CS188 Fall 2015 Probability Section 2

1 Marginalization

S	R	$P(S,R)$
s	r	0.05
s	$\neg r$	0.75
$\neg s$	r	0.10
$\neg s$	$\neg r$	0.10

Consider the binary random variables, S and R, from last week. Their joint distribution is given in the table above. Calculate the following quantities:

(a) $\sum_{r' \in (r, \neg r)} P(s, r')$

(b) $\sum_{r' \in (r, \neg r)} P(\neg s, r')$

(c) $\sum_{s' \in (s, \neg s)} P(s', r)$

(d) $\sum_{s' \in (s, \neg s)} P(s', \neg r)$

2 Product Rule and Chain Rule

Suppose that, if we randomly choose a student, the probability that they like to play basketball is 0.01. Now, suppose that, if we randomly choose a student that likes to play basketball, the probability that they are tall is 0.3. In other words, the probability that a student is tall *given that* they like to play basketball is 0.3.

(a) Intuitively, would you expect the probability that a student likes to play basketball *and* is tall to be smaller or larger than 0.01? Why?

(b) Now, calculate the probability that a student likes to play basketball and is tall.

Consider two binary random variables, L and T. L takes on values l and $\neg l$, corresponding to whether or not you're late for work. T takes on values t and $\neg t$ and corresponds to whether or not there's a traffic jam. So, l, t means that it's you're late for work and there's a traffic jam. We are given the following probability tables:

T	$P(T)$
t	0.4
$\neg t$	0.6

L	T	$P(L \mid T)$
l	t	0.8
l	$\neg t$	0.25
$\neg l$	t	0.2
$\neg l$	$\neg t$	0.75

(c) Calculate $P(L, T)$ from the tables given.

L	T	$P(L,T)$
l	t	
l	$\neg t$	
$\neg l$	t	
$\neg l$	$\neg t$	

Now suppose that we add another random variable M representing whether or not your boss is mad, and in addition you are given the following table:

L	T	M	$P(M \mid L, T)$
l	t	m	0.95
l	t	$\neg m$	0.05
l	$\neg t$	m	0.75
l	$\neg t$	$\neg m$	0.25
$\neg l$	t	m	0.40
$\neg l$	t	$\neg m$	0.60
$\neg l$	$\neg t$	m	0.10
$\neg l$	$\neg t$	$\neg m$	0.90

(d) Calculate $P(L, T, M)$ from the tables given.

L	T	M	$P(L, T, M)$
l	t	m	
l	t	$\neg m$	
l	$\neg t$	m	
l	$\neg t$	$\neg m$	
$\neg l$	t	m	
$\neg l$	t	$\neg m$	
$\neg l$	$\neg t$	m	
$\neg l$	$\neg t$	$\neg m$	

3 Normalization Trick

Color	Number
red	6
blue	3
green	1

Suppose you have a bag that contains 10 balls of different colors. You pull all of the balls out of the bag and find the above numbers for each color. Now, you want to figure out what the probability of picking a ball of a certain color is. To do this, you fill in the following probability table:

Color	Number
red	0.6
blue	0.3
green	0.1

This may seem trivial, but there's actually two ways you could have done this. One way is to go through each entry in the table, and figure out the probability of that entry individually. The other way is to take the original table, and simply *normalize* it so that the entries sum to one. This is what is known as the *normalization trick*, and it will come in handy when we look more in depth at conditional probabilities. The normalization trick states that, for a table that gives a set of frequencies that don't necessarily sum to one, *but should*, we can normalize it to get a valid probability distribution, rather than having to recalculate each probability individually.

Let's consider again the probability table from Part 1:

S	R	$P(S, R)$
s	r	0.05
s	$\neg r$	0.75
$\neg s$	r	0.10
$\neg s$	$\neg r$	0.10

Using the normalization trick, calculate $P(S \mid r)$ and $P(S \mid \neg r)$.

S	$P(S \mid r)$
s	
$\neg s$	

S	$P(S \mid \neg r)$
s	
$\neg s$	

4 Bayes' rule

Recall the product rule, which is written as $P(A, B) = P(A \mid B)P(B)$. Notice that we can also use commutativity to obtain:

$$\begin{aligned} P(A, B) = P(B, A) &= P(B \mid A)P(A) \\ \Rightarrow P(A \mid B)P(B) &= P(B \mid A)P(A) \end{aligned}$$

This leads us to Bayes' rule, which is written as follows:

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

For two random variables A and B. This rule in general can be very useful, because it relates conditional probabilities together and allows us to calculate one from the other.

As an example, let's take a look at one variant of what is commonly known as the false positive paradox. In a population of 1000 people, 2% have a deadly disease. You are administering a test for this disease, which has a false positive rate of 5% (i.e., it tests positive when a person doesn't have the disease 5% of the time) and a false negative rate of 0%. Let T be a random variable indicating whether or not the person tests positive, and D indicate whether or not the person actually has the disease. We then have the following tables:

D	$P(D)$
d	0.02
$\neg d$	0.98

D	T	$P(T \mid D)$
d	t	1
d	$\neg t$	0
$\neg d$	t	0.05
$\neg d$	$\neg t$	0.95

Using this information, calculate the probability table $P(D \mid T)$.

D	T	$P(D \mid T)$
d	t	
d	$\neg t$	
$\neg d$	t	
$\neg d$	$\neg t$	