Agenda

Professor Hamlin
Day 13

1. Admin
2. Joint Probability Dist.
3. Marginalization
4. Conditional Probability
5. Bayes Rule
6. Independence

Review

Probability - Experiment, outcomes, sample space distribution, random variable

Expected Value : "average value"

$$E[X] = \sum_{x \in S}' x \cdot Pr[X=x]$$

Variance: how much things vary

$$Var(X) = E[X^2] - (E[X])^2$$

1. What is the probability of rolling ≤ 2 on the die w/ following #'s 1, 1, 1, 2, 6, 6 → Event

$$Pr[X \leq 2] = \frac{4}{6}$$

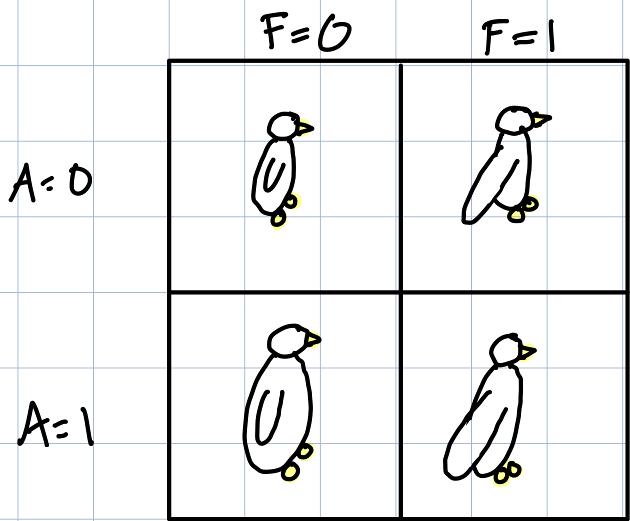
2. What is Expected Value of the die?

$$E[X] = 1 \cdot \tfrac{1}{2} + 2 \cdot \tfrac{1}{6} + 6 \cdot \tfrac{1}{3}$$

$$\boxed{2 \tfrac{5}{6}}$$

Joint Probability: a dist. over more than 1 Random variable at a time

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



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

adds up to 1

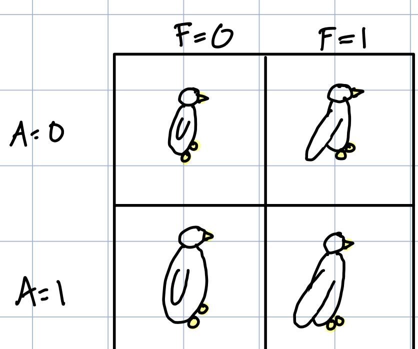
Adult penguins w/ large fins

events happen at same time

$\Pr[A=0, F=1]$ is how we express it w/ math notation

$\Pr[A=0, F=0] = 3/12$

Marginalization: removing a random variable from probability dist. (e.g. what fraction of penguins are adults)



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

this row of all adult penguins

$\Pr[A=1] = 7/12$

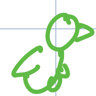
$$Pr[A=1, F=0] + Pr[A=1, F=1]$$

$$\frac{1}{12} + \frac{6}{12} = \boxed{7/12}$$

To compute $Pr[A=a]$ sum up $Pr[A=a, B=?]$ for all outcomes in sample space of B

$$Pr[A=a] = \sum_{b \in S} Pr[A=a, B=b]$$

<u>Exercise</u> C = color of penguin (red, blue, green)
A = penguin is adult (1) or 0 otherwise

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

1) $Pr[C = blue]$

$Pr[C=B, A=0] + Pr[C=b, A=1]$

$3/12 + 1/12 = \boxed{4/12}$

2) $P[C=red] + Pr[C=green]$

$1/12 + 2/12 + 0/12 + 5/12$

$\boxed{8/12}$

3) $Pr[A=1]$

$2/12 + 1/12 + 5/12 = \frac{8}{12}$

$1 - Pr[C = blue]$

Conditional Probability: "If x then the probability of y is?"

$C=1$ indicate if person has covid
$T=1$ indicates test is positive

1) What is prob person has positive test

$Pr[T=1]$

2) If person has covid then what is prob of positive test?

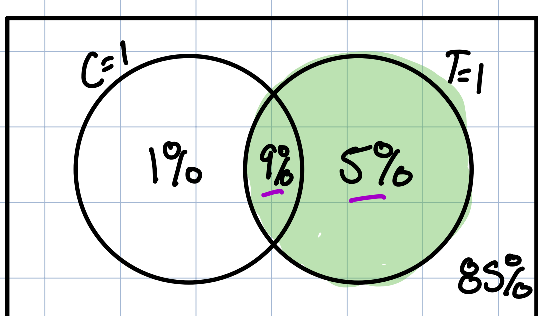
$Pr[T=1 \mid C=1]$ → This is read as "Given C=1"

3) Prob person has covid given positive test?
"if pos test then prob of covid is?

$Pr[C=1 \mid T=1]$

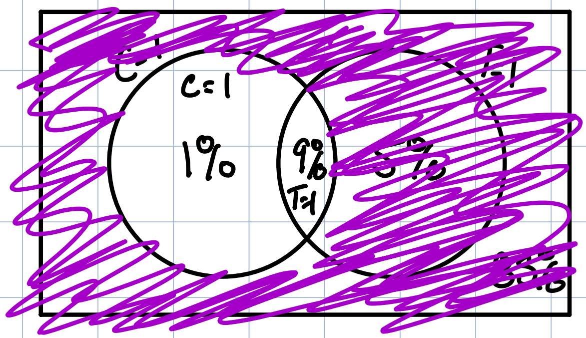
So how do we calculate it? Let's talk intiution

1) What is prob person has positive test



$Pr[T=1]$
$.09 + .05 = 14\%$

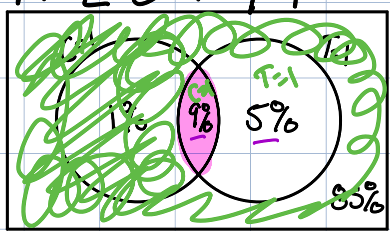
2) If person has covid then what is prob of positive test? $\Pr[T=1 \mid \underline{C=1}]$



$$= \frac{.09}{\underbrace{.01+.09}_{\Pr[C=1]}} = 90\%$$

3) Prob person has covid given positive test?

$\Pr[C=1 \mid T=1]$



$$\frac{9\%}{5\% + 9\%} = \boxed{\frac{9}{14}}$$

Conditional $\Pr[X=x \mid Y=y]$ is prob of $\Pr[X=x]$ when we constrain ourselves to the world of $Y=y$

Formally:

Prob x & y happen together

$$\Pr[X=x \mid Y=y] = \frac{\Pr[X=x, Y=y]}{\Pr[Y=y]}$$

Prob x happens given y happening

Prob y happens

Exercise! S: twitter sentiment score (1 = good, 0 = neutral, −1 = bad)

B: Bitcoin price (1 = up, −1 = down)

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

$\Pr[B=-1, S=1]$

1) Compute $\Pr[S=-1 \mid B=1]$ and $\Pr[S=-1]$ and explain # in english.

$$\Pr[S=-1 \mid B=1] = \frac{\Pr[S=-1, B=1]}{\Pr[B=1]} = \frac{8}{8+21+20} = \boxed{16\%}$$

$$\Pr[S=-1] = 19+8 = \boxed{27\%}$$

Given improved bitcoin prices reduces chance of negative sentiment

2) Compute $\Pr[B=1 \mid S=-1]$ and $\Pr[B=1]$ and explain # in english.

$$\Pr[B=1] = 8+21+20 = \boxed{49\%}$$

$$\Pr[B=1 \mid S=-1] = \frac{\Pr[B=1, S=-1]}{\Pr[S=-1]} = \frac{8\%}{27\%} \approx \boxed{30\%}$$

It less likely for bitcoin prices to go up when negative sentiment exists

<u>Note</u> we can manipulate the conditional prob. expression

$$\Pr[A=a \mid B=b] = \frac{\Pr[A=a, B=b]}{\Pr[B=b]}$$

~or~

$$\Pr[A=a, B=b] = \Pr[A=a \mid B=b] \cdot \Pr[B=b]$$

Multiplying conditional prob w/ the probability of condition yields prob both outcomes happen together

<u>Bayes Rule</u>: if given $\Pr[A=a \mid B=b]$ how to get $\Pr[B=b \mid A=a]$?

$$\Pr[A=a \mid B=b] \cdot \Pr[B=b] = \Pr[A=a, B=b]$$

and

$$\Pr[B=b \mid A=a] \cdot \Pr[A=a] = \Pr[A=a, B=b]$$

$$\Rightarrow \Pr[A=a \mid B=b] \cdot \Pr[B=b] = \Pr[B=b \mid A=a] \cdot \Pr[A=a]$$

$$\Rightarrow \boxed{\Pr[A=a \mid B=b] = \frac{\Pr[B=b \mid A=a] \cdot \Pr[A=a]}{\Pr[B=b]}}$$

Bayes Rule

e.g. if given variables in one order find them in another

Helpful Note:

Marginalization

$$\Pr[B=b] = \sum_{a \in S} \Pr[A=a, B=b]$$

cond. prob definition

$$= \sum_{a \in S} \Pr[B=b \mid A=a] \cdot \Pr[A=a]$$

Why is this helpful?

$$\Pr[A=a \mid B=b] = \frac{\Pr[B=b \mid A=a] \cdot \Pr[A=a]}{\Pr[B=b]}$$

we now can calculate this!

$$\boxed{\Pr[A=a \mid B=b] = \frac{\Pr[B=b \mid A=a] \cdot \Pr[A=a]}{\sum_{x \in S} \Pr[B=b \mid A=x] \cdot \Pr[A=x]}}$$

Another form of Bayes Rule

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

$F=0$  (healthy

$T=0$ $-$ negative

$\Pr[T=0 \mid F=0] = .9$

$\Pr[T=0 \mid F=1] = .01$

$F=1$  Flu

$T=1$ $+$ positive

$\Pr[T=1 \mid F=0] = .1$

$\Pr[T=1 \mid F=1] = .99$

$$\Pr[F=1] = .04 \Rightarrow \Pr[F=0] = .96$$

Asking $\Pr[F=1 \mid T=1] = ?$

$$\Pr[F=1 \mid T=1] = \frac{\Pr[T=1 \mid F=1] \cdot \Pr[F=1]}{\Pr[T=1]}$$

$$\Pr[T=1] = \Pr[T=1 \mid F=0] \cdot \Pr[F=0] + \Pr[T=1 \mid F=1] \cdot \Pr[F=1]$$

$$.1 \cdot .96 + .99 \cdot .04$$

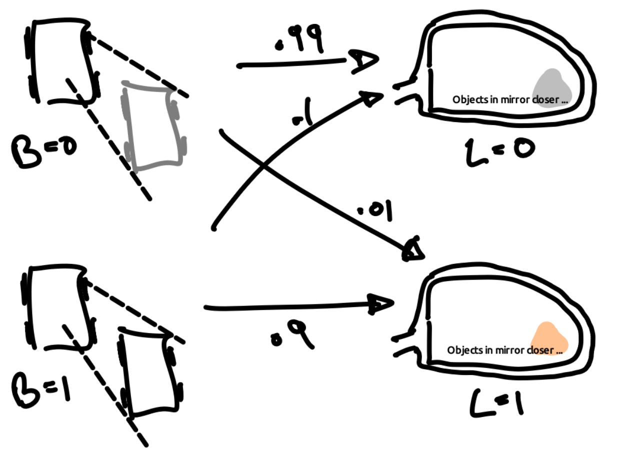
$$= .1356$$

$$= \frac{.99 \cdot .04}{.1356} \cong \boxed{29\%}$$

Exercise

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.)

$\rightarrow \Pr[B=1] = .02 \quad \Pr[B=0] = .98$



$$\Pr[L=0 \mid B=0] = .99$$
$$\Pr[L=0 \mid B=1] = .1$$

$$\Pr[L=1 \mid B=0] = .01$$
$$\Pr[L=1 \mid B=1] = .9$$

$$\Pr[B=1 \mid L=0] = \frac{\Pr[L=0 \mid B=1] \cdot \Pr[B=1]}{\Pr[L=0]}$$

$$\Pr[L=0] = \Pr[L=0 \mid B=0] \cdot \Pr[B=0] + \Pr[L=0 \mid B=1] \cdot \Pr[B=1]$$

$$= .99 \cdot .98 + .1 \cdot .02$$

$$\frac{.1 \cdot .02}{.99 \cdot .98 + .1 \cdot .02} \cong .00205$$

Independence

Intution: if RV X & Y are independent if observing any outcome of one does not impact the outcome of the other

Math: $Pr[X=x, Y=y] = Pr[X=x] \cdot Pr[Y=y]$

~and~

$Pr[X=x \mid Y=y] = P[X=x]$

Example

$Pr[X=1] = 1/4$, 0 otherwise

$Pr[Y=1] = 1/10$, 0 otherwise

	$X=0$	$X=1$
$Y=0$	$\frac{3}{4} \cdot \frac{9}{10}$	$\frac{1}{4} \cdot \frac{9}{10}$
$Y=1$	$\frac{3}{4} \cdot \frac{1}{10}$	$\frac{1}{4} \cdot \frac{1}{10}$

Exercise) Two unfair coins $Pr[C1=H] = 3/4$

$Pr[C2=H] = 2/3$

What is the probability of getting...

... $Pr[C1=H, C2=T] = 3/4 \cdot 1/3 = 1/4$

... $Pr[C1=T, C2=T] = 1/4 \cdot 1/3 = 1/12$