Problem Set #1

(1) The following table gives the joint probability distribution $p(X, Y)$ of random variables X and Y.

	X		
Y	1	2	3
1	.02	.04	.12
2	.03	.18	.04
3	.00	.02	.10
4	.09	.18	.18

Determine the following:

(a) Do the entries of the table satisfy the conditions for a bivariate density function?

(b) The marginal (or unconditional) probability distributions of X and Y. [Note: These will be a collection of probabilities: the probabilities associated with the 3 values of X and the probabilities associated with the 4 values of Y].

(c) The conditional probability distributions $p(X|Y = 3)$ and $p(Y|X = 1)$. (Note: The first conditional probability distribution is the collection of three numbers, $\Pr(X = 1|Y = 3), \Pr(X = 2|Y = 3), \Pr(X = 3|Y = 3)$.)

(2) (Based on a newspaper article I recently read). Suppose you want to study the impact of a certain *trait* T on the probability of having a heart attack. Let $D = 1$ denote the event of a heart attack and $D = 0$ denote the event of no heart attack. The variable T is also binary (i.e., it takes on only one of two possible values), and $T = 1$ indicates that an individual has the trait, and $T = 0$ indicates that he or she does not have the trait. (You might think of T as representing if you have a particular gene, if you have high cholesterol, etc).

The article claimed that the trait T had no effect on the prevalence of heart attacks

($D = 1$) because "among those individuals who had heart attacks, 50 percent of them had the trait ($T = 1$) while 50 percent of them did not ($T = 0$)."

Critically evaluate this statement. Does this result necessarily imply that the probability of having a heart attack is unrelated to having the trait T? (Hint: Think about what the statement means formally in terms of conditional probability. You might try and construct some 2×2 probability tables to investigate this issue).

(3) Extra Credit: (Worth 2 percentage points on the first exam).

(*Based on The Monte Hall Problem*) In a popular game show, a contestant is asked to pick one of three boxes. One of the three boxes contains a prize, while the other two are empty. After the contestant makes her initial selection, the game show host opens one of the two boxes which was not chosen, and shows her that this box is empty. (The host knew that this box was empty all along). The contestant then has the option to either stay with her original choice, or switch to the other unopened box. If the contestant wishes to maximize her probability of getting the prize, what should she do? (Stay with her original choice, or switch?)

(Hint: This is all conditional probability! To get credit for your answer, you MUST support your argument using laws of conditional probability. Guesses or intuition will not be sufficient to earn the extra credit).