Stat 516, 2014 Homework 2

Due date: Tuesday, October 14.

Note: Do this homework in *pairs*; two students turning in a single joint solution. No two Statistics or Biostatistics Ph.D. students may work together. No two QERM Ph.D. students may work together, etc. (Exceptions as needed based on the make-up of the class.)

1. *Conditional independence.*

   (a) Let (X, Y, Z) be a random vector (to be clear, taking values in $\mathbb{R}^3$). Suppose (X, Y, Z) is multivariate normal with $X \perp\!\!\!\perp Y$ and $X \perp\!\!\!\perp Y \mid Z$. Show that it must hold that $X \perp\!\!\!\perp (Y, Z)$ or $(X, Z) \perp\!\!\!\perp Y$.

   (b) Give an example of a joint distribution for three binary random variables (r.v.) X, Y and Z such that $X \perp\!\!\!\perp Y$ and $X \perp\!\!\!\perp Z$ but X is not independent of the pair (Y, Z).

   (c) Let X and Y be two r.v. taking values in the set $\{1, 2, 3\}$, with the following matrix of joint probabilities $p_{ij} = P(X = i, Y = j)$:

$$P = (p_{ij}) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} \frac{1}{23} & \frac{1}{23} & \frac{1}{23} \\ \frac{1}{23} & \frac{2}{23} & \frac{3}{23} \\ \frac{1}{23} & \frac{4}{23} & \frac{9}{23} \end{bmatrix} \end{array}.$$

   Does there exist a binary r.v. Z such that $X \perp\!\!\!\perp Y \mid Z$? (*Hint*: Think about the rank of P.)

2. *Conditional independence and graph separation.*

   Solve the following problem from Guttorp's book using what you have learned about conditional independence and graph separation for distributions that factorize according to a graph.

   (Guttorp, 2.14.1) Keeping with the notation used in the lectures, prove that the requirement we made in the definition of a Markov chain $(X_n)_{n \geq 0}$ (compare the slide entitled "Markov chains") is equivalent to each of the following two statements:

   (a) Let $T_1 \subset \{n+1, n+2, \dots\}$ be a finite set of times later than n, and $T_0 \subseteq \{0, \dots, n\}$ a set of times less than or equal to n. Let $t_0 = \max T_0$. Then

$$\Pr(X_k = i_k \,\forall k \in T_1 \mid X_l = i_l \,\forall l \in T_0) = \Pr(X_k = i_k \,\forall k \in T_1 \mid X_{t_0} = i_{t_0})$$

   for all collections of states i_k, i_l, i_{t_0} for which the conditional probabilities are well-defined.

   (b) Let $T_1 \subset \{n+1, n+2, \dots\}$ be a finite set of times later than n, and $T_0 \subseteq \{0, \dots, n-1\}$ a set of times prior to n. Then

$$\begin{aligned} &\Pr(X_k = i_k \,\forall k \in T_1,\ X_l = i_l \,\forall l \in T_0 \mid X_n = i_n) \\ &\qquad = \Pr(X_k = i_k \,\forall k \in T_1 \mid X_n = i_n) \Pr(X_l = i_l \,\forall l \in T_0 \mid X_n = i_n). \end{aligned}$$

   for all collections of states i_k, i_l, i_n for which the conditional probabilities are well-defined.

(Turn page)

3. *Recognizing Markov chains.*

   A fair die is thrown over and over again.

   (a) Let X_n be the maximum reading obtained in the first n throws. Show that X_n is a Markov chain, and find p_{ij} and $p_{ij}^{(n)}$.

   (b) Let Z_n be the second largest reading obtained in the first n throws, $n \geq 2$. Is Z_n a Markov chain? Give an 'intuitive' explanation and then describe what you might do for a rigorous proof of your claim.

4. *Markov chain with two states.*

   Let $P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$ be the transition probability matrix of a Markov chain with two states. Write the n-th power of P as

   $$P^n = \begin{pmatrix} 1-a_n & a_n \\ b_n & 1-b_n \end{pmatrix}.$$

   Give a formula for a_n, b_n, and use it to find the limit of P^n as $n \to \infty$. Does P^n always converge?

5. *Simulating gambler's ruin.*

   Write a routine to simulate realizations of the gambler's ruin chain $\{X_n\}$ with probabilities $p_{i,i+1} = p$, $p_{i,i-1} = q$, $p + q = 1$. The routine should stop simulations as soon as you hit one of the absorbing states. Your input will consist of an initial state i, state space size N, and probability of increasing gambler's fortune p. The routine should return a vector of Markov chain states until absorption.

   (a) Provide the source code in any computer language of your choice and output of your routine in the form of 20 random realizations of the Markov chain for input parameters $N = 10$, $i = 3$, and $p = 0.32$.

   (b) Use your simulation routine to estimate the probability of reaching the largest state $N = 10$ starting at state 4, denoted $h(4, p)$, for probabilities $p_{i,i+1} = p \in \{0.1, 0.2, \ldots, 0.9\}$. Turn in a graph with estimated $h(4, p)$ plotted against p.