Stat 104: Quantitative Methods for Economics
Homework 3 SOLUTIONS: Due Monday, September 25

1) A journal article reported in September 2003 that a Swiss dermatologist established a link between smoking and gray hair in women under 40, with results as shown in this table. Let G denote the event a women in the study has gray hair, and S denote the event a women in the study is a smoker.

	Gray	Not Gray	Total
Smoker	13	10	23
Nonsmoker	0	32	32
Total	13	42	55

a) Was this an observational study or an experiment?
This was an Observational study since the dermatologist is not assigning any treatment to the participants; he or she is simply observing whether or not these women have gray hair in conjunction with whether or not they are smoking.

b) Find P(G)
P(G) = 13/55

c) Find P(S or G)
P(S or G) = P(S) + P(G) - P(S and G) = 23/55 + 13/55 - 13/55 = 23/55

d) Find P(S|G)
P(S|G) = 13/13 = 1

e) Are having Gray hair and being a smoker independent or dependent events? Explain
They are dependent events since $P(S \mid G) \neq P(S)$. ($P(S \mid G) = 1$, $P(S) = 23/55$)
In other words, the probability that a randomly chosen woman is a smoker is affected by whether or not we know she has gray hair.

2) Suppose $P(A) = 0.76$, $P(B \mid A) = 0.30$ and $P(B \mid \bar{A}) = 0.02$. Find $P(\bar{A})$, P(A and B) , and P($\bar{A}$ and B) . Use these to construct a probability table. Now use the table to find the following:

$P(\bar{A}) = 1 - P(A) = 1 - 0.76 = 0.24$

$P(A \text{ and } B) = P(B \mid A)P(A) = 0.30(0.76) = 0.228$

$P(\bar{A} \text{ and } B) = P(B \mid \bar{A})P(\bar{A}) = 0.02(0.24) = 0.0048$

	A	**$\bar{A}$**	
B	**0.228**	**0.0048**	**0.2328**
$\bar{B}$	**0.532**	**0.2352**	**0.7672**
	0.76	**0.24**	**1**

a) $P(\bar{B} \mid A) = P(A \text{ and } \bar{B}) / P(A) =$ 0.7

b) $P(\bar{B} \mid \bar{A}) = P(\bar{A} \text{ and } \bar{B}) / P(\bar{A}) =$ 0.98

c) $P(B) =$ 0.2328

d) $P(\bar{B}) =$ 0.7672

e) $P(A \mid B) = P(A \text{ and } B) / P(B) =$ 0.98

f) $P(\bar{A} \mid B) = P(\bar{A} \text{ and } B) / P(B) =$ 0.02 or $P(\bar{A} \mid B) = 1 - P(A \mid B) =$ 0.02

g) $P(A \mid \bar{B}) = P(A \text{ and } \bar{B}) / P(\bar{B}) =$ 0.69

h) $P(\bar{A} \mid \bar{B}) = P(\bar{A} \text{ and } \bar{B}) / P(\bar{B}) =$ 0.31 or $P(\bar{A} \mid \bar{B}) = 1 - P(A \mid \bar{B}) =$ 0.31

3) Suppose the probability of having schizophrenia P(s) = 0.01 in the population, and the conditional probability of "hearing voices" given schizophrenia P(hv | s) = 0.66, and the probability of "hearing voices" P(hv) = 0.75. Find the probability of having schizophrenia given "not hearing voices":

Note: S^C = S complement (i.e. not S)

P(S) = 0.01
P(HV | S) = 0.66
P(HV) = 0.75

	S	S^C	
HV	**0.66×0.01**	**0.7435**	**0.75**
$(HV)^C$	**0.0034**	**0.2466**	**0.25**
	0.01	**0.99**	**1**

$$\boldsymbol{P(S|(HV)^C) = \frac{P(S \text{ and } (HV)^C)}{P((HV)^C)} = \frac{0.0034}{0.25} = 0.0136}$$

4) In a class on probability, a statistics professor flips two balanced coins. Both fall to the floor and roll under his desk.

a) A student in the first row informs the professor that he can see both coins. He reports that at least one of them shows tails. What is the probability that the other coin is also tails?

There are four possible scenarios: HH HT, TH, TT. If one coin is tails, then the only scenarios left are HT, TT, or TH. And since all possibilities are equally likely, the probability that there are two tails is 1/3.

b) Suppose the student informs the professor that he can see only one coin and it shows tails. What is the probability that the other coin is also tails?

If the student is seeing the first coin, then the possibilities are TT or TH. If the student is seeing the second coin, then the possibilities are TT or HT. And since all possibilities are equally likely, the probability that there are two tails is 2/4 = 1/2.

5) There are two boxes, Box B1 and Box B2. Box B1 contains 2 red balls and 8 blue balls. Box B2 contains 7 red balls and 3 blue balls. Suppose Jane first randomly chooses one of two boxes B1 and B2, with equal probability, 1/2, of choosing each. Suppose Jane then randomly picks one ball out of the box she has chosen (without telling you which box she had chosen), and shows you the ball she picked. Suppose you only see that the ball Jane picked is red. Given this information, what is the probability that Jane chose box B1?

Let B1 be the event that Jane chooses Box B1
B2 be the event that Jane chooses Box B2
R be the event that Jane selects a red ball
B be the event that Jane selects a blue ball

From the information given,
P(B1) = P(B2) = 1/2
P(R | B1) = 2/10
P(R | B2) = 7/10

Then,

	B1	**B2**	
R	**0.1**	**0.35**	**0.45**
B	**0.4**	**0.15**	**0.55**
	0.5	**0.5**	**1**

$$\mathbf{P(B1 \mid R) = \frac{P(B1\ and\ R)}{P(R)} = \frac{P(R \mid B1)\ \times\ P(B1)}{P(R)} = \frac{\frac{2}{10} \times \frac{1}{2}}{\frac{9}{20}} = \frac{\frac{2}{20}}{\frac{9}{20}} = \frac{2}{9} \approx 0.222}$$

Alternatively, without creating a table:

$$\mathbf{P(B1 \mid R) = \frac{P(B1\ and\ R)}{P(R)} = \frac{P(R \mid B1) \times P(B1)}{P(R)} = \frac{P(R \mid B1) \times P(B1)}{P(R \mid B1) \times P(B1) + P(R \mid B2) \times P(B2)} =}$$

$$\mathbf{\frac{\frac{2}{10} \times \frac{1}{2}}{\frac{2}{10} \times \frac{1}{2} + \frac{7}{10} \times \frac{1}{2}} = \frac{\frac{2}{20}}{\frac{9}{20}} = \frac{2}{9}}$$

6) Suppose it has been observed empirically that the word "Congratulations" occurs in 1 out of 10 spam emails (that is, P(congratulations|spam) = 0.1), but that "Congratulations" only occurs in 1 out of 1000 non-spam emails. Suppose it has also been observed empirically that about 4 out of 10 emails are spam. In Bayesian Spam Filtering, these empirical probabilities are interpreted as genuine probabilities in order to help estimate the probability that an incoming email is spam. Suppose we get a new email that contains "Congratulations". Let C be the event that a new email contains "Congratulations". Let S be the event that a new email is spam. We have observed C. Calculate P(S | C)

From the information given,
P(C | S) = 0.1
P(C | S^C) = 0.001
P(S) = 0.4

	S	**S^C**	
C	**0.04**	**0.0006**	**0.0406**
C^C	**0.36**	**0.5994**	**0.9594**
	0.4	**0.6**	**1**

$$\mathbf{P(S \mid C) = \frac{P(S\ and\ C)}{P(C)} = \frac{P(C \mid S) \times P(S)}{P(C)} = \frac{0.1 \times 0.4}{0.0406} = 0.9852}$$

7) McDonald's is planning on opening a new location. They must decide how big of a restaurant to build at the location: small, medium, or large. Demand for the McDonald's in this location is uncertain, and will affect profitability. They have projected profitability for weak, moderate, and strong demand as shown in the following table:

	Demand		
Size	Weak	Moderate	Strong
Small	400	500	660
Medium	-250	650	800
Large	-400	580	990
Best	400	650	990

a) What is the maximax decision?

The maximax decision is the option which maximizes the best-case scenario. Thus, the maximax decision is for McDonald's to build a large restaurant since that gives the highest best-case scenario of 990.

b) What is the maximin decision?

The minimax decision is the option which maximizes the worst-case scenario. Thus, the maximin decision is for McDonald's to build a small restaurant, since that gives the highest worst-case scenario of 400.

8) Video Tech is considering marketing one of two new video games for the coming season: Battle Pacific or Space Pirates. Battle Pacific is a unique game and appears to have no competition. Estimated profits (in thousands of dollars) under high, medium, and low demand are as follows:

	Demand		
Battle Pacific	High	Medium	Low
Profit	$1000	$700	$300
Probability	0.2	0.5	0.3

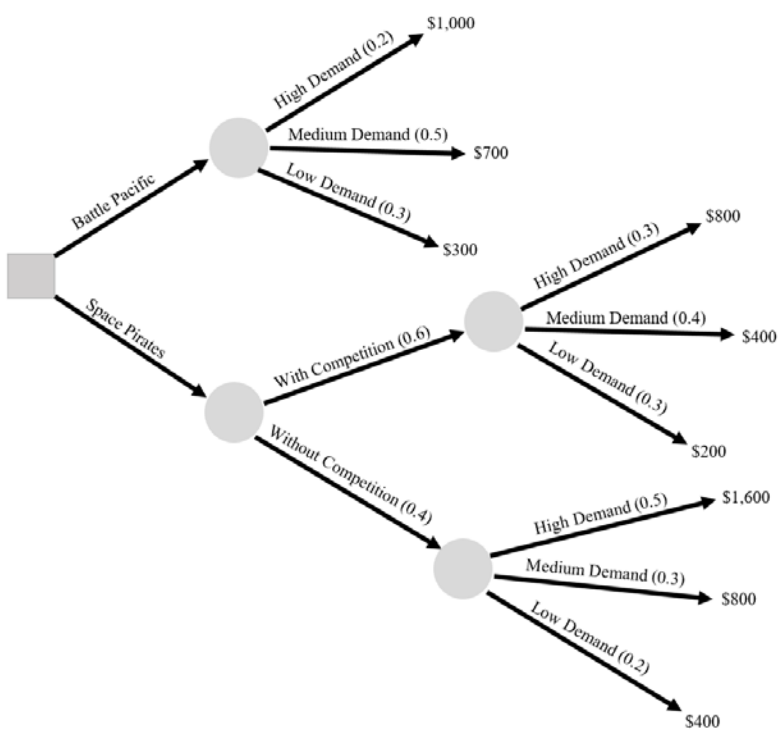
Video Tech is optimistic about its Space Pirates game. However, the concern is that profitability will be affected by a competitor's introduction of a video game viewed as similar to Space Pirates. Estimated profits (in thousands of dollars) with and without competition are as follows:

	Demand		
Space Pirates			
With Competition	High	Medium	Low
Profit	$800	$400	$200
Probability	0.3	0.4	0.3

Space Pirates			
Without Competition	High	Medium	Low
Profit	$1600	$800	$400
Probability	0.5	0.3	0.2

Video Tech believes there is a 0.6 probability that its competitor will produce a new game similar to space Pirates

a) Draw the decision tree



b) What is the optimal decision?

Let P be the profit of Video Tech (in thousands of dollars)

Battle Pacific: E(P) = 0.2($1,000) + 0.5($700) + 0.3($300) = $640
Space Pirates: E(P) = 0.6[0.3($800) + 0.4($400) + 0.3($200)] +
0.4[0.5($1,600) + 0.3($800) + 0.2($400)] = $724

Since the expected profit is higher for Space Pirates, marketing Space Pirates is the optimal decision

c) Right now the chance of competition is 60%. How much higher or lower would this probability need to be in order for your decision to change?

Let p be the probability of competition. We should change our mind if and only if the expected profit for marketing Space Pirates was lower than the expected profit for marketing Battle Pacific. So,

E(P) = p[0.3(800) + 0.4(400) + 0.3(200)] +
(1 - p)[0.5(1,600) + 0.3(800) + 0.2(400)] = 460p + 1,120(1 - p) < 640

460p + 1,120(1 - p) < 640 ⇒ -660p + 1,120 < 640 ⇒ -660p < -480 ⇒ p > 8/11 ≈ .72727

9. Let X be a discrete random variable with PMF (probability mass function) given by

$$p_X(x) = \begin{cases} x^2/a, & if\ x = -3, -2, -1, 0, 1, 2, 3 \\ 0, & otherwise \end{cases}$$

a) Find the value of *a*
Since all probabilities must add up to 1,
$(-3)^2/a + (-2)^2/a + (-1)^2/a + 0^2/a + 1^2/a + 2^2/a + 3^2/a = 28/a = 1 \Rightarrow a = 28$

b) Calculate E(X)
E(X) = -3(9/28) + -2(4/28) + -1(1/28) + 0(0/28) + 1(1/28) + 2(4/28) + 3(9/28) = 0

Also, because the PMF is symmetric around X = 0, E(X) = 0.

10) The night watchman in a factory cannot guard both the safe in back and the cash register in front. The safe contains $6000, while the register has only $1000. Tonight the guard fears a robbery; the probability that the thief will try the cash register is 0.8 and the probability the thief will try the safe is 0.2. If the guard is not present, the thief will take all the money. If the guard is present, the thief will go away empty handed. Where should the guard be positioned in order to minimize the thief's gains?

Let L be the amount of money that the factory loses

Guard is at safe: E(L) = 0.8($1,000) + 0.2(0) = $800
Guard is at cash register: E(L) = 0.8(0) + 0.2($6,000) = $1,200

So in order to minimize the expected amount lost, the guard should be positioned at the safe.

11) The probability that a cellular phone company kiosk sells *X* number of new phone contracts per day is shown below.

X	4	5	6	8	10
P(X)	0.3	0.15	0.35	0.15	0.05

a) Find the mean, variance, and standard deviation for this probability distribution.

Mean(X) = E(X) = 4(0.3) + 5(0.15) + 6(0.35) + 8(0.15) + 10(0.05) = 5.75

$Var(X) = (4 - 5.75)^2 \times 0.3 + (5-5.75)^2 \times 0.15 + (6-5.75)^2 \times 0.35 + (8-5.75)^2 \times 0.15 + (10-5.75)^2 \times 0.05 = 2.6875$

$SD(X) = \sqrt{Var(X)} \approx 1.6394$

b) Suppose the kiosk salesperson makes $80/day (8 hours at $10/hour), plus a $25 bonus for each new phone contract sold. What is the mean and variance of the salesperson's daily salary?

Let Salary = S = 25X + 80

Then,
E(S) = E(25X + 80) = 25E(X) + 80 = 25(5.75) + 80 = 223.75
$Var(S) = Var(25X + 80) = 25^2 Var(X) = 625(2.6875) = 1679.6875$

12) In a population of students, the number of calculators owned is a random variable X with P(X = 0) = 0.2, P(X = 1) = 0.6, P(X = 2) = 0.2.

a) Find E(X)
E(X) = 0(0.2) + 1(0.6) + 2(0.2) = 1

b) Find Var(X)
$Var(X) = (0 - 1)^2 \times 0.2 + (1 - 1)^2 \times 0.6 + (2 - 1)^2 \times 0.2 = 0.4$

13) You roll two dice.

a) What is the probability of two sixes? Of exactly one 6? Of no sixes?

1, 1	2, 1	3, 1	4, 1	5, 1	6, 1
1, 2	2, 2	3, 2	4, 2	5, 2	6, 2
1, 3	2, 3	3, 3	4, 3	5, 3	6, 3
1, 4	2, 4	3, 4	4, 4	5, 4	6, 4
1, 5	2, 5	3, 5	4, 5	5, 5	6, 5
1, 6	2, 6	3, 6	4, 6	5, 6	6, 6

Since there is 6 possible outcomes for the first die and 6 possible outcomes for the second die, there are a total of 6 × 6 = 36 total possible combinations. Since all combinations are equally likely, and there is only one way to get two sixes (6, 6), P(Two sixes) = 1/36.

Likewise, since there are 10 different ways to get exactly one 6, P(Exactly one six) = 10/36.

And P(No sixes) = 1 - P(at least one six) = 1 - 11/36 = 25/36.

b) What is the expected number of sixes that will show?

E(Number of sixes) = 25/36 × 0 + 10/36 × 1 + 1/36 × 2 = 12/36 = 1/3

14) We can simulate the expected value result in part (b) above. Follow the following steps in R:

i. Simulate two dice rolls using (use similar code for die2)
die1=sample(1:6,10000,replace=TRUE)

ii. Combine the two dice rolls into a matrix using dicerolls=cbind(die1,die2)

iii. Each row of dicerolls represents the outcome of rolling two dice. We want to count how many 6's appear each time we roll two dice. We do that as follows.
num6=rowSums(dicerolls==6)

iv. Take the mean of the num6 variable and compare it to part (b) above. How does this mean change if we instead use 1000000 rolls?

Answers will vary, but both answers should be fairly close to 1/3, and as we increase the number of rolls, the closer the average should be to 1/3.

15) If random variable X has mean μ and variance σ^2, show (using the a + bX rule) what the mean and variance of Z = (X - μ) / σ are.

$$\boldsymbol{E(Z) = E\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma}E(X - \mu) = \frac{1}{\sigma}E(X) - \frac{1}{\sigma}E(\mu) = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0}$$

$$\boldsymbol{Var(Z) = Var\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma^2}Var(X - \mu) = \frac{1}{\sigma^2}Var(X) = \frac{\sigma^2}{\sigma^2} = 1}$$

16) Find the variance of each of the following bets from the class notes. Which bet is riskiest and which best is safest?

i. You get \$5 with probability 1.0.
The higher the variance, the riskier the bet!

Variance = $(5 - 5)^2$ x 1 = 0. This is the safest bet as you will always get \$5 no matter what. Thus this is the safest bet.

ii. You get \$10 with probability 0.5, or \$0 with probability 0.5.
Variance = $(10 - 5)^2$ × 0.5 + $(0 - 5)^2$ x 0.5 = 25.

iii. You get \$5 with probability 0.5, \$10 with probability 0.25 and \$0 with probability 0.25.
Variance = $(5 - 5)^2$ x 0.5 + $(10 - 5)^2$ x 0.25 + $(0 - 5)^2$ x 0.25= 12.5.

iv. You get \$5 with probability 0.5, \$105 with probability 0.25 or lose \$95 with probability 0.25.
Variance = $(5 - 5)^2$ x 0.5 + $(105 - 5)^2$ x 0.25 + $(-95 - 5)^2$ x 0.25 = 5000. Since this one has the highest variance, it is the riskiest bet.

17) Let X be a random variable with E(X) = 120 and Var(X) = 20. Find the following.

a) $E(X^2)$
Var(X) = $E(X^2)$ - $E(X)^2$ ⇒ $E(X^2)$ = Var(X) + $E(X)^2$ = 20 + 120^2 = 14,420

b) E (3X + 10)
E(3X + 10) = 3E(X) + 10 = 3(120) + 10 = 370

c) E (-X)
E(-X) = E (-1X) = -1E(X) = -1 × 120 = -120

d) Standard deviation of –2X
Var(-2X) = $(-2)^2$Var(X) = 4Var(X) = 4(20) = 80 ⇒ SD(-2X) = √80 ≈ 8.94