CSC 535 – Probabilistic Graphical Models

Assignment Two

Due: 11:59pm (*) Wednesday, September 30.

Weight about 7 points

This assignment should be done individually

The purpose of this assignment is to get familiar with basic probability, and continue practicing implement probabilistic computations. Key concepts that are featured include conditional independence, Bayes′ rule in action, and properties of the normal distribution.

Deliverables

Deliverables are specified below in more detail. For the high level perspective, you are to provide a program to output a few numbers and to create figures. You also need to create a PDF document that tells the story of the assignment including output, plots, and images that are displayed when the program runs. Even if the question does not explicitly remind you to put the resulting image into the PDF, if it is flagged with **($)**, you should do so. The instructor should not need to run the program to verify that you attempted the question. See

http://kobus.ca/teaching/grad-assignment-instructions.pdf

for more details about preparing write-ups. While it takes work, it is well worth getting better (and more efficient) at this. **A substantive part of each assignment grade is reserved for exposition.**

For this first problem, minimal writeup (e.g., just answers) is fine.

1. Consider throwing a 6 sided die two times in succession. Let A be the number of the first throw and B be the number of the second throw. Let S be the random variable that is the sum of the numbers. What is (**$**):

a) P(A=3, B=3 | S=6)
b) P(S=6 | A=3, B=3)
c) P(A=1, B=1 | S=6)
d) P(A=1|S=2)
e) P(A=3|S=2)
f) P(A=2|S=6)
g) P (S=12)
h) P(S=6)

2. Suppose you roll an N-sided dice, each side labeled with a number 1,2,…,N. Further assume the dice is *fair* (i.e. any side is equally likely to be rolled). What is the expected number of rolls needed before every side comes up at least once?

Hint: Let the random variable X be the number of rolls needed to see all sides at least once. Let X_i be the number of times you rolled while having seen (i-1) sides. Express X as a function of the X_i's and note that we covered the distribution of X_i in lecture (see the slides).

3. Answer 8.3 in Bishop reproduced below (**$**). In Bishop this problem is a double star one, but I don't think it is so bad (although a bit tedious, so think a bit about how to organize things).

Table 8.2 The joint distribution over three binary variables.

a	b	c	$p(a,b,c) \times 1000$
0	0	0	192
0	0	1	144
0	1	0	48
0	1	1	216
1	0	0	192
1	0	1	64
1	1	0	48
1	1	1	96

8.3 ($\star\star$) Consider three binary variables $a, b, c \in \{0, 1\}$ having the joint distribution given in Table 8.2. Show by direct evaluation that this distribution has the property that a and b are marginally dependent, so that $p(a, b) \neq p(a)p(b)$, but that they become independent when conditioned on c, so that $p(a, b|c) = p(a|c)p(b|c)$ for both $c = 0$ and $c = 1$.

4. Consider these two propositions for conditional independence of X and Y, given Z:

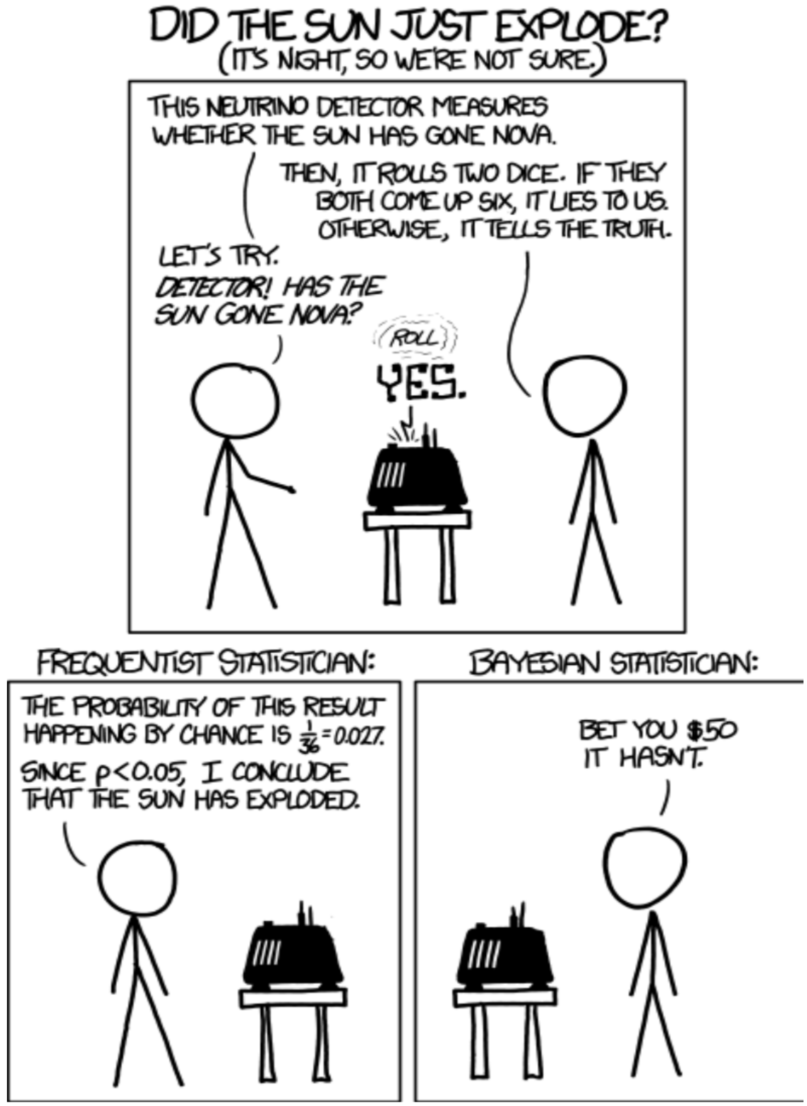
(A) P(X|Y,Z)=P(X|Z) or P(Y,Z)=0.

(B) P(X,Y|Z)=P(X|Z)P(Y|Z)

Show that (A) implies (B) and that (B) implies (A) (**$**). If you like, simply pretend that probabilities are never zero, rather than fretting about the special cases of P(Y,Z) being zero. I am not looking for a tight proof. The point of this problem is to get used to manipulating expressions using basic probability definitions.

5. As the Bayesian in the following picture, what are the chances that you will win the bet if the prior probability that the sun has exploded is **($)**: a) 1/6; b) 1/36; c) 1/1000; d) 1/1,000,000. Be sure to show your work **($)**!

Hint. There are 3 random variables here. You might find it easiest to marginalize out the throw of the two die at a convenient point. However, due to the fancy construction of the setup, you could also do this implicitly by focusing on the cases that can occur, thereby implicitly ignoring terms in the marginal that are zero. For example, the probability of the sun exploding, the detector saying yes, and the roll being two sixes is zero.



6.

For this problem you should code up the formula for the univariate and bivariate Gaussian distributions rather than hunt for existing options in your software environment (in Matlab the they are `normpdf()` and `mvnpdf()`). You are welcome to use such functions to check your answer. The reason for entering the formula yourself is to get some feel for the form of the function. For the multivariate version, you need to find the determinant of a matrix, which is non-trivial in general, but very easy for 2x2, where it is given by $det(A)=A_{11}A_{22}-A_{12}A_{21}$. You can also make use of an external function for the inverse of a matrix, but again, for 2x2 it is a simple formula that you can look up.

a) Plot three univariate Gaussians on one graph. Specifically, plot Gaussians with means and variances (2, 0.2), (1, 0.5), and (0, 2) (**$**). Be sure to provide a caption.

b) Make a 3D plot of p(x,y) for a bivariate Gaussian with mean (0,0) and covariance matrix.

$$\begin{vmatrix} 0.5 & 0.8 \\ 0.8 & 2.0 \end{vmatrix}$$

Be sure to provide a caption. (**$**).

For the next two problems, the main point is getting familiar with the joint distribution for continuous domains, and also seeing two very interesting properties about Gaussian distributions in action. Once you have done the first one, the second one should not take too much time as you will have built some of the programming infrastructure.

7. For the bivariate Gaussian in part (b) of problem 6, using numerical integration, approximate p(x) for a sensible stepping of x values and plot the result (**$**). Be sure to provide a caption. Does it have the shape you expect, and what is that shape (**$**)?

Alternative: If proof by programming is not for you, you can instead show analytically that for a bivariate Gaussian p(x,y), p(x) is also Gaussian (**$**). What do you expect the area of the curve to be (**$**)? It may be easier to work with the precision matrix which is the inverse of the covariance matrix and which is also symmetric.

8. For the bivariate Gaussian in part (b) of problem 6, plot $p\left(x \middle| y=2.0\right)$ (**$**). Be sure to provide a good caption.

Does it have the shape you expect, and what is that shape (**$**)? What do you expect the area of the curve to be (**$**)?

Alternative: If proof by programming is not for you, show analytically that for a bivariate Gaussian p(x,y), p(x|y) is also Gaussian ($). It may be easier to work with the precision matrix which is the inverse of the covariance matrix and which is also symmetric.

What to Hand In

Hand in a program hw2.<suffix> (e.g., hw2.m if you are working in Matlab) and the PDF file hw2.pdf with the story of your efforts into D2L.