Stat 516, Homework 2

Due date: Thursday, October 12.

Note: Do this homework in *pairs*; two students turning in a single joint solution. No two students from the same department/program may work together. (Exceptions as needed based on the make-up of the class.)

1. *Guttorp, 1.6.2: Consistent finite-dimensional distributions*

   Let $\lambda \in (0, \infty)$. For any $n \in \mathbb{N}$, real numbers $0 \le t_1 < \cdots < t_n$ and integers $r_1, \ldots, r_n$, define the probability

$$p_{t_1,\ldots,t_n}(r_1, \ldots, r_n) = \begin{cases} \dfrac{\lambda^{r_n} e^{-\lambda t_n} t_1^{r_1} (t_2 - t_1)^{r_2 - r_1} \ldots (t_n - t_{n-1})^{r_n - r_{n-1}}}{r_1!(r_2 - r_1)! \ldots (r_n - r_{n-1})!} & \text{if } 0 \le r_1 \le \cdots \le r_n, \\ 0 & \text{otherwise.} \end{cases}$$

   For unordered reals $t_1, \ldots, t_n \ge 0$, define $p_{t_1,\ldots,t_n}(r_1, \ldots, r_n) = p_{t_{\pi(1)},\ldots,t_{\pi(n)}}(r_{\pi(1)}, \ldots, r_{\pi(n)})$ where π is the permutation such that $t_{\pi(1)} < t_{\pi(2)} < \cdots < t_{\pi(n)}$.

   (a) Show that these probabilities determine a consistent family of finite-dimensional distributions that define a stochastic process $(X_t : t \in [0, \infty))$, where $P(X_t \in \{0, 1, 2, \ldots\}) = 1$ for all $t \ge 0$.

   A finite-dimensional distribution can be shown to exist if a stochastic process satisfies symmetry and marginalization. Symmetry can be shown to exist as follows

   if

$$p_{t_1,\ldots,t_n}(r_1, \ldots, r_n) = p_{t_{\pi(1)},\ldots,t_{\pi(n)}}(r_{\pi(1)}, \ldots, r_{\pi(n)})$$

   and

$$p_{t_{\zeta(1)},\ldots,t_{\zeta(n)}}(r_{\zeta(1)}, \ldots, r_{\zeta(n)} = p_{t_{\eta(\zeta_1)},\ldots,t_{\eta(\zeta_n)}}(r_{\eta(\zeta_1)}, \ldots, r_{\eta(\zeta_n)})$$

   then

$$p_{t_1,\ldots,t_n}(r_1, \ldots, r_n) = p_{t_{\phi(1)},\ldots,t_{\phi(n)}}(r_{\phi(1)}, \ldots, r_{\phi(n)})$$

   Where ϕ is any permutation function as it can be seen that any permuation which is ordered is equal to itself in probability and therefore all permutations are equal to each other.

   The process can be shown to satisy marginlization as it is a poisson point process where the individual terms are independent and thus the following holds.

$$\begin{aligned}
&P(X_{a_1} \in A_1, ..., X_{a_n} \in A_n, X_{a_{n+1}} \in \mathbb{R}) \\
&= P(X_{a_1} \in A_1, ..., X_{a_n} \in A_n) P(X_{a_{n+1}} \in \mathbb{R}) \\
&= P(X_{a_1} \in A_1, ..., X_{a_n} \in A_n) \int_0^\infty f(x)dx \\
&= P(X_{a_1} \in A_1, ..., X_{a_n} \in A_n)
\end{aligned}$$

   (b) What is the joint distribution of X_t and $X_{t+s} - X_t$ for $s, t > 0$?

   Because this is a consistent family of finite-dimensional distributions. Then

$$P(X_{t+s} - X_t = k) = (X_s - X_0 = k) = P(X_s = k)$$

therefore

$$\begin{aligned}P(X_t = x, X_{t+s} - X_t = k) &= P(X_t = x, X_s = k)\\ &= \frac{\lambda^k s^k e^{-\lambda s}}{k!}\frac{\lambda^x t^x e^{-\lambda t}}{x!}\\ &= \frac{\lambda^{k+x} s^k t^x e^{-\lambda(s+t)}}{x!k!}\end{aligned}$$

2. *Conditional independence*

   (a) Let (X_1, X_2, X_3) be a multivariate normal random vector with positive definite covariance matrix $\Sigma = (\sigma_{ij})$ and correlations $\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$. Show that $X_1 \perp\!\!\!\perp X_2 \mid X_3$ if and only if $\rho_{12} = \rho_{13}\rho_{23}$.
   if

$$X_1 \perp\!\!\!\perp X_2 | X_3$$

   then

$$\begin{aligned}\Sigma^{-1}_{1,2} = 0 &= \frac{\sigma_{1,2}\sigma_{3,3} - \sigma_{1,3}\sigma_{2,3}}{Det(\Sigma)}\\ &= \sigma_{1,2}\sigma_{3,3} - \sigma_{1,3}\sigma_{2,3}\\ \sigma_{1,2}\sigma_{3,3} &= \sigma_{1,3}\sigma_{2,3}\\ \rho_{1,2}\sqrt{\sigma_{1,1}\sigma_{2,2}}\sigma_{3,3} &= \rho_{1,3}\sqrt{\sigma_{1,1}\sigma_{3,3}}\rho_{2,3}\sqrt{\sigma_{2,2}\sigma_{3,3}}\\ \rho_{1,2} &= \rho_{1,3}\rho_{2,3}\end{aligned}$$

   (b) Let (X_1, X_2, X_3, X_4) be a multivariate normal random vector with positive definite covariance matrix $\Sigma = (\sigma_{ij})$. Use the fact from (a) to derive that

$$X_1 \perp\!\!\!\perp X_2 \mid X_3,\ X_2 \perp\!\!\!\perp X_3 \mid X_4,\ X_3 \perp\!\!\!\perp X_4 \mid X_1,\ X_1 \perp\!\!\!\perp X_4 \mid X_2 \quad \Longrightarrow \quad X_1 \perp\!\!\!\perp X_2.$$

$$\begin{aligned}\Sigma_{1,2} &= \frac{\sigma^{-1}_{1,2}\sigma^{-1}_{3,4}\sigma^{-1}_{3,4} + \sigma^{-1}_{1,3}\sigma^{-1}_{2,3}\sigma^{-1}_{4,4} + \sigma^{-1}_{1,4}\sigma^{-1}_{3,3}\sigma^{-1}_{2,4} - \sigma^{-1}_{1,2}\sigma^{-1}_{3,3}\sigma^{-1}_{4,4} - \sigma^{-1}_{1,3}\sigma^{-1}_{3,4}\sigma^{-1}_{2,4} - \sigma^{-1}_{1,4}\sigma^{-1}_{2,3}\sigma^{-1}_{3,4}}{det(\Sigma^{-1})}\\ &= \frac{0 * 0 * \sigma^{-1}_{3,4} + \sigma^{-1}_{1,3}0\sigma^{-1}_{4,4} + 0\sigma^{-1}_{3,3}\sigma^{-1}_{2,4} - 0\sigma^{-1}_{3,3}\sigma^{-1}_{4,4} - \sigma^{-1}_{1,3}0\sigma^{-1}_{2,4} - 0 * 0\sigma^{-1}_{3,4}}{det(\Sigma^{-1})}\\ &= 0\\ &\therefore\\ &X_1 \perp\!\!\!\perp X_2\end{aligned}$$

   Alternatively, given the independence statements made above we have

$$\begin{aligned}\rho_{12} &= \rho_{13}\rho_{23}\\ \rho_{23} &= \rho_{24}\rho_{34}\\ \rho_{34} &= \rho_{13}\rho_{14}\\ \rho_{14} &= \rho_{12}\rho_{24}\\ \rho_{12}(\rho_{13}^2\rho_{24}^2 - 1) &= 0\end{aligned}$$

We know that $(\rho_{13}^2\rho_{24}^2 - 1)$ can not be zero because $|\rho_{13}|$ and $|\rho_{24}|$ would need to be 1 which would imply linear dependence in Σ. Therfore, ρ_{12} must be 0 which implies $X_1 \perp\!\!\!\perp X_2$.

(c) 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).

for $0 < p < 1$ and $p \neq .5$

if

$$\begin{aligned} Z &\perp\!\!\!\perp Y \\ P(Z = z) &= .5 \\ P(Y = y) &= .5 \end{aligned}$$

$$P(X = x, Z = x, Y = y) = \begin{cases} .25p^x(1-p)^{1-x} & \text{if } Z == Y, \\ .25(1-p)^x p^{1-x} & \text{otherwise.} \end{cases}$$

then

$$\begin{aligned} P(X = x) &= .5 \\ P(X = x|Z) &= .5 \\ P(X = x|Y) &= .5 \end{aligned}$$

and

$$P(X = x|(Z, Y)) = \begin{cases} p^x(1-p)^{1-x} & \text{if } Z == Y, \\ (1-p)^x p^{1-x} & \text{otherwise.} \end{cases}$$

3. *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

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

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

If we take $T_1 = \{n+1\}$ and $T_0 = \{0...n\}$ we have

$$Pr(X_{n+1} = i_{n+1}|X_n = i_n ... X_0 = i_0) = Pr(X_{n+1} = i_{n+1}|X_n = i_n)$$

Which is equivalent to the definition of a markov chain.

This random vector obeys the Markov property as shown by the following graph.

$$X_L, L \in T_0, L \neq t_0 \qquad X_{t_0} \qquad X_k, k \in T_1$$

There is no path between T_0 and T_1 with out going through x_{t_0} for the entire set of values that T_0, T_1 and t_0 can take and therefore any value that they take in their respective sets must also hold the same property.

(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

$$\Pr(X_k = i_k \,\forall k \in T_1,\ X_l = i_l \,\forall l \in T_0 \,|\, X_n = i_n)$$
$$= \Pr(X_k = i_k \,\forall k \in T_1 \,|\, X_n = i_n) \Pr(X_l = i_l \,\forall l \in T_0 \,|\, X_n = i_n).$$

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

If we take $T_1 = \{n+1\}$ and $T_0 = \{0, ..., n-1\}$, we have

$$Pr(X_{n+1} = i_{n+1}, X_{n-1} = i_{n-1}, ..., X_0 = i_0|X_n = i_n) =$$
$$Pr(X_{n+1} = i_{n+1}|X_n = i_n)Pr(X_{n-1} = i_{n-1}, ..., X_0 = i_0|X_n = i_n)$$

Which is equivalent to the definition of a Markov chain because it implies

$$X_k, k \in T_0 \perp\!\!\!\perp X_L, L \in T_1 | X_n$$

The random vector obeys the Markov property as shown by the following graph:

$$X_L, L \in T_0 \qquad X_n \qquad X_k, k \in T_1$$

Again there is no path between T_0 and T_1 without passing through X_n