CSC 535 – Probabilistic Graphical Models

Assignment Two

Due: 11:59pm (*) Monday, September 10.

(*) There is grace until 8am the next morning, as the instructor will not grade assignment before then. However, once the instructor starts grading assignments, no more assignments will be accepted.

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.

---

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. 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), and so it has no stars for us.

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 $c = 1$.

3.

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.

$$\begin{vmatrix} 0.5 & 0.3 \\ 0.3 & 2.0 \end{vmatrix}$$

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.

4. For the bivariate Gaussian in part (b) of problem 3, 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.

5. For the bivariate Gaussian in part (b) of problem 3, and using the result of question 4, plot p(y|X=0.5) ($). Be sure to provide a caption. Does it have the shape you expect, and what is that shape ($)? 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.

6. Consider these propositions for conditional independence of X and Y, given Z, and do either (a) or (b) below:

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

(B) $P(Y|X,Z)=P(Y|Z)$ or $P(X,Z)=0$

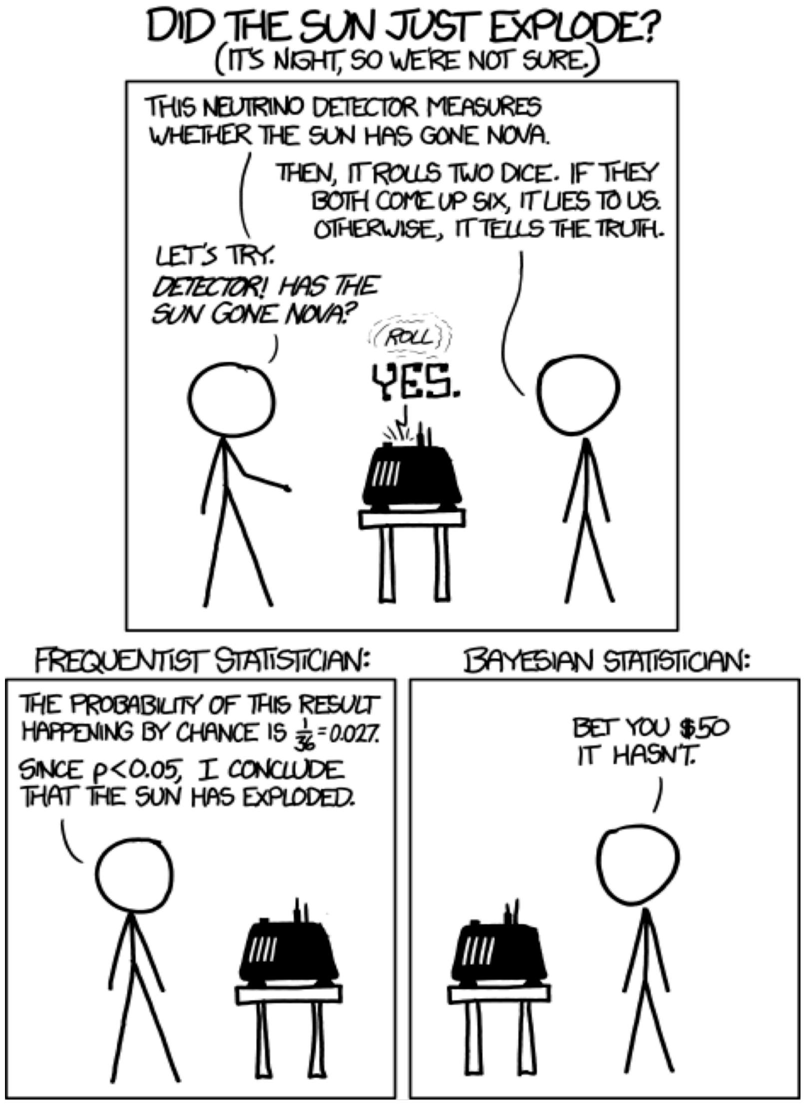
(C) $P(X,Y|Z)=P(X|Z)P(Y|Z)$

a) Show that (A) implies (B) ($), that (A) implies (C) ($), and that (C) implies (A) ($). (By symmetry, the third case means that (C) also implies (B), so no need to show that). If you like, simply pretend that probabilities are never zero, rather than fretting about the special cases of P(X,Z) or 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.)

b) Describe a problem of interest to you, ideally from your own work, where there is conditional independence, using the first definition (A) or (B, which is symmetric with A) to explain/illustrate the conditional independence of two arbitrary data points ($). Then write down a mathematical expression for your problem, using (C), introducing any notation you need ($). Quite likely your problem will be better described by the more general form of (C) for multiple variables instead of the two (X,Y), namely $P(\{X_i\}|Z)=\prod P(X_i|Z)$.

Hint. If you are having trouble thinking of your own problem in this form, start with the simple problem of points on a line, whose parameters are Z, with X and Y (or $\{X_i\}$) being the observed data (points on the line).

7. As the Bayesian in the following picture, what is the chances you will win the bet, under the assumption that the prior probability that the sun has exploded is **($)**: a) 1/6; b) 1/36; c) 1/1000; d) 1/1,000,000. Be sure to show your work **($)**! Note that unlike the TB example in class, there are 3 random variables here. You might find it easiest to marginalize out the throw of the two die at a convenient point.



If you have done all the questions until here, then congratulations, you are done! However, if you are looking for alternatives, or extra problems, feel free to keep reading.

8. (**) Consider the simple D dimension Gaussian that have mean vector **0** and variance vector **1**. Derive an expression, $p(r)$, for the probability mass at a radius r ($). In other words, $p(r)\Delta r$ is the probability that a sampled point is between *r* and $r+\Delta r$, where, for a D dimensional point, **X**, $r^2 = X^T X$. Plot p(r) for D=1, D=2, D=10, D=20, and D=30. Comment on the result ($).

9. (**) Transform (x,y) so that the bivariate Gaussian in problem 3.a) has diagonal covariance in a different space. A question on the previous homework might be relevant here. Report the transformation matrix, and redo the plots ($). Be sure to have an informative caption.

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.