CS1800 Day 12

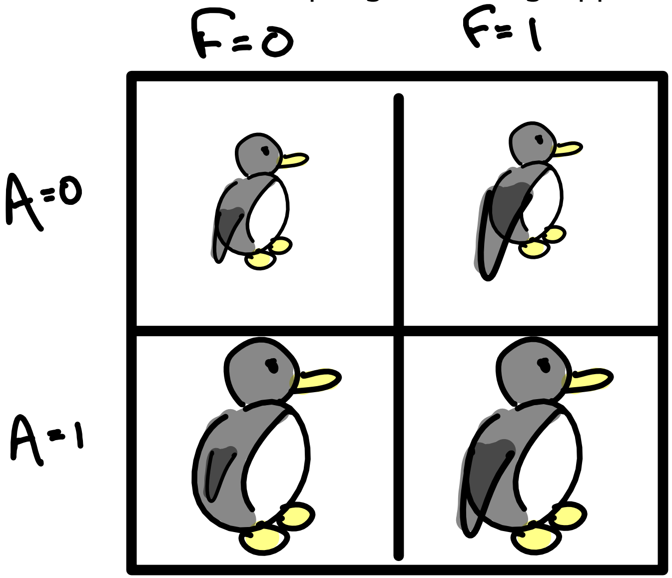
Admin:
- exam / HW4 results
- grade estimates by this Friday (likely sooner)
- tuning up your study process in CS1800

Content:

Joint Probability Distribution
Marginalization
Conditional Probability
Bayes Rule
Independence

Joint Probability Distribution: A distribution over more than 1 variable at a time

Let A=1 indicate if a penguin is an adult (0 otherwise)
Let F=1 indicate if a penguin has big flippers (0 otherwise)



	F=0	F=1
A=0	3/12	2/12
A=1	1/12	6/12

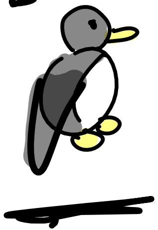
HALF OF PENGUINS ARE ADULTS w/ BIG FLIPPERS

NOTATION

$P(A=0\ F=1)$ IS PROB OF BOTH OUTCOMES OCCURING AT SAME TIME

$A=0$ (NOT ADULT)

$F=1$ (BIG FLIPPER)



$A=0$

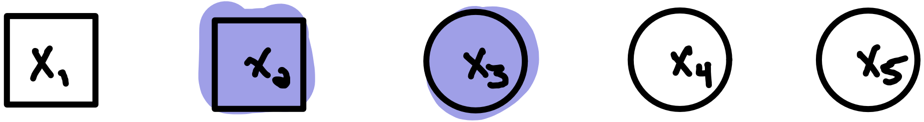
$F=1$

Marginalizing (removing a random variable from a joint distribution)

Consider selecting one of the following 5 shapes, with equal probability:

Let B=1 be the event that the shape is blue

Let C=1 be the event that the shape is a circle



	B=0	B=1
C=0	1/5	1/5
C=1	2/5	1/5

WHAT IS $P(B=1) = P(B=1\ C=0) + P(B=1\ C=1)$

$$= 1/5 + 1/5$$

$$= 2/5$$

Remember: To compute P(B) we can sum P(B, A) for all outcomes in sample space of A

$$P(B=b) = \sum_a P(B=b, A=a)$$

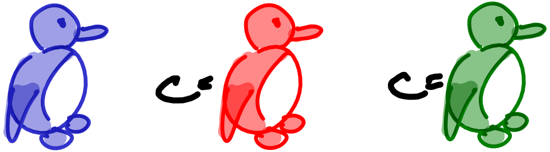
In Class Activity

Let C be a random variable representing penguin color (sample space: blue, red or green)
Let A=1 indicate if a penguin is an adult (0 otherwise)

Given the following distribution of A, C

Compute each of the follow probabilities:

- P(C=blue)
- P(C=red) + P(C=green)
  (how is this related to prob above?)
- P(A=1)



	C=	C=	C=
A=0	1/12	3/12	0/12
A=1	2/12	1/12	5/12

In Class Activity

adding over all outcomes of A

- P(C=blue) $= \sum_a P(C=\text{BLUE}, A=a)$

$$= P(C=\text{BLUE}, A=0) + P(C=\text{BLUE}, A=1)$$
$$= 1/12 + 2/12$$
$$= 1/4$$

- P(C=red) + P(C=green)

  (how is this related to prob above?)

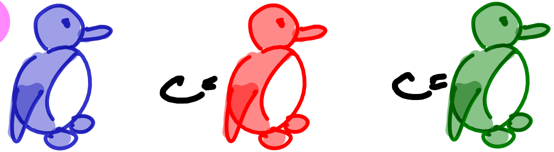
$$P(C=r) = P(C=r\ A=0) + P(C=r\ A=1)$$
$$= 3/12 + 1/12$$

$$P(C=g) = P(C=g\ A=0) + P(C=g\ A=1)$$
$$= 0/12 + 5/12$$
$$= 5/12$$

$$P(C=\text{RED OR GREEN}) = 9/12$$

- P(A=1)

$$P(A=1) = P(A=1\ C=b) + P(A=1\ C=r) + P(A=1\ C=g)$$
$$= 2/12 + 1/12 + 5/12 = 8/12$$



	C=	C=	C=
A=0	1/12	3/12	0/12
A=1	2/12	1/12	5/12

Conditional Probability (intuition & motivation)

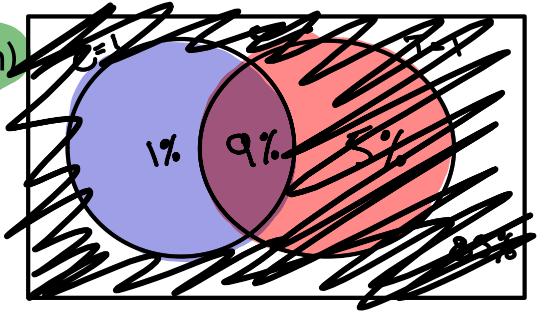
$C=1$ indicates a person has covid ($C=0$ otherwise)
$T=1$ indicates a person has positive test ($T=0$ otherwise)

Let us discuss (and express) the following probabilities:

- probability person has a positive test $P(T=1)$
- probability person has positive test given they have covid
  within the group that has covid, what fraction is testing positive? $P(T=1 \mid C=1)$
- probability person has covid given a positive test
  with the group that has a positive test, how many have covid? $P(C=1 \mid T=1)$

Intuition:
Conditional probability $P(X=x|Y=y)$ is the probability of event $X=x$ if we constrain ourselves to a world where $Y=y$.

Conditional Probability (motivating our formula from intuition)



C=1 indicates a person has covid (C=0 otherwise)
T=1 indicates a person has positive test (T=0 otherwise)

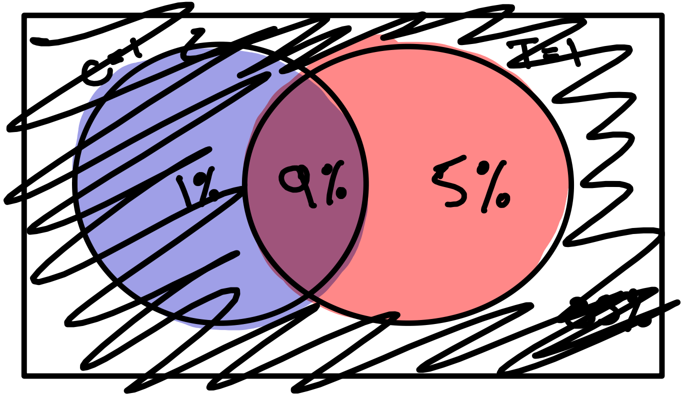
Let us discuss (and express) the following probabilities:

- probability person has a positive test $P(T=1) = 9+5 = 14\%$

- probability person has positive test given they have covid $P(T=1 \mid C=1) = \frac{9}{1+9} = 90\%$

- probability person has covid given a positive test

Intuition:
Conditional probability P(X=x|Y=y) is the probability of event X=x if we constrain ourselves to a world where Y=y.



COMPUTE

$$P(C=1 \mid T=1)$$

$$P(C=1 \mid T=1) = \frac{9}{9+5} = \frac{P(C=1\ T=1)}{P(T=1)}$$

Conditional Probability (Formula version1: from our intuition)

$$P(a|b) = \frac{P(ab)}{P(b)}$$

PROB a HAPPENS GIVEN CONDITION b

PROB a b HAPPEN TOGETHER

PROB b HAPPENS

In Class Activity

Compute each of the probabilities from the distribution below. For each, write a sentence explaining its meaning which is easily understood by a non-technical reader

Let S be a twitter sentiment score about bitcoin (1=good, 0=neutral, -1=bad)
Let B be the movement of bitcoin price (1=up, -1=down)

P(S=-1, B=1)
P(S=-1 | B=1)
P(B=1 | S=-1)
P(B=1)

	S=-1	S=0	S=1
B=-1	19%	27%	5%
B=1	8%	21%	20%

P(S=-1, B=1) $= 8\%$

P(S=-1 | B=1) $= \dfrac{P(S=-1, B=1)}{P(B=1)}$

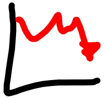
$= 8/(8+21+20) = 16.3\%$

P(B=1 | S=-1)

$= \dfrac{P(B=1, S=-1)}{P(S=-1)} = \dfrac{8}{19+8} = 29.6\%$

P(B=1)

$= 8+21+20 = 49\%$



	S=-1	S=0	S=1
B=-1	19%	27%	5%
B=1	8%	21%	20%

Conditional Probability (formula version 2: often more useful in our algebraic manipulations)

$$P(a|b) = \frac{P(a\,b)}{P(b)} \quad \longleftrightarrow \quad P(a|b)\,P(b) = P(a\,b)$$

Takeaway above:

Multiplying
- a conditional probability
- the probability of condition

Will yield
- prob both outcomes happen together

BAYES RULE (GLORIFIED CONDITIONAL PROBABILITY)

SEE PREVIOUS SLIDE

$$P(a|b)P(b) = P(ab) = P(b|a)P(a)$$

$$\Rightarrow \quad P(a|b) = \frac{P(b|a)P(a)}{P(b)}$$

Notice: this formula "swaps" the order of the conditioning: P(A|B) on left P(B|A) on right
Its typical in a Bayes question to be given variables in one order while question asks for other.

A Helpful Manipulation

Marginalization

$$P(b) = \sum_a P(a\,b)$$

Conditional Prob Definition

$$= \sum_a P(b|a)\,P(a)$$

WHY WAS THAT HELPFUL?

BAYES RULE 1

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)}$$

BAYES RULE 2

$$P(a|b) = \frac{P(b|a)\,P(a)}{\sum_i P(b|a_i)P(a_i)}$$

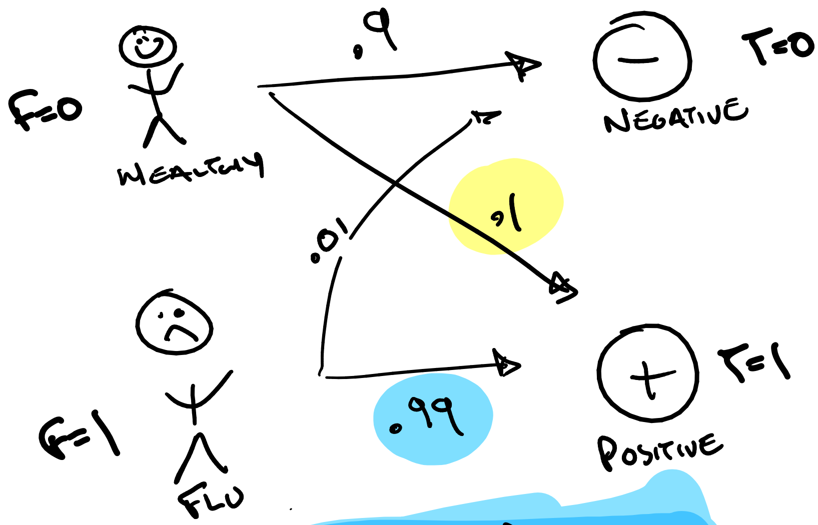
Notice:
all terms of form P(b|a) and P(a) here

Bayes Rule Ex

$P(F=1)=.04$

$P(F=0)=1-.04$

Given flu occurs in 4% of population, what is the probability one has flu given they test positive?



$$P(F=1|T=1)=\frac{P(T=1|F=1)P(F=1)}{P(T=1)}$$

$$=\frac{P(T=1|F=1)P(F=1)}{P(T=1\,F=0)+P(T=1\,F=1)}$$

$$=\frac{P(T=1|F=1)P(F=1)}{P(T=1|F=0)P(F=0)+P(T=1|F=1)P(F=1)}$$

$P(T=1|F=1)=.99$

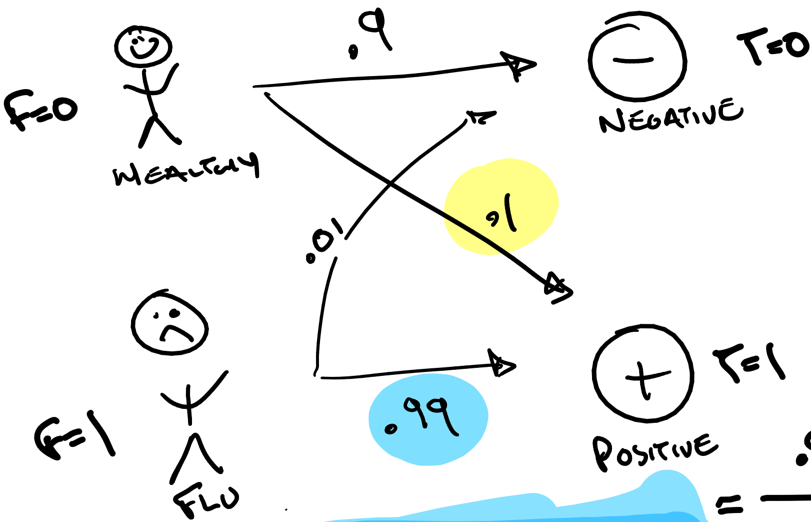
$P(T|F)$

Bayes Rule Ex

$P(F=1)=.04$

Given flu occurs in 4% of population, what is the probability one has flu given they test positive?

$P(F=0)=1-.04$



$$P(F=1|T=1)=\frac{P(T=1|F=1)P(F=1)}{P(T=1)}$$

$$=\frac{P(T=1|F=1)P(F=1)}{P(T=1|F=0)+P(T=1|F=1)}$$

$$=\frac{P(T=1|F=1)P(F=1)}{P(T=1|F=0)P(F=0)+P(T=1|F=1)P(F=1)}$$

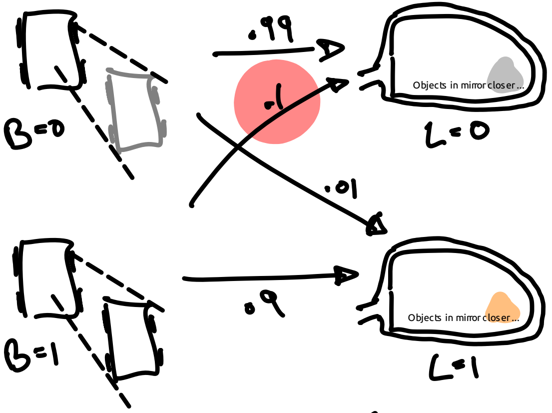
$$P(T=1|F=1)=.99$$

$$=\frac{.99\cdot .04}{.1\cdot .96 \quad .99\cdot .04}$$

$$\approx 29.4\%$$

In Class Assignment

A blind spot monitor produces a warning light (L=1) when it estimates that a car is in one's blind spot (B=1). Given that the light is off, whats the probability that a car is one's blind spot? (Assume that a car is in your blindspot 2 percent of the time while driving.)



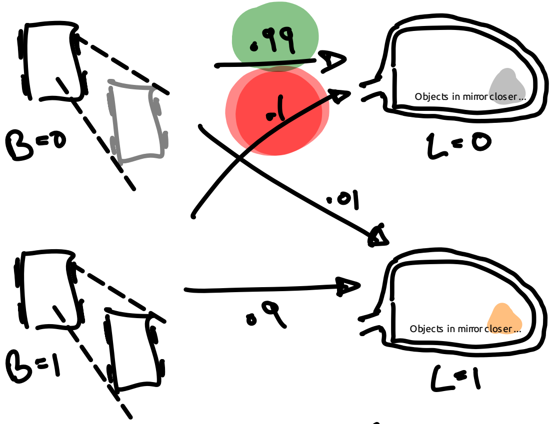
$$P(L|B)$$

$$P(B=1|L=0) = \frac{P(L=0|B=1)\,P(B=1)}{P(L=0)}$$

$$= \frac{(.1)(.02)}{.9722} \approx .002$$

In Class Assignment

A blind spot monitor produces a warning light (L=1) when it estimates that a car is in one's blind spot (B=1). Given that the light is off, whats the probability that a car is one's blind spot? (Assume that a car is in your blindspot 2 percent of the time while driving.)



$$P(L|B)$$

$$P(L=0) = P(L=0\ B=0) + P(L=0\ B=1)$$

$$= P(L=0|B=0)P(B=0) + P(L=0|B=1)P(B=1)$$

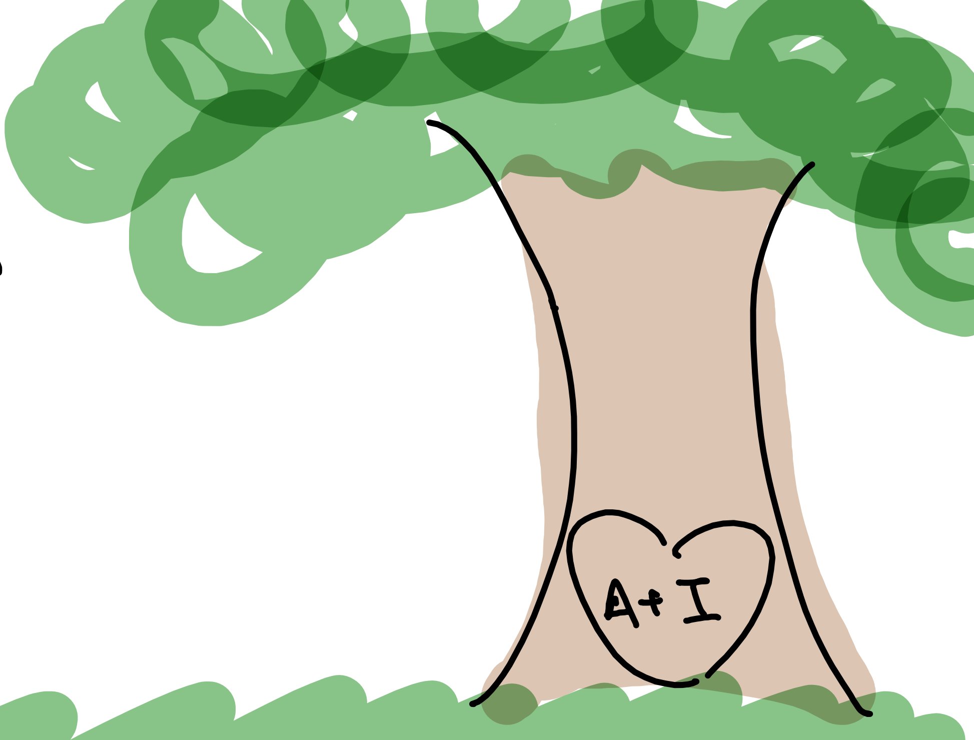
$$= (.99)(1-.02) + (.1)(.02)$$

$$= .9722$$

ALGEBRA + INTUITION

LOVE STORY

AHEAD...



INDEPENDENCE + CONDITIONAL PROB

INDEPENDENCE

INTUITION:

Random variables x, y are independent
if observing any outcome of one doesn't
impact our beliefs about the other.

ALGEBRA:

FOR EACH OUTCOME PAIR x, y

$$P(X=x, Y=y) = P(X=x)\,P(Y=y)$$

Bayes Rule shows the equivilence of the algebraic and intuitive definitions above!

INDEPENDENCE + CONDITIONAL PROB

INDEPENDENCE

INTUITION:

Random variables x, y are independent if observing any outcome of one doesn't impact our beliefs about the other.

ALGEBRA:

FOR EACH OUTCOME PAIR x, y

$$P(X=x\ Y=y) = P(X=x)\,P(Y=y)$$

$$P(X|Y) = \frac{P(X\,Y)}{P(Y)} = \frac{P(X)P(Y)}{P(Y)} = P(X)$$

Notice that P(X|Y) = P(X). Observing Y has no impact on the prob of X!