STAT 515 Lec 03 slides

Conditional probability, independence, Bayes' rule

Karl B. Gregory

University of South Carolina

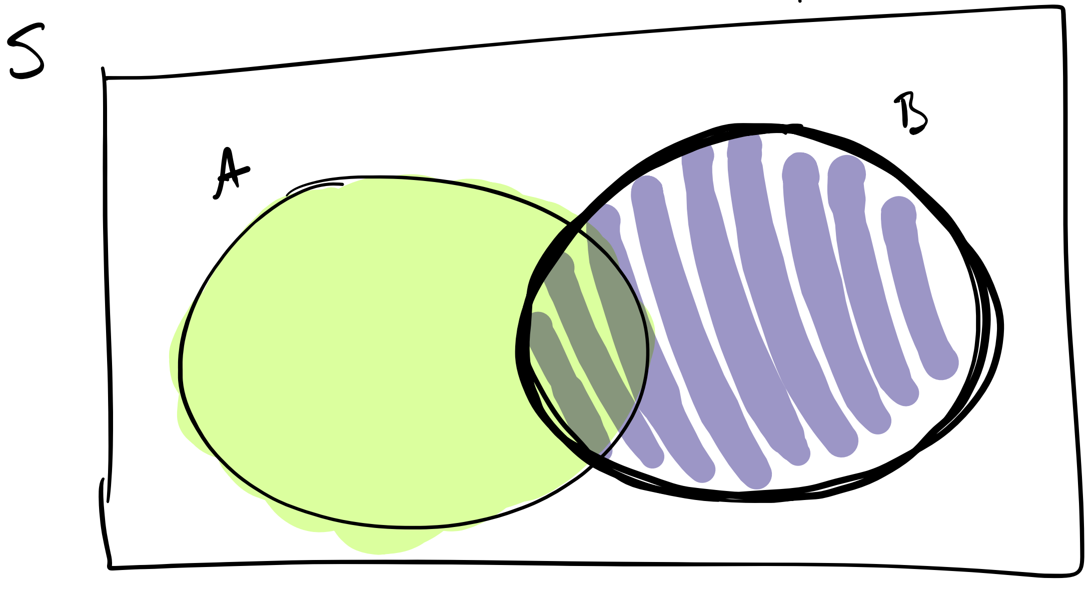
These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

$A \cap S = A$

$P(S) = 1$

$$P(A) = \frac{P(A \cap S)}{P(S)}$$

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



If we know that B happens, then we can treat B as the whole sample space.

Conditional probability

"The probability of A given B"

The *conditional probability* of A given B is

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

conditioning event

This is the probability that the event A occurs given that B occurs.

Exercise: Roll two dice. Find

1. $P($ doubles $)$
2. $P($ sum $\geq 10)$
3. $P($ doubles $|$ sum $\geq 10)$
4. $P($ sum $\geq 10|$ doubles $)$

① $P(\text{doubles}) = \frac{6}{36} = \frac{1}{6}$

② $P(\text{sum} \geq 10) = \frac{1}{6}$

$$\mathcal{S} = \left\{ \begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array} \right\}$$

③ $P(\text{Doubles} \mid \text{sum} \geq 10) = \dfrac{P(\text{Doubles} \cap \text{sum} \geq 10)}{P(\text{sum} \geq 10)}$

$= \dfrac{2/36}{1/6} = \dfrac{6}{18} = \boxed{\dfrac{1}{3}}.$

(4) $P\left(\text{sum} \geq 10 \,\middle|\, \text{Doubles} \right) = \frac{2}{6} = \frac{1}{3}.$

Exercise: From STAT 515 fa 2019:

- 40 students in class
- 10 students got an 'A' on final exam
- 12 students got an 'A' hw average
- 23 students did not get an 'A' on the final or an 'A' hw average.

If a student is drawn at random from the class, give

1. P('A' on final)
2. P('A' on final | 'A' hw average)
3. P('A' on final | less than 'A' hw average)
4. P(less than 'A' on final | less than 'A' hw average)
5. P('A' on final $\cap$ 'A' hw average)

E : Get A on Exam

H : Get A avg on Hw.

	E	E^c	total
H	5	7	12
H^c	5	23	28
total	10	30	40

① $P(E) = \frac{10}{40} = 0.25$

② $P(E \mid H) = \frac{P(E \cap H)}{P(H)} = \frac{5/40}{12/40} = \frac{5}{12} = 0.417$

③ $P(E \mid H^c) = \frac{P(E \cap H^c)}{P(H^c)} = \frac{5/40}{28/40} = \frac{5}{28} = 0.179$

④ $P(E^c \mid H^c) = \frac{23}{28} = \frac{P(E^c \cap H^c)}{P(H^c)} = \frac{23/40}{28/40} = \frac{23}{28} = 0.821$

⑤ $P(E \cap H) = \frac{5}{40}$.

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

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

Intersection prob. as conditional times unconditional

For any two events A and B,

$$P(A \cap B) = P(A|B)P(B) \quad \text{or} \quad P(A \cap B) = P(B|A)P(A).$$

Exercise: Suppose that on a safari, the probabilities of seeing a giraffe (G), a wildebeest (W), and a crocodile (C) are as follows:

$P(W) = 0.40$

$P(C) = 0.60$

$P(G) = 0.20$

$P(C|W) = 0.775$

$P(C|G) = 0.65$

$P(G \cap W) = 0.06$ ✓

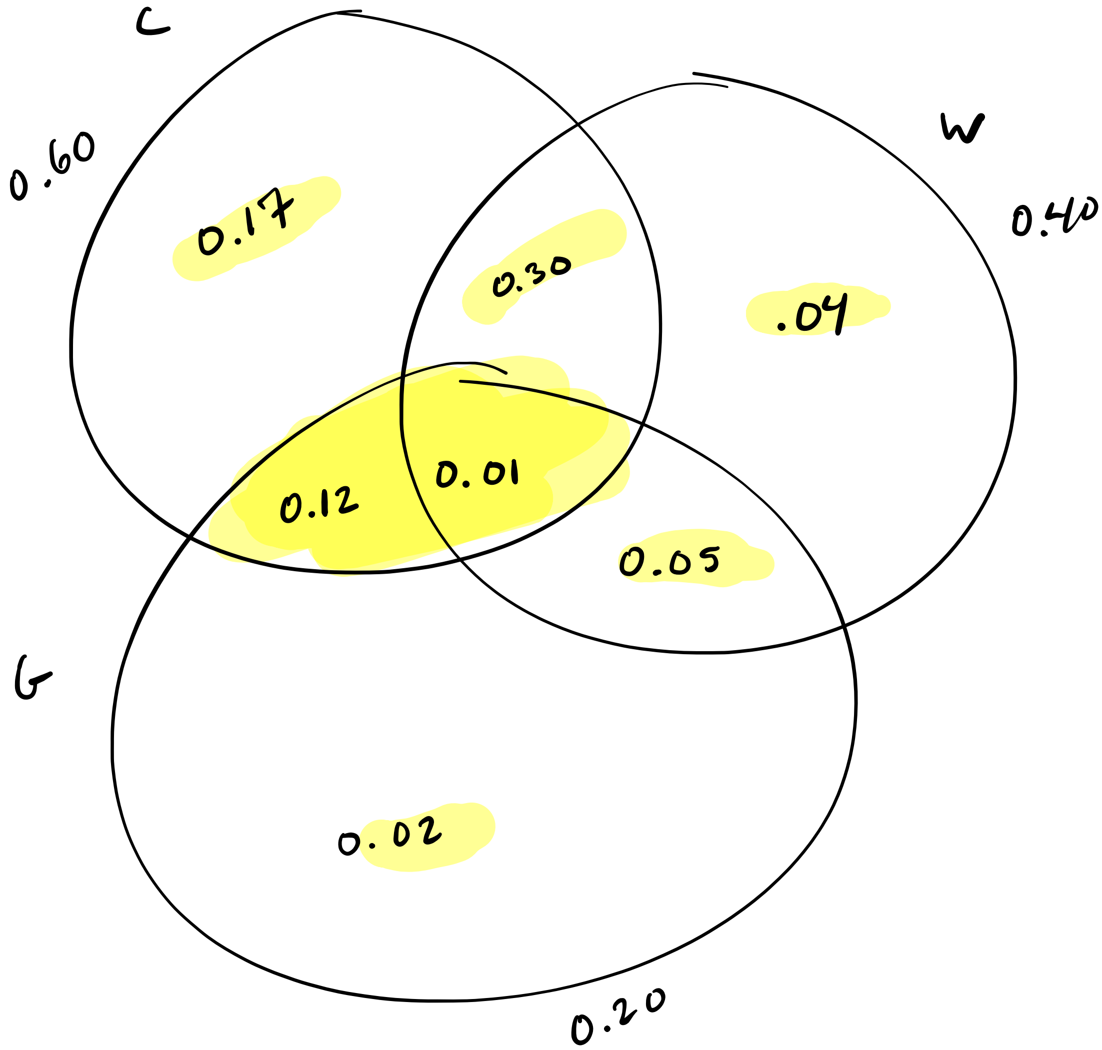
$P(G \cap W \cap C) = 0.01$ ✓

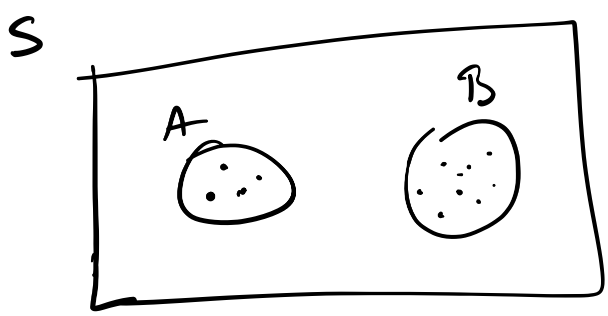
$p(C \cap W) = p(C|W)p(W) = (0.775)(0.40)$

$= 0.31$

$p(C \cap G) = p(C|G)p(G) = (0.65)(0.2) = 0.13$

Fill out a Venn diagram with the probabilities of all possibilities.





$$P(A|B) = 0$$

$$\underset{>0}{P(A)} \neq \underbrace{P(A|B)}_{=0}$$

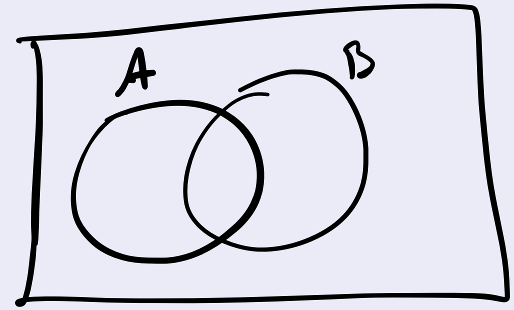
If A and B are mutually exclusive, can they be independent?
Assume $P(A) > 0$, $P(B) > 0$.

A and B are mutually exclusive if $A \cap B = \emptyset$, which means A and B can't happen at the same time. $\Rightarrow \boxed{P(A \cap B) = 0}$

Independence

Two events A and B are called *independent* if

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



Equivalent definitions of independence

The following statements are equivalent:

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

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

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

Also: If A, B independent, so are the pairs of events A, B^c and A^c, B and A^c, B^c.

Exercise: Flip a coin twice and let

$$\begin{aligned} H_1 &= \text{heads on first flip} \\ H_2 &= \text{heads on second flip} \end{aligned}$$

Find $P(H_1 \cap H_2)$ assuming that the flips are independent.

$$P(H_1 \cap H_2) = P(H_1)P(H_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Exercise: Let

- $A =$ flat tire
- $B =$ forgot spare tube

Suppose that $P(A) = 0.02$ and $P(B) = 0.10$ and $P(A \cap B) = 0.002$.

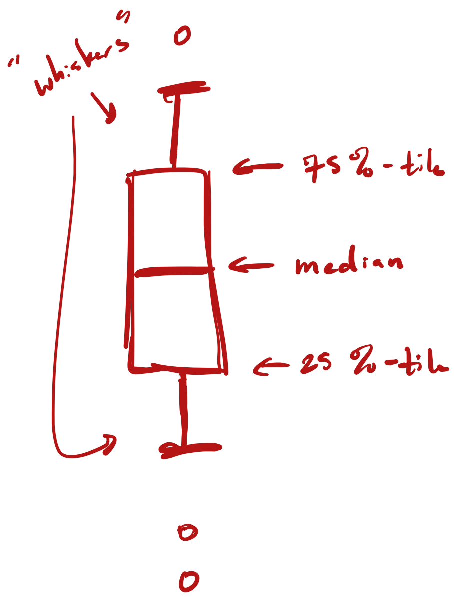
Are the events independent?

Check $\quad P(A)\,P(B) \stackrel{?}{=} P(A \cap B)$

$(0.02)\,(0.10) = 0.002.$

So yes, independent.

Boxplot



$R_1 \quad R_2 \quad \dots \quad R_{10}$

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

Exercise: Send survey to 10 people. Let R_i = person i responds, $i = 1, \dots, 10$. Assume independence with probability of response 0.20. Give

1. $P($ Everyone completes survey $)$
2. $P($ No one completes survey $)$
3. $P($ At least one person completes the survey $)$

① $P(R_1 \cap R_2 \cap R_3 \cap \dots \cap R_{10}) \overset{\text{independence}}{=} P(R_1)P(R_2)\cdot \dots \cdot P(R_{10})$

$= (0.20)^{10} = 1.024 \times 10^{-7}$

② $P(R_1^c \cap R_2^c \cap \dots \cap R_{10}^c) = P(R_1^c)\cdot P(R_2^c) \cdot \dots \cdot P(R_{10}^c) = (0.8)^{10} = 0.107$

③ $P(R_1 \cup R_2 \cup \cdots \cup R_{10}) = 1 - P((R_1 \cup \cdots \cup R_{10})^c)$

$$= 1 - P(R_1^c \cap \cdots \cap R_{10})$$

$$= 1 - (.8)^{10} = 1 - 0.107 = 0.893.$$

Exercise: From STAT 515 fa 2019:

- 40 students in class
- 10 students got an ‘A’ on final exam
- 12 students got an ‘A‘ hw average
- 23 students did not get an ‘A‘ on the final or an ‘A’ hw average.

If a student is drawn at random from the class, are the events $\underbrace{\text{‘A’ on final}}_{E}$ and $\underbrace{\text{‘A’ hw average}}_{H}$ independent?

$$P(E \cap H) \overset{?}{=} P(E)P(H) \qquad P(H) \overset{?}{=} P(H \mid E)$$

$$P(E) \overset{?}{=} P(E \mid H) \longleftarrow \text{check this one}$$

① $P(E) = \frac{10}{40} = 0.25$

not equal

② $P(E \mid H) = \frac{P(E \cap H)}{P(H)} = \frac{5/40}{12/40} = \frac{5}{12} = 0.417$

So E and H are not independent.

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

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

Bayes' Rule (simplified)

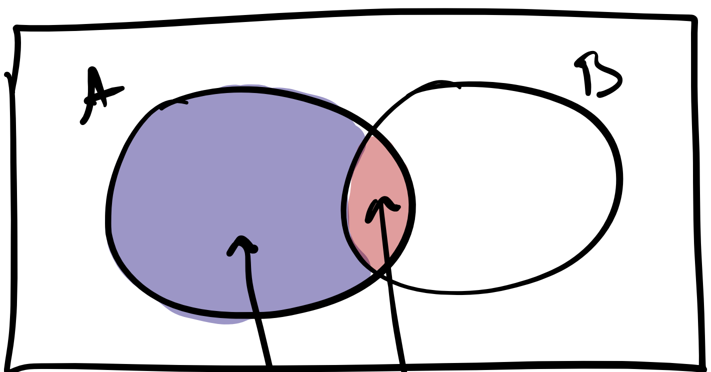
For any two events A and B such that $P(A) > 0$,

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

Bayes Rule is just another way to express this.

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

S



Show why...

$A \cap B^c$ $A \cap B$

Thomas Bayes

$$A = (A \cap B^c) \cup (A \cap B)$$

$$P(A) = P\big((A\cap B^c)\cup(A\cap B)\big)$$
$$= P(A\cap B^c) + P(A\cap B)$$
$$= P(A|B^c)P(B^c) + P(A|B)P(B)$$

If $A\cap B=\emptyset$

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

Exercise: Suppose 20% of a population has an infection. Suppose a test is positive with probability 0.70 when the infection is present and negative with probability 0.95 when the infection is absent.

I = infection

$+$ = positive test

$-$ = negative test

$P(I) = 0.20$

$P(+|I) = 0.70$

$P(-|I^c) = 0.95$

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

$P(A^c|B) = 1 - P(A|B)$

complementation works for conditional probs.

① $$P(I|+) = \frac{P(+|I)\,P(I)}{P(+|I)\,P(I) + \underbrace{P(+|I^c)}_{1-P(-|I^c)}\,P(I^c)}$$

$$= \frac{(0.70)(0.20)}{(0.70)(0.20) + (1-0.95)(1-0.20)}$$

$$= 0.78$$

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

② $$P(I^c \mid -) = \frac{P(- \mid I^c)\,P(I^c)}{P(- \mid I^c)\,P(I^c) + \underbrace{P(- \mid I)}_{=\,1 - P(+ \mid I)}\,P(I)}$$

$$= \frac{(0.95)(1 - 0.20)}{(0.95)(1 - 0.20) + (1 - 0.70)(0.20)}$$

$$= \frac{(.95)(0.80)}{(.95)(0.8) + (0.30)(0.20)}$$

$$= 0.927.$$

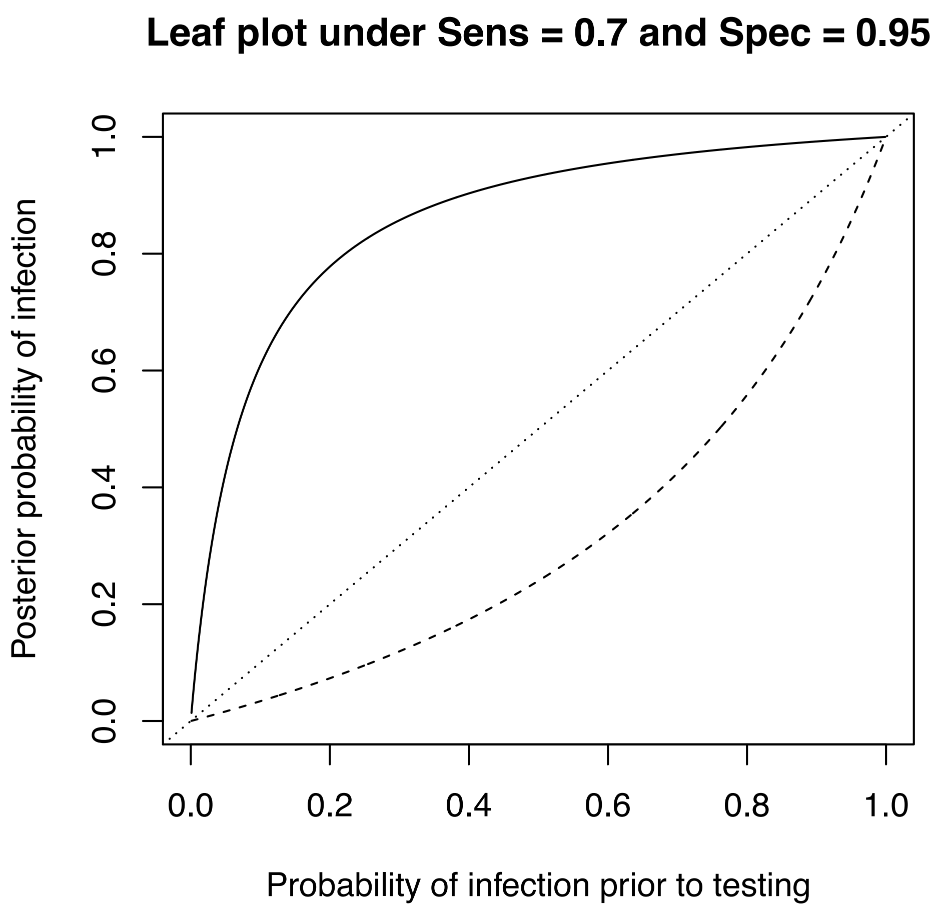
Sensitivity : $P(+ \mid I)$

Specificity : $P(- \mid I^c)$

Exercise: Suppose 20% of a population has an infection. Suppose a test is positive with probability 0.70 when the infection is present and negative with probability 0.95 when the infection is absent.

For a randomly selected person from the population, find:

1. $P($ infection $|$ positive test $)$
2. $P($ no infection $|$ negative test $)$
3. If 100 people are tested, among whom 20 have the infection, how many do you expect of
   - False positives
   - True positives
   - False negatives
   - True negatives
4. Suppose a person is tested twice, with test outcomes independent. Find
   - $P($ infection $|$ two positive tests$)$
   - $P($ infection $|$ two negative tests$)$


Leaf plot under Sens = 0.7 and Spec = 0.95
Posterior probability of infection
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Probability of infection prior to testing

Bayes' Rule

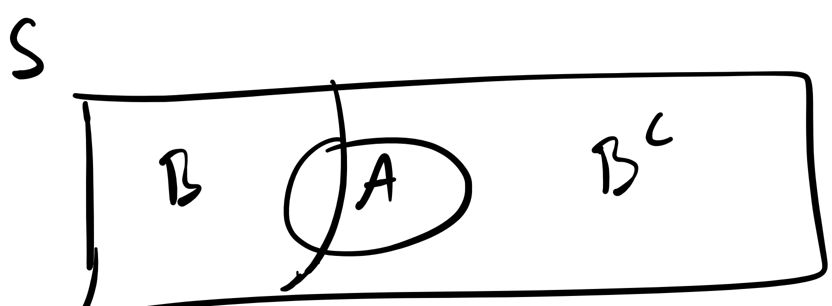
For an event A and a set of mutually exclusive events $B_1, \ldots, B_K$ such that $P(B_1) + \cdots + P(B_K) = 1$, we have

$B_1 \cup \cdots \cup B_k = S$

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{P(A|B_1)P(B_1) + \cdots + P(A|B_K)P(B_K)}.$$

Show why...

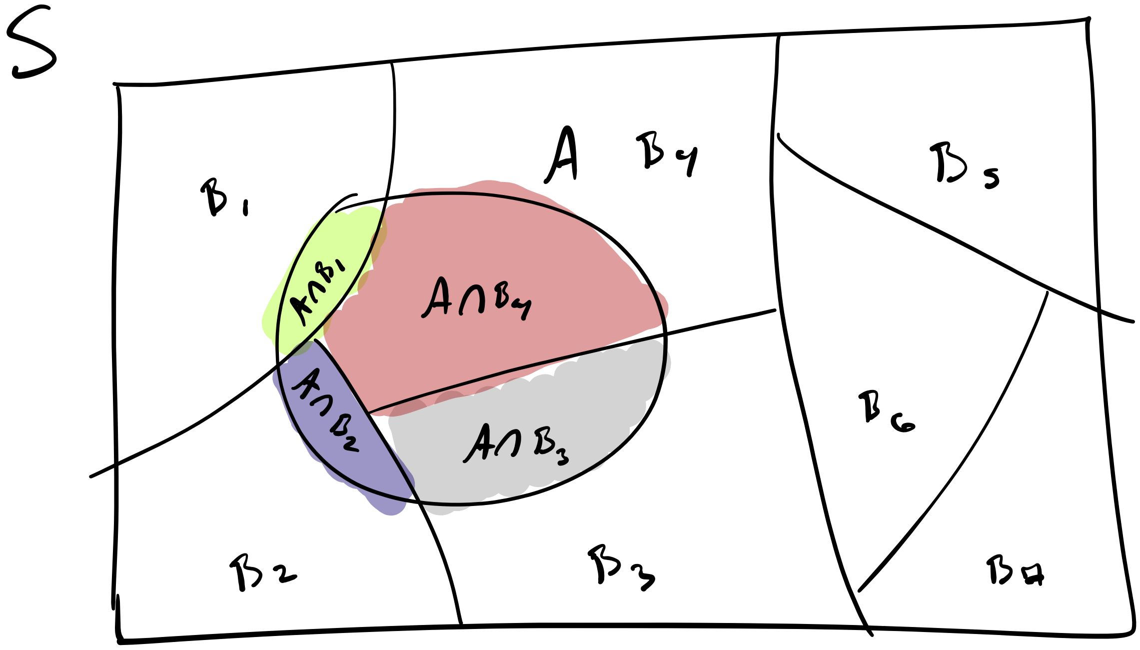
Let $K = 2$ with $B_1 = B$ and $B_2 = B^c$ to get simplified version.



Events $B_1, B_2, \ldots, B_K$ are a <u>partition</u> of S.

$$B_1 \cup B_2 \cup \cdots \cup B_K = S$$

$$B_j \cap B_k = \emptyset \quad \text{for all} \quad j \neq k$$



$$P(B_2 \mid A) = \frac{P(A \cap B_2)}{P(A)}$$

$$= \frac{P(A \mid B_2)\,P(B_2)}{\underbrace{P(A \cap B_1)} + \cdots + P(A \cap B_K)}$$

$$= \frac{P(A \mid B_2)\,P(B_2)}{P(A \mid B_1)\,P(B_1) + \cdots + P(A \mid B_K)\,P(B_K)}$$

Exercise: Roll a die and draw one bill from a bag as follows:

Roll ⚀, ⚁, ⚂	$\longrightarrow$	draw from bag 1:	3 \$5 bills and 1 \$20 bill
Roll ⚃, ⚄	$\longrightarrow$	draw from bag 2:	2 \$5 bills and 2 \$20 bills
Roll ⚅	$\longrightarrow$	draw from bag 3:	1 \$5 bill and 3 \$20 bills

1. What is the probability that you get \$20?
2. Given that you get \$20, what is the probability that you drew from bag 1?
3. If you did this 1000 times:
   - How many times would you expect to get \$20?
   - Of the times you get \$20, on how many do you expect it to be from bag 1?

$p(\$20) =$

Roll ⚀, ⚁, ⚂ $\longrightarrow$ draw from bag 1: 3 \$5 bills and 1 \$20 bill
Roll ⚃, ⚄ $\longrightarrow$ draw from bag 2: 2 \$5 bills and 2 \$20 bills
Roll ⚅ $\longrightarrow$ draw from bag 3: 1 \$5 bill and 3 \$20 bills

Draw \$20 or \$5

B_1 = draw from bag 1 (roll 1, 2, 3)
B_2 = draw from bag 2 (roll 4, 5)
B_3 = draw from bag 3 (roll 6)

$P(B_1) = \frac{1}{2}$ $\quad$ $P(B_2) = \frac{1}{3}$ $\quad$ $P(B_3) = \frac{1}{6}$

$P(\$20 \mid B_1) = \frac{1}{4}$ $\quad$ $P(\$20 \mid B_2) = \frac{1}{2}$ $\quad$ $P(\$20 \mid B_3) = \frac{3}{4}$

②

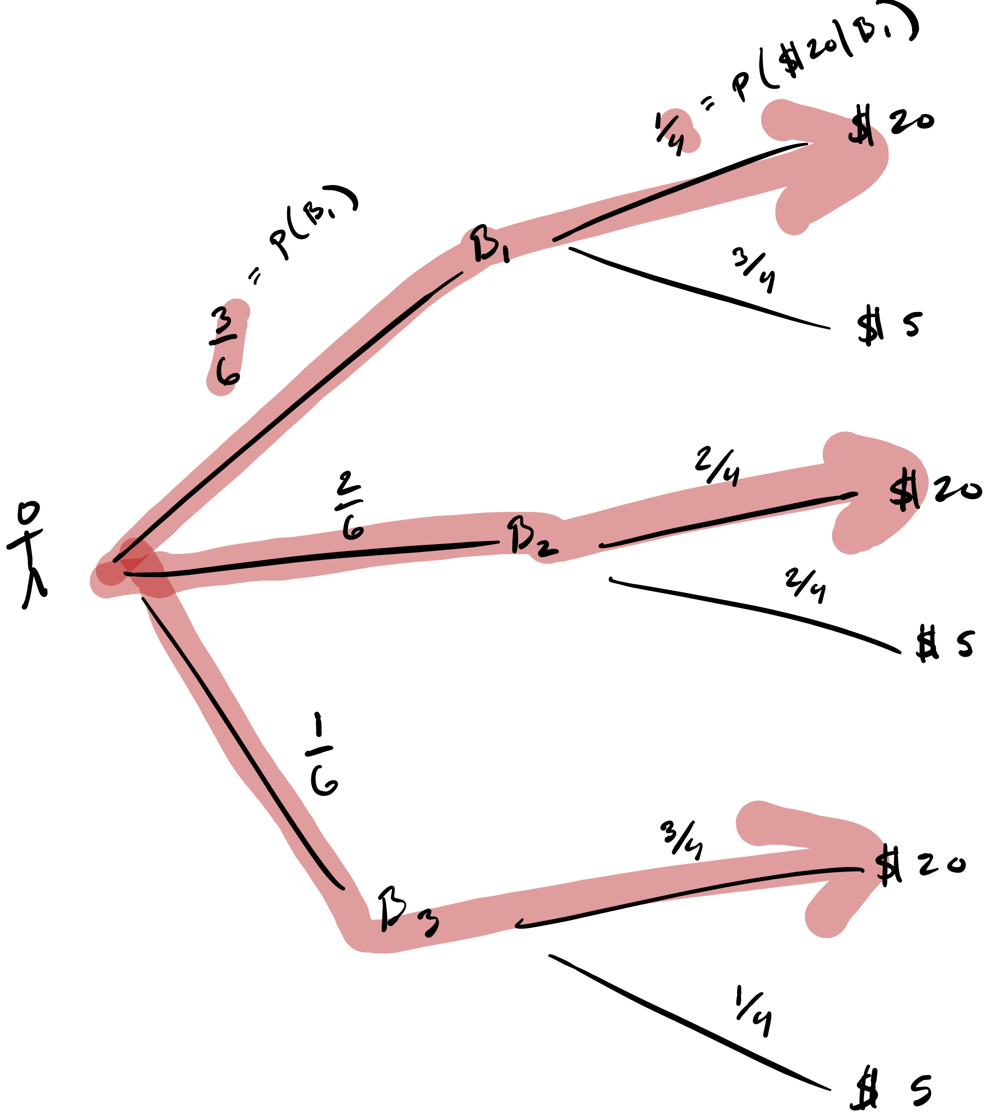
$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{P(A|B_1)P(B_1) + \cdots + P(A|B_K)P(B_K)}.$$

$$P(B_1 \mid \$20) = \frac{P(\$20 \mid B_1)\, P(B_1)}{P(\$20 \mid B_1) P(B_1) + P(\$20 \mid B_2) P(B_2) + P(\$20 \mid B_3) P(B_3)}$$

$$= \frac{\frac{1}{4} \cdot \frac{1}{2}}{\frac{1}{4} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} + \frac{3}{4} \cdot \frac{1}{6}}$$

$$= \frac{1/8}{1/8 + 1/6 + 3/24} = \frac{1/8}{\frac{3 + 4 + 3}{24}}$$

$$= \frac{3}{10}$$



$\underline{P(B_1 \cap \$20)} = \frac{3}{6} \cdot \frac{1}{4}$

$P(B_2 \cap \$20) = P(\$20 | B_2)\, P(B_2)$

$= \frac{2}{6} \cdot \frac{2}{4} = \frac{1}{6}$

$P(B_3 \cap \$20) = \frac{1}{6} \cdot \frac{3}{4}$

$$P(\$20) = P(\$20 \cap B_1) + P(\$20 \cap B_2) + P(\$20 \cap B_3)$$
$$= \frac{3}{6} \cdot \frac{1}{4} + \frac{2}{6} \cdot \frac{2}{4} + \frac{1}{6} \cdot \frac{3}{4}$$
$$= \frac{3 + 4 + 3}{24}$$
$$= \frac{10}{24}$$
$$= \frac{5}{12}.$$

Go to tree!

$$P(B_1 | \$20) = \frac{P(B_1 \cap \$20)}{P(\$20)} = \frac{\frac{3}{6} \cdot \frac{1}{4}}{\frac{5}{12}} = \frac{3}{10}.$$

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

4. a) (i) $P(2 \text{ passenger cars}) = \dfrac{\# \text{ ways to choose } 2 \text{ passenger cars}}{\# \text{ ways to choose } 2 \text{ from } 9 \text{ cars}} = \dfrac{\binom{4}{2}}{\binom{9}{2}}$

(ii) $P(1 \text{ luggage}, 1 \text{ passenger}) = \dfrac{\binom{2}{1}\binom{4}{1}}{\binom{9}{2}}$

(iii) $P(\text{At least 1 cattle}) = 1 - P(\text{No cattle})$

$= 1 - \dfrac{\binom{6}{2}}{\begin{matrix}9\\2\end{matrix}}$

$(\quad)$

4. A train to the western frontier will consist of 4 passenger cars, 3 cattle cars, and 2 luggage cars, which are to be put in order at random.

   (a) Bandits plan to mount the train and enter the two rearmost cars. Find the probability that they enter

   Bandits select 2 cars at random

      i. two passenger cars.

      ii. a luggage car and a passenger car.

      iii. at least one cattle car.

   5

   1 2 3 4 1 2 3 4 6

   (b) A lady's shawl flies out a window of the foremost passenger car, reenters a window in the rearmost passenger car, and is seized by a gentleman who gallantly vows to return it while the train is in motion. With what probability can he make his way from the rearmost passenger car to the foremost, passing only through passenger cars?

   $$\frac{6! \cdot 4!}{9!}$$

   (c) There are 14 head of cattle to be transported in the three cattle cars.

      i. In how many ways can 5, 5, and 4 head of cattle, respectively, be put into the three cattle cars?

      ii. The gallant gentleman owns 3 of the 14 head of cattle. If 5, 5, and 4 of the 14 head of cattle are put into the three cattle cars at random, with what probability will the gallant gentleman's cattle all be placed in the same cattle car?

5. Suppose there are 5 bowling balls which are identical except that one is magical and delivers, no matter what, a strike with probability 3/4. Suppose you get a strike 1 out of 4 times on average when using non-magical bowling balls. You select one of the 5 balls at random and send it down the lane...

(a) Give the probability that you get a strike.

(b) Given that you got a strike, what is the probability you chose the magic bowling ball?

(c) Suppose you choose a ball and with the same ball you get two strikes in a row. What is the probability that you chose the magic ball?

M = Magic ball

S = strike

$P(M) = \frac{1}{5}$

$P(S \mid M) = \frac{3}{4}$

$P(S \mid M^c) = \frac{1}{4}$

(b) $$P(M \mid S) = \frac{P(S \mid M)\,P(M)}{P(S \mid M)\,P(M) + P(S \mid M^c)\,P(M^c)}$$

$$= \frac{(3/4)\,1/5}{(3/4)\,1/5 + 1/4 \cdot 4/5}$$

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

(c) S_1 = strike

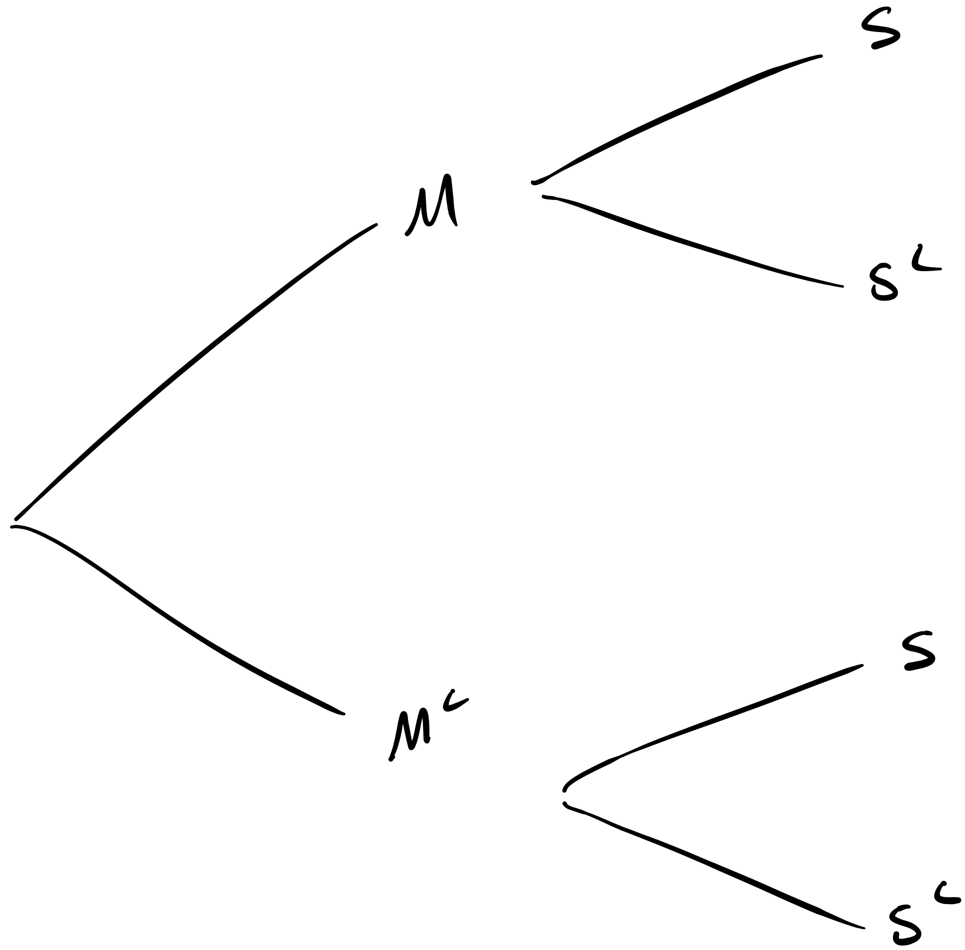
S_2 = another strike

$P(M) = \frac{1}{5}$

$P(S_1 \cap S_2 \mid M) = P(S_1 \mid M) \cdot P(S_2 \mid M)$

$P(S_1 \cap S_2 \mid M^c) = P(S_1 \mid M^c) \cdot P(S_2 \mid M^c)$

$P(M \mid S_1 \cap S_2) =$


M
S
S c
M c
S
S c