Section 3.1: Conditional Probability

If an event with $P(A)>0$ occurs, this reduces the sample space from Ω to A, and the prob. that B will occur given that A has occurred is

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

Ex: 2

Candidates a, b, & c are running for one seat on the city council. Suppose a and b have an equal chance of being elected, but c is half as likely as a to be elected.

Let A = 'a wins', B = 'b wins', C = 'c wins'

$$\Rightarrow P(A)+P(B)+P(C) = 1 = P(A)+P(A)+\tfrac{1}{2}P(A) \Rightarrow P(A)=\tfrac{2}{5}$$

Suppose A drops out. Then

$$P(B|A^c) = \frac{P(B\cap A^c)}{P(A^c)} = \frac{P(B)}{P(B\cup C)} = \frac{2/5}{3/5} = \frac{2}{3}$$

$$P(C|A^c) = \frac{P(C)}{P(B\cup C)} = \frac{1/5}{3/5} = \frac{1}{3}$$

Assuming a former supporter of a is equally as likely to support b as c. If a supporters more likely to choose c, then $P(C|A^c) > \frac{1}{3}$.

Ex 1:

Roll a pair die once. Let X be the outcome, and let

F = '$X=6$' E = '$X>4$'

Then $P(F) = 1/6$

Suppose the die is rolled once and we are told event E

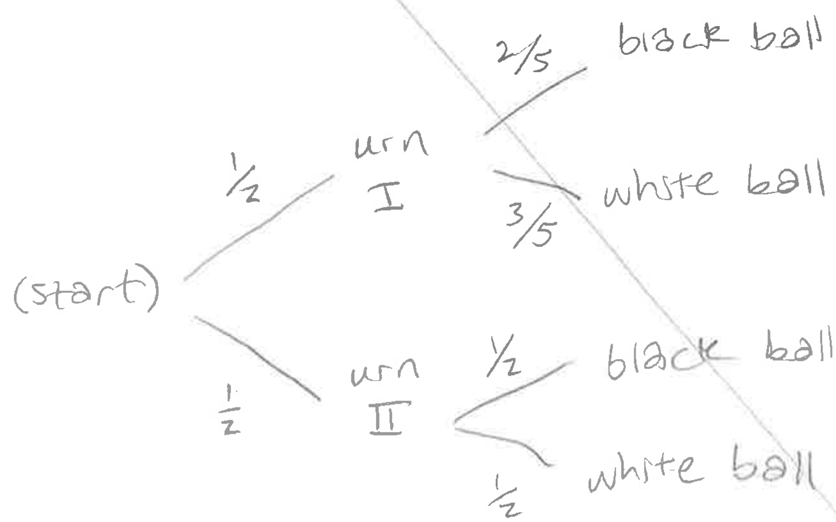
has occured. I.e., a roll of 5 or 6. These are equally likely, so $P(F|E) = \frac{1}{2} = \frac{P(F\cap E)}{P(E)} = \frac{P(F)}{P(E)} = \frac{1/6}{2/6} = \frac{1}{2}$

EX: 4 (Next time)

Suppose we have 2 urns. In urn I there are 2 black balls and 3 white balls; in urn II there is 1 black and 1 white ball. We choose an urn at random, then a ball at random.

What is prob. of getting a black ball from urn I?

we can draw a tree diagram to represent this experiment



Let I = 'urn I is chosen'

B = 'black ball is chosen'

Totals: 3 black
4 white

Then

$P(B|I) = \frac{2}{5}$ ← branch weight $\left(= \frac{P(B\cap I)}{P(I)} = \frac{\frac{1}{5}}{\frac{1}{2}} = \frac{2}{5}\right)$

and

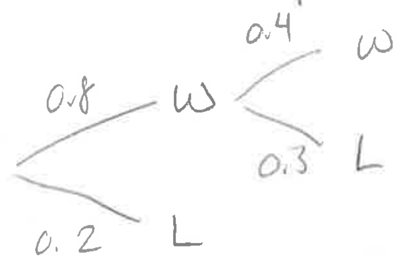
$$P(I|B) = \frac{P(I\cap B)}{P(B)} = \frac{P(B|I)P(I)}{P(B\cap I) + P(B\cap I^c)} = \frac{\frac{2}{5}\left(\frac{1}{2}\right)}{\frac{1}{5}+\frac{1}{4}}$$

$$= \frac{1/5}{4/20 + 5/20} = \frac{1/5}{9/20} = \frac{4}{9}$$

EX 3:

The MN Wild is playing a turnament. In the first round they have an 80% chance of beating their opponent, but if they advance they will only have a 40% chance of beating their next opponent.

Tree diagram:



Prob. of winning tournament:

let A = 'first game W'
B = '2nd game W'

Then
$$P(A \cap B) = P(B|A)\,P(A) = 0.4(0.8) = 0.32$$

EX 5: Monty Hall

You pick door 1, monty reveals a donkey behind door 3. Should you switch?

If you pick curtain 1:

	#1	#2	#3	Monty opens
case 1:	Donkey	D	C	2
case 2:	Donkey	C	D	3
case 3:	Car	D	D	2 or 3

① $P(\text{case } 2, \text{ M opens } \#3) = \frac{1}{3}$

② $P(\text{case } 3, \text{ M opens } 3) = P(\text{case } 3)\, P(\text{open door } 3 \mid \text{case } 3) = \frac{1}{3}\left(\frac{1}{2}\right) = \frac{1}{6}$

$\Rightarrow P(\text{M opens } 3) = ① + ② = \frac{1}{2}$

$$\Rightarrow P(\text{case } 3 \mid \text{M opens } 3) = \frac{P(\text{case } 3, \text{ M opens door } 3)}{P(\text{M opens } 3)}$$

$$= \frac{1/6}{1/2} = 1/3$$

Section 3.2: Two-Stage experiments

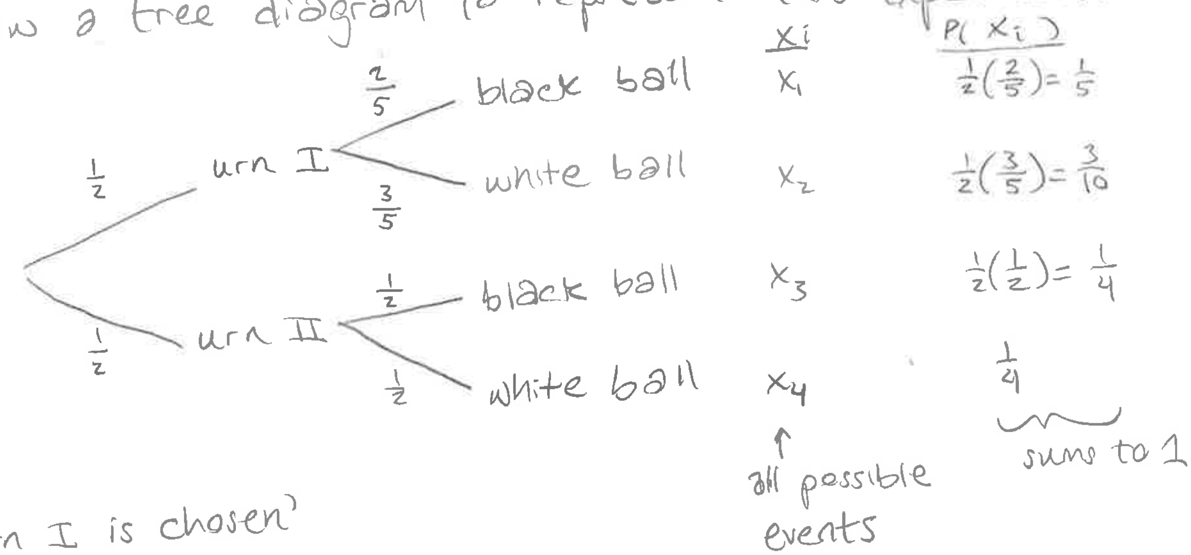
This section is all about calculating probabilities for experiments where 1 event occurs, then another event (hence 'two-stage'). We talked about a very simple two-stage experiment last time with the MN Wild problem; this time we'll discuss some more complex 2-stage experiments.

Example:

Suppose we have 2 urns. Urn I has 2 black balls and 3 white balls, and Urn II has 1 black and 1 white ball. We choose an urn at random, then a ball at random.

Q: What is the probability of getting a black ball from urn I?

We can draw a tree diagram to represent this experiment.



Let I = 'Urn I is chosen'

B = 'Black ball is chosen'

Then $P(B|I)$ is given by the branch weight between urn I and the black ball: 2/5 (the sample space is reduced). Also,

$$P(B|I) = \frac{P(B \cap I)}{P(I)} = \frac{1/5}{1/2} = 2/5$$

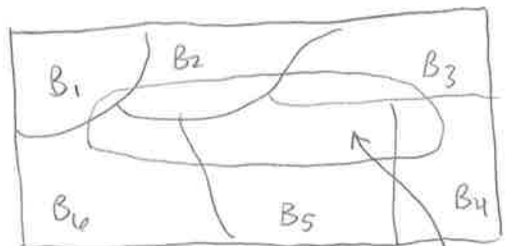
Q: What is the probability that a black ball is chosen from any urn?

$P(x_1) + P(x_3) = \frac{1}{5} + \frac{1}{4} = \frac{9}{20}$ (by counting all possible events w/ a black ball)

We thought about the last question intuitively, but it also follows from a result called the Law of Total Probability.

Law of Total Probability

Let $B_1, B_2, \ldots, B_K$ be a set of disjoint events whose union is Ω (that is, $B_1, \ldots, B_K$ form a partition). Let A be an event that intersects each of $B_1, \ldots, B_K$.



$\Omega = B_1 \cup B_2 \cup \cdots \cup B_K$ A

Then

$$P(A) = \sum_{i=1}^{K} P(A \cap B_i) = \sum_{i=1}^{K} P(A \mid B_i)\, P(B_i)$$

since $A \cap B_i$ is disjoint with $A \cap B_j$, $\forall i \neq j$

def'n conditional probability

Back to example:

$$P(B) = P(B \cap I) + P(B \cap I^c) = \tfrac{1}{5} + \tfrac{1}{4} = \tfrac{9}{20}$$

Q: Given that a black ball was chosen, what is the probability it came from urn I?

$$P(I \mid B) = \frac{P(I \cap B)}{P(B)} = \frac{1/5}{P(B \cap I) + P(B \cap I^c)} = \frac{1/5}{9/20} = \frac{4}{9}$$

By the Law of Total Probability.

The probability that an initial event occured given that you know that a subsequent event occured is sometimes called an inverse probability. More frequently, it's called a Bayesian (or a Bayes) probability.

Section 3.3: Bayes' Formula

Bayesian probabilities are an essential component in real-life probability: we frequently know the result of a series of events and would like to construct the probability of an earlier event in the series.

Example: Exit polls (from book)

In the 1986 CA gubernatorial (governor's) election, exit polls dramatically misjudged the election: Tom Bradley was indicated as winning the election by exit polling, but George Deukmejian won by a landslide.

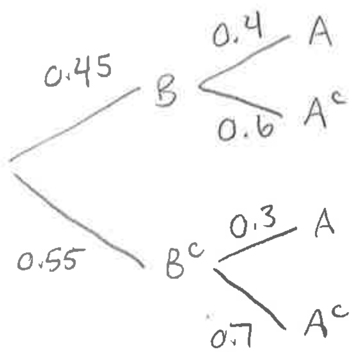
Let B = 'person votes for Bradley'

A = 'voter stops and answers how they voted'

Suppose $P(B) = 0.45$

$P(A|B) = 0.4$ ← 40% Bradley voters will stop

$P(A|B^c) = 0.3$ ← 30% Deukmejon voters will stop.



by def'n conditional probability

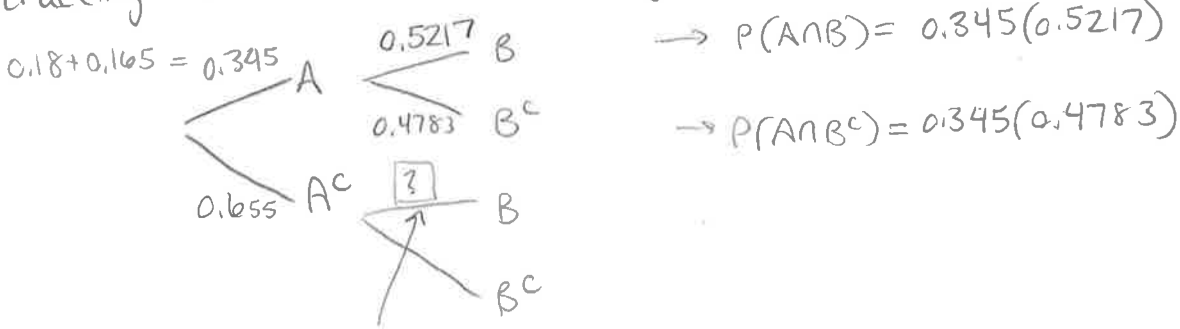
Then

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

$$= \frac{0.45(0.4)}{0.45(0.4) + 0.55(0.3)} = \frac{0.18}{0.18 + 0.165} \approx 0.5217$$

So from this sample, it looks like Bradley will win. But the problem with the exit poll is that the difference in response rates makes the sample not representative of the population as a whole.

Note:
We could also determine the Bayes probability by constructing the reverse tree diagram:



what information would we need to determine this?

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

$$\Rightarrow P(A^c \mid B),\ P(A^c \mid B^c)$$

Bayes' Formula:
Let $B_1, \dots, B_n$ form a partition of the sample space, Ω. We are given $P(B_i)$ and $P(A \mid B_i)$ for $i \in [1, n]$, and we want to calculate $P(B_1 \mid A)$. Then

$$P(B_1 \mid A) = \frac{P(B_1 \cap A)}{P(A)} = \frac{P(B_1)P(A \mid B_1)}{\sum_{i=1}^{n} P(B_i)P(A \mid B_i)}$$

Note:
The book recommends not memorizing this formula, but rather think of it as a conditional probability. I tend to agree, though it's good to have an idea of what you expect.

Section 3.4: Discrete joint distributions

Let's use an example to motivate thinking about joint distributions.

Example:
Toss a coin 3 times. Let the r.v. X_i be the outcome of the i^{th} toss. If the coin is fair, each X_i will have identical distributions.

Let's consider the space of all possible outcomes of the three tosses by thinking of the X_i's as a 3-tuple. We call the distribution of all possible outcomes for this 3-tuple the joint distribution of X_1, X_2, and X_3.

Definition:
Let $X_1, X_2, \dots, X_n$ be r.v.'s associated with an experiment. Suppose the sample space of each X_i is R_i. Then the joint distribution function of $(X_1, X_2, \dots, X_n)$ is the function in $\Omega = R_1 \times R_2 \times \cdots \times R_n$ that gives the probability of each of the outcomes of $(X_1, X_2, \dots, X_n)$.

Example:
Suppose we choose a person at random from a group of 60 people. Of these 60, we can summarize some characteristics about them using the table below:

	Don't smoke	Smoke	Total:
Don't have cancer	40	10	50
Have cancer	7	3	10
Total:	47	13	60

Now consider the random variables

C: whether a person has cancer ($C=1$) or not ($C=0$)

S: whether a person smokes ($S=1$) or not

We can represent the joint distribution of (C, S) as

	$S=0$	$S=1$	
$C=0$	$40/60$	$10/60$	$50/60$
$C=1$	$7/60$	$3/60$	$10/60$
	$47/60$	$13/60$	

write in after defining marginal dist

keep on board

When you have a joint distribution, sometimes all you want is the distribution of an individual r.v.. We can recover this as the marginal distribution.

<u>Definition</u>:
Given the joint distribution of (X, Y), the <u>marginal distributions</u> of X and Y are given by

$$P(X=x) = \sum_y P(\{X=x\} \cap \{Y=y\})$$

$$P(Y=y) = \sum_x P(\{X=x\} \cap \{Y=y\})$$

<u>Question</u>:
In our example, then, what are the marginal distributions?

x	0	1
C	$50/60$	$10/60$

x	0	1
S	$47/60$	$13/60$

When you have two random variables with different events possible, then a natural question to ask me is, are these events independent?

Not only can we talk about the independence of the events, but we can talk about properties of the random variables themselves based on the independence of their respective events.

<u>Definition:</u>

Two random variables X and Y are <u>independent</u> if

$$P(\{X=x\}\cap\{Y=y\}) = P(X=x)P(Y=y)$$

for each outcome x and y. i.e., if their joint distribution is equal to the product of the two marginal distributions.

From our example (it's a hard-working one today), we see that

$$P(C=0) = 50/60 \qquad P(S=0) = 47/60$$
$$P(C=1) = 10/60 \qquad P(S=1) = 13/60$$

} from marginal dist.'s

$$P(\{C=0\}\cap\{S=0\}) = 40/60 \approx 0.6667$$
$$P(\{C=0\}\cap\{S=1\}) = 10/60 \approx 0.1667$$
$$P(\{C=1\}\cap\{S=0\}) = 7/60 \approx 0.1167$$
$$P(\{C=1\}\cap\{S=1\}) = 3/60 \approx 0.05$$

from joint distr.

But, e.g.,

$$P(C=1)P(S=1) = \frac{10}{60}\left(\frac{13}{60}\right) \approx 0.0361 \neq P(\{C=1\}\cap\{S=1\}).$$

so these r.v.'s are <u>not</u> independent.

Continuing along the lines of 'things we can think about in general when we have two or more random variables in a problem, we can talk about conditional probabilities with random variables.

Definition:
Let X and Y be random variables. Suppose the event $Y=y$ occurs (i.e., y is fixed). Then the <u>conditional distribution of X given that $Y=y$</u> is given by

$$P(X=x \mid Y=y) = \frac{P(\{X=x\}\cap\{Y=y\})}{P(Y=y)}$$

$$= \frac{P(\{X=x\}\cap\{Y=y\})}{\sum_u P(\{X=u\}\cap\{Y=y\})}$$

Example:
Insurance companies keep track of how likely various car brands are to be stolen. Suppose a company in Minneapolis has computed the joint distribution

stolen, X \ Brand, Y	1	2	3	4	5
0 (not stolen)	0.129	0.298	0.161	0.280	0.108
1 (stolen)	0.010	0.010	0.001	0.002	0.001

what is the conditional distribution of X (being stolen) given that $Y=1$?

x	0	1
X	$\frac{0.129}{0.129+0.01}$ $=0.9281$	$\frac{0.01}{0.129+0.01}$ $=0.0719$