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

Homework policy: Homework is due by 8:00 AM (EST) on the due date. Homework is to be handed in via the course website in pdf format. You do not need to type the homework; there are many ways (scanner in the library or phone apps) to convert written homework into a pdf file. Ask the teaching staff if you need assistance.

Late homework will not be accepted. You are encouraged to discuss homework problems with other students (and with the instructor and TFs, of course), but you must write your final answer in your own words. Solutions prepared "in committee" or by copying someone else's paper are not acceptable.

- Please submit your homework in pdf format; this can be done in Word, or OpenOffice or via cellphone apps that will scan and turn into pdf.
- Please make your homework solutions legible by **bolding** or using circles to identify your solution.
- Since we are not printing out anything, use lots of s p a c e for your solutions, and put each answer on a different page if it makes the solution easier to read.
- Please make sure your submitted solutions are in numerical order [problem 1, problem 2 and so on].
- Please keep your computer output to a minimum and focus on the required answer. The easiest way to put your computer output into your homework is to cut and paste it into a Word file and use the font "courier new".
- Please keep in mind the course rules on Academic Honesty and Collaboration

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?
b) Find P(G)
c) Find P(S or G)
d) Find P(S|G)
e) Are having Gray hair and being a smoker independent or dependent events? Explain

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

a) $P(\bar{B} \mid A)$
b) $P(\bar{B} \mid \bar{A})$
c) $P(B)$
d) $P(\bar{B})$
e) $P(A \mid B)$
f) $P(\bar{A} \mid B)$
g) $P(A \mid \bar{B})$
h) $P(\bar{A} \mid \bar{B})$

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”:

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?
   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?

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?

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 $P(S \mid C)$

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 ?
   b) What is the maximin decision?

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?
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?

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*
b) Calculate E(X)

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?

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.
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?

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)
b) Find Var(X)

13) You roll two dice.

a) What is the probability of two sixes? Of exactly one 6? Of no sixes?
b) What is the expected number of sixes that will show?

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=head(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?

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

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

a) You get \$5 with probability 1.0.
b) You get \$10 with probability 0.5, or \$0 with probability 0.5.
c) You get \$5 with probability 0.5, \$10 with probability 0.25 and \$0 with probability 0.25.
d) You get \$5 with probability 0.5, \$105 with probability 0.25 or lose \$95 with probability 0.25.

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

a) $E(X^2)$
b) E (3X + 10)
c) E (-X)
d) Standard deviation of –2X ?