random process.
- Let A = the sum is a prime number (2, 3, 5, or 7). Find $P(A)$.

AP® EXAM TIP

List, don't guess. Writing out the full sample space is often the fastest path to a correct probability on the AP® exam.

Basic probability rules. Rules a valid probability model must obey:

- If all outcomes are equally likely, $P(A) = (\text{outcomes in } A) \div (\text{total outcomes in } S)$.
- The probability of any event is a number between _____ and _____.
- All possible outcomes together must have probabilities that add up to _____.
- The probability that an event does **not** occur is _____ minus the probability that it does occur.

Complement rule $P(A^C) = \text{_____}$, where A^C is the complement of event A — the event that A does _____ occur.

Three accepted notations for the complement of A : A' , $\bar{A}$ (A -bar), and A^C . We will use A^C .

AP® EXAM TIP

The complement rule is **NOT** on the AP® Statistics formula sheet. Know it by heart.

Don't lose free points. A probability is ALWAYS a number from 0 to 1, inclusive. No exceptions, ever.

AP® EXAM TIP

You lose any chance for partial credit on a probability free-response question if you state a probability below 0 or above 1. Be careful with scientific notation: **2.32445E-4** means **0.000232445** — not 2.32445.

Mixed Review. Practice using proper notation. Show the probability statement before you find the number.

1. A card is drawn from an ordinary deck. Find the probability of getting:
 - a) the ace of spades
 - b) a black card
 - c) a heart
 - d) a queen
 - e) a red 6

2. A couple has had 3 babies in the last four years. Find the probability that exactly 2 of them are girls.

3. When a single standard die is rolled, what is the probability of getting a 9?

4. When a single standard die is rolled, what is the probability of getting a number less than 7?

Mixed Review, continued.

5. Find the **complement** of each event:
 - a) selecting a month that has 30 days
 - b) selecting a day of the week that begins with the letter S
 - c) rolling two dice and getting a sum of 7
 - d) selecting a letter of the alphabet that is a vowel

6. In a study, 24% of people who were victims of a violent crime were age 20 to 24. If a victim is selected at random, what is the probability they are younger than 20 *or* older than 24?

7. In a sample of 50 people, 21 had type O blood, 22 had type A, 5 had type B, and 2 had type AB.

BLOOD TYPE	O	A	B	AB	TOTAL
Frequency					
Probability					

- a) Set up the frequency distribution above.
- b) $P(\text{a person has type O blood})$
- c) $P(\text{a person has type A or type B blood})$
- d) $P(\text{a person has neither type A nor type O blood})$
- e) $P(\text{a person does not have type AB blood})$

Whorls, loops, and arches. Every fingerprint is unique, but each fits a basic pattern. Britannica gives this probability model for a randomly selected fingerprint:

PATTERN	LOOP	WHORL	ARCH
Probability	0.65	0.30	0.05

- a) Explain why this is a valid probability model.
- b) Find the probability that a random fingerprint is NOT a loop.
- c) Find the probability that a random fingerprint is a loop or an arch.

TOPIC 2.5 · UNIT 2A

Mutually Exclusive Events

"Or" and unions · the addition rules · two-way tables · Venn diagrams

One event or another. At a large political gathering, a person is selected at random.

- a) You want the probability the person is female **or** a Republican. How many possibilities are there to consider?
- b) Same gathering: what about the probability they are a Democrat **or** an Independent? How many now?

SOUP OR SALAD?

The waiter wants you to choose one or the other, but _____ both.

CREAM OR SUGAR?

It is fine to ask for one, the other, _____ both.

In mathematics and probability, “A or B” always means one, the other, _____ both.

| Mutually exclusive (disjoint) Two events A and B are mutually exclusive if they have _____ outcomes in common and so can never occur together — that is, if $P(A \cap B) = \underline{\hspace{2cm}}$. Read $\cap$ as “and.”

Addition rule for mutually exclusive events:

$$P(A \cup B) = \underline{\hspace{2cm}} \quad (\text{read } \cup \text{ as “or”})$$

CAUTION: this rule works *ONLY* for mutually exclusive events.

Practice — Mutually Exclusive Events. Determine whether the two events are mutually exclusive. Write YES or NO and give a one-line reason.

1. Randomly selecting a female student; randomly selecting a student who is a junior
2. Randomly selecting a person with type A blood; randomly selecting a person with type O blood
3. Rolling a die and getting an odd number; rolling a die and getting a number less than 3
4. Randomly selecting a person who is under 21 years of age; randomly selecting a person who is over 30 years of age

Prove it. Computing is not enough — the AP® exam asks you to **justify** using the joint probability.

- a) One random fingerprint. Events: “loop” and “arch.” Mutually exclusive? Prove it.
- b) One card from a standard 52-card deck. Events: “king” and “heart.” Mutually exclusive? Prove it.

The basic rules in symbolic form. Fill each one in — these are the ones you owe me cold.

- For any event A: _____ $\leq P(A) \leq$ _____
- $P(S) =$ _____ — the sample space always has probability 1
- Equally likely outcomes: $P(A) =$ _____
- Complement rule: $P(A^C) =$ _____
- Mutually exclusive: $P(A \cap B) =$ _____
- Addition rule (ME events only): $P(A \cup B) =$ _____

Taste the rainbow? Wrong candy. Mars, Inc. gives this probability model for M&M'S Milk Chocolate Candies from its Cleveland factory:

COLOR	BLUE	ORANGE	GREEN	YELLOW	RED	BROWN
Probability	0.207	0.205	0.198	0.135	0.131	0.124

- Explain why this is a valid probability model.
- Find the probability that you do NOT get a blue M&M.
- What is the probability you get an orange or a brown M&M?

Joint probability The value where a row and a column come together, $P(A \cap B)$. It is the same number you called a joint _____ frequency earlier in this unit — new name, same idea. And if $P(A \cap B) = 0$, the events are _____.

	<\$50K	\$50-\$100K	>\$100K	TOTAL
Very likely	184	114	84	382
Somewhat likely	84	65	41	190
Not very likely	13	27	5	45
Not at all likely	22	9	7	38
Not sure	113	55	28	196
Total	416	270	165	851

Survey: how likely is it that alien technology has visited Earth? By family income, 851 adults.

Order does not matter: "A and B" is the same event as "B and A," so $P(A \cap B) = P(\text{_____} \cap \text{_____})$.

Pierced ears. A statistics class recorded gender and pierced-ear status for all 178 people. Choose a student at random. **A:** male. **B:** has a pierced ear.

PIERCED EAR	MALE	FEMALE	TOTAL
Yes	19	84	103
No	71	4	75
Total	90	88	178

- Find $P(B)$.

- b) Find $P(A \cap B)$. Interpret this value in context.
- c) Find $P(A \cup B)$.

For (c): what goes wrong if you just add $P(A) + P(B)$?

The double-counting fix. Adding $P(A) + P(B)$ counts the overlap _____ . Subtract it back out once.

General addition rule:

For any two events A and B, $P(A \cup B) =$ _____

Lesson: you cannot use the mutually-exclusive addition rule unless A and B have _____ outcomes in common.

Facebook and Instagram. A survey of all residents in a large apartment complex: **68%** use Facebook, **28%** use Instagram, and **25%** do both. Select a resident at random.

Let F = uses Facebook and I = uses Instagram. Find the probability the person uses Facebook or Instagram.

Aw, nuts! During one month, **35%** of a candy maker’s chocolate bars contained nuts, **20%** contained caramel, and **12%** contained both. Grab one bar at random.

	CONTAIN NUTS	NO NUTS	TOTAL
Caramel			
Not caramel			
Total			

Fill the table in as you work — the grand total is 100%.

- a) Which cell does the 12% go in?
- b) Where do the 35% and the 20% live?
- c) Find $P(N \cup C)$.

Venn diagram Consists of one or more _____ surrounded by a rectangle. Each circle represents an _____. The region inside the rectangle represents the _____ of the chance process.

Label each region for the pierced-ears table:

REGION	IN WORDS	COUNT
In the intersection	Male and pierced ear	19
Inside A, outside B	Male and no pierced ear	71
Inside B, outside A	Female and pierced ear	84
Outside both circles	Female and no pierced ear	4

A^c — COMPLEMENT

Everything _____ in A.

Shaded region = _____

$A \cap B$ — INTERSECTION

Outcomes in _____ events.

The overlap.

$A \cup B$ — UNION

Outcomes in A, in B, or in _____.

Back to Facebook and Instagram. 68% Facebook, 28% Instagram, 25% both.

a) Make a Venn diagram to display the sample space using F and I. Put a number in every region.

DRAW AND LABEL YOUR VENN DIAGRAM

b) Find the probability the person uses neither Facebook nor Instagram.

Understanding check — table or Venn, dealer’s choice. One adult is chosen at random from the alien-technology survey above. **A:** answered “not sure.” **B:** family income under \$50,000.

- a) Find $P(A)$.
- b) Find $P(A \cap B)$.
- c) Find $P(A \cup B)$.
- d) Are A and B mutually exclusive? Justify using a joint probability.

TOPIC 2.6 · UNIT 2A

Conditional Probability

Probability given that something has happened · the conditional formula
· independence

Can you do both? One student is chosen at random from a college stats class. **A**: can roll their tongue. **B**: can raise one eyebrow. *Try both right now.*

	TONGUE: YES	TONGUE: NO	TOTAL
Eyebrow: Yes	42	20	62
Eyebrow: No	107	31	138
Total	149	51	200

- a) Given that the student can raise one eyebrow, what is the probability they can roll their tongue?
- b) Given that the student can roll their tongue, what is the probability they can raise one eyebrow?

Same 42 in the numerator both times. What changed?

| Conditional probability The probability that one event happens **given that** another event is known to have happened. The conditional probability that event B happens given that event A has happened is written

_____ .

Read the vertical bar | as " _____ " or "under the condition that."

Calculating a conditional probability:

$P(A \mid B) =$ _____

The choice of which event is A and which is B is arbitrary, so it is also true that $P(B \mid A) =$ _____

Titanic. Randomly select one adult passenger. **F**: first-class. **S**: survived.
T: third-class.

	FIRST	SECOND	THIRD	TOTAL
Survived	197	94	151	442
Died	122	167	476	765
Total	319	261	627	1207

- a) Find $P(T \mid S)$. Interpret this value in context.
- b) Given that the chosen person is NOT a first-class passenger, what is the probability they survived? Write your answer as a probability statement using correct symbols.

You already knew this. A conditional probability is just a conditional _____ frequency that comes from a random process — you have been computing these since two-way tables opened this unit.

Counts: $P(A \mid B) = 42/62 = \underline{\hspace{2cm}}$ Probabilities: $0.21/0.31$
 $= \underline{\hspace{2cm}}$

Joint probability divided by the _____ probability — exactly the formula above.

Facebook and Instagram, one more time. 68% Facebook, 28% Instagram, 25% both. Given that the person uses Facebook, what is the probability they use Instagram?

Practice — you don't look so good. 16 people were photographed twice: once after a placebo injection, once after a bacterium that causes inflammation. 2945 volunteers rated each photo "healthy" or "sick" in under 5 seconds. **H**: rated healthy. **S**: rated sick. **B**: bacterium photo.

	BACTERIUM	PLACEBO	TOTAL
Rated healthy	698	1032	1730
Rated sick	775	440	1215
Total	1473	1472	2945

- a) Find $P(S \mid B)$. Describe it in words.
- b) Given the rating was "healthy," what is the probability the photo was the placebo shot? Write it in correct notation.

No memory. Toss a fair coin twice. A: first toss is a head. B: second toss is a head. The coin has no memory.

Independent events A and B are independent if knowing whether or not one event has occurred does _____ change the probability that the other will happen.

A and B are independent if
 $P(A \mid B) = P(A \mid B^C) =$ _____
or, equivalently,
 $P(B \mid A) = P(B \mid A^C) =$ _____

AP® EXAM TIP
The formulas for determining whether two events are independent are **NOT** on the AP® Statistics formula sheet. Know them cold.

Pierced ears again. Are “male” and “pierced ear” independent? Work both conditionals and compare.

PIERCED EAR	MALE	FEMALE	TOTAL
Yes	19	84	103
No	71	4	75
Total	90	88	178

- a) $P(\text{pierced ear} \mid \text{male}) = P(B \mid A) =$
- b) $P(\text{pierced ear} \mid \text{female}) = P(B \mid A^C) =$
- c) Verdict — independent or not? Say why.

The CED check: compare to the unconditional. The exam’s favourite version — compare a conditional probability to the plain unconditional one, straight off the table.

- No condition at all: $P(\text{pierced ears}) = P(B) =$ _____
- $P(B \mid A) =$ _____ and $P(B \mid A^C) =$ _____
- Independent would mean $P(B \mid A)$ _____ $P(B)$. Does it?

Handedness. An SRS of 100 Australian high school students. Choose one at random.

DOMINANT HAND	MALE	FEMALE	TOTAL
Right	39	51	90
Left	7	3	10
Total	46	54	100

- a) Are the events “male” and “left-handed” independent? Justify your answer.
- b) Now run the CED-style check: compare $P(\text{left})$ to $P(\text{left} \mid \text{male})$. Same verdict?

Independence means no association. If events from a two-way table are independent, the two variables have _____ association — every bar in the segmented bar graph splits the _____ way.

- a) Is there an association between longer foot and longer finger in the sample shown in class?
- b) Does that prove an association in the population of all seniors who took the survey? Explain.

TOPIC 2.7 · UNIT 2A

Independent Events and Unions of Events

The multiplication rules · tree diagrams · reversing the condition · at-least-one problems

The general multiplication rule. About 55% of high school students play a school sport, and about 6% of those athletes go on to play in the NCAA.

For any chance process:

$$P(A \cap B) = \underline{\hspace{2cm}}$$

$$P(\text{high school sport} \cap \text{NCAA}) = (0.55)(\underline{\hspace{1cm}}) = \underline{\hspace{1cm}}$$

AP® EXAM TIP

The general multiplication rule is **NOT** on the AP® Statistics formula sheet. Know it cold.

Social media and pressure. Pew reports that **79%** of teens use social media, and **39%** of teens who use social media feel pressure to post

content that will be popular.

Find the probability a randomly selected teen uses social media AND feels that pressure.

Candy grab: no put-backs. A bag holds 20 candies — 15 cherry, 5 lemon. You grab 3 at random, eating as you go. Nothing goes back.

- a) Find $P(\text{all three cherry})$.
- b) Find $P(\text{no cherry at all})$.

Think in **stages**: each draw's probability is conditional on what is already gone. Always ask — with or _____ replacement?

| Tree diagram Shows the sample space of a chance process involving _____ stages. The probability of each outcome is shown on the corresponding _____ of the tree.

Key fact: all probabilities after the first stage are _____. probabilities. Every set of branches from a single point sums to _____.

Multiply along the branches. **Add across** them.

Snooze you lose. Shannon hits snooze on **60%** of school days. If she snoozes, she makes it to first period on time with probability **0.70**. If she gets right up: **0.90**. Select a school day at random.

DRAW AND LABEL THE TREE — PUT A PROBABILITY ON EVERY BRANCH

- a) Why do the second-stage branches read 0.70 / 0.30 and 0.90 / 0.10?
- b) Find $P(\text{snooze } \cap \text{ late})$.
- c) Find $P(\text{late})$. *How many ways can Shannon be late?*

Ebooks by generation. Harris Interactive: 20% of millennials, 25% of Gen Xers, 21% of baby boomers and 17% of matures read more ebooks than print. Census Bureau: of U.S. adults, 34% are millennials, 22% Gen X, 30% boomers, 14% matures. Select one U.S. adult at random.

a) Draw a tree diagram to model this chance process.

FOUR FIRST-STAGE BRANCHES

- b) Find the probability the person reads more ebooks than print books.
- c) Suppose the chosen person reads more ebooks than print. What is the probability they are a millennial?

Going in reverse. Shannon was late today. What is the probability she hit snooze? The condition now sits at the _____ of the tree — we are driving it backward.

$P(\text{snooze} \mid \text{late}) =$ _____

Covid scent dogs. July 2020: dogs were trained to detect Covid-19 in human saliva. They alerted on **83%** of infected subjects and stayed quiet on **96%** of uninfected subjects. About **2%** of U.S. residents were infected. A random resident gets sniffed — and the dog alerts.

Before any math — gut check. How sure are you that this person really has the virus?

TREE DIAGRAM

- Find $P(\text{virus detected})$.
- Find $P(\text{infected} \mid \text{virus detected})$.
- What does this tell you about screening for a *rare* condition?

Practice — reverse the condition. Use the tree given in class.

- Label the tree.
- Make a two-way table of counts (use 100 people).
- Determine whether M and G are independent events. Justify.

	G	G ^C	TOTAL
M			
F			
Total			

The multiplication rule for independent events. When A and B are independent, $P(B \mid A) = \underline{\hspace{2cm}}$, so the general rule simplifies.

If A and B are independent:
 $P(A \cap B) = \underline{\hspace{2cm}}$

CAUTION: this rule applies *ONLY* to independent events.

AP® EXAM TIP

The multiplication rule for independent events is **NOT** on the AP® Statistics formula sheet. Know it cold.

Independence check #2: multiply and compare. If A and B are independent then $P(A \cap B) = P(A) \cdot P(B)$. So multiply the unconditional probabilities and compare to what you actually observe.

	TONGUE: YES	TONGUE: NO	TOTAL
Eyebrow: Yes	42	20	62
Eyebrow: No	107	31	138
Total	149	51	200

- a) $P(A) = \underline{\hspace{2cm}}$ $P(B) = \underline{\hspace{2cm}}$
b) Independence would predict $P(A \cap B) = \underline{\hspace{2cm}}$
c) Observed $P(A \cap B) = \underline{\hspace{2cm}}$. Verdict?

Or still means union. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Nothing new — but now you have more ways to find the overlap.

$P(\text{rain tomorrow}) = 0.3$, $P(\text{your bus runs late}) = 0.2$, and the two events are **independent**. Find $P(\text{rain} \cup \text{bus late})$.

Triple defect. The probability that a randomly selected Type A product is defective is 0.1. You select three Type A products **independently**. Find the probability all three are defective.

Rapid HIV tests. Applied to people who do NOT have HIV, one rapid test has probability about **0.004** of producing a false positive. A clinic tests **200** randomly selected people who are free of HIV antibodies. Assume results for different individuals are independent.

Find the probability that **at least one** false positive occurs.

PRO TIP

When asked for $P(\text{at least one})$, compute $1 - P(\text{none})$. It saves an enormous amount of calculation.

Mutually exclusive $\neq$ independent. Choose a U.S. resident at random. A: the person is younger than 10. B: the person has a driver’s licence.

MUTUALLY EXCLUSIVE?
_____ — they cannot happen together, so $P(A \cap B) =$ _____.

INDEPENDENT?
_____ ! If A happens, B is impossible: $P(B \mid A) =$ _____ $\neq P(B)$.

Mutually exclusive events with nonzero probabilities can _____ be independent — one happening guarantees the other does not.

The tragic case of Sally Clark. England, 1999: Sally Clark was convicted of murdering her two infant sons after an expert testified that two SIDS deaths in one family had probability $(1/8543)^2$ — about 1 in 73 million. Her conviction was overturned in 2003, after more than three years in prison.

What went wrong — in the exact language of this section?

Understanding check — sports and jobs. 200 students at a school; one is chosen at random. **S**: plays a school sport. **J**: has a part-time job.

	JOB: YES	JOB: NO	TOTAL
Sport: Yes	30	50	80
Sport: No	60	60	120
Total	90	110	200

- a) Find $P(J \mid S)$.
- b) Are S and J independent? Compare $P(J \mid S)$ to $P(J)$.
- c) Two *different* students are chosen at random. Find $P(\text{both play a sport})$.
- d) Find $P(S \cup J)$.