PSM Spring 2019, homework 5

1. (The following probability numbers are my invention)

   - The probability that an adult person is suffering from schizophrenia is 0.01.
   - The probability that an adult person wearing a hearing aid hears voices is 0.1.
   - The probability that a schizophrenic adult person who has a hearing aid hears voices is 0.99.
   - The probability that an adult person has a hearing aid is independent of whether this person is schizophrenic.

   Mr. Thompson, who wears a hearing aid, hears voices. What is the probability that he suffers from schizophrenia?

2. Let $\Omega = \{\omega_1, \omega_2\}$ and $X, Y : \Omega \to \{0, 1\}$. Specify a probability measaure P on Ω and concrete values $X(\omega_i), Y(\omega_i)$ (where $i = 1, 2$) such that X, Y are identically distributed but not for all $\omega \in \Omega : X(\omega) = Y(\omega)$.

3. Verify the claim made in the last bullet point in the list after Definition 8.2.1 in the lecture notes.

4. Prove the equivalence of Definitions 12.0.1 and 12.0.2 in the lecture notes. *Note 1:* old version of the lecture notes had a wrong Definition 12.0.1. Use the definition given in the latest release (published online March 29). *Note 2:* This requires some work. My solution uses 1 full page of Latex-formatted formulas. But it all can be done with elementary transformations of conditional probabilities. If you master this one, you know how to play with those guys. If you don't find a way after some effort, check out the first half of the solution (which proves the first direction of the claimed equivalence). The "tricks" that you find there can be used also for the second direction - try your best. *Hint:* for both directions it is helpful to prove that a third equivalent definition of the Markov chain property follows from either 12.0.1 or 12.0.2, namely,

$$P(x_0, x_1, \ldots, x_n) = P(x_0)P(x_1 \mid x_0)P(x_2 \mid x_1) \cdots P(x_n \mid x_{n-1}).$$

Solution to 1. With an obvious dirty notation, we compute

$$\begin{aligned}
P(S \mid V, H) &= \frac{P(V, H \mid S)\; P(S)}{P(V, H)} \quad \text{;;; Bayes' rule} \\
&= \frac{P(H \mid S)\; P(V \mid H, S)\; P(S)}{P(V \mid H)\; P(H)} \\
&= \frac{P(V \mid H, S)\; P(S)}{P(V \mid H)} \quad \text{;;; independence of H from S} \\
&= \frac{0.99 \cdot 0.01}{0.1} \\
&= 0.099
\end{aligned}$$

Solution to 2. The obvious solution is $X(\omega_1) = 0, X(\omega_2) = 1, Y(\omega_1) = 1, X(\omega_1) = 0$ (or its mirror version).

Solution to 3. ... not given, — just a little mechanical verification.

Solution to 4. First, 12.0.1 implies 12.0.2: Assume that 12.0.1 holds, that is, for all $n \geq 2, 0 < m < n, (x_0, \ldots, x_n) \in S^{n+1}$:

$$P(x_0, x_1, \ldots, x_{m-1},\ x_{m+1}, \ldots, x_n \mid x_m) = P(x_0, \ldots, x_{m-1} \mid x_m) \cdot P(x_{m+1}, \ldots, x_n \mid x_m). \quad (1)$$

Using (1) we first show that for all n:

$$P(x_0, x_1, \ldots, x_n) = P(x_0, \ldots, x_{n-1}) P(x_n \mid x_{n-1}). \quad (2)$$

This is derived via

$$\begin{aligned} P(x_0, x_1, \ldots, x_n) &= P(x_{n-1}) P(x_0, \ldots, x_{n-2}, x_n \mid x_{n-1}) \\ &=^1 P(x_{n-1}) P(x_0, \ldots, x_{n-2} \mid x_{n-1}) P(x_n \mid x_{n-1}) \\ &= P(x_0, \ldots, x_{n-1}) P(x_n \mid x_{n-1}), \end{aligned}$$

where for $=^1$ (1) was used. An iterated application of this argument on the left factor in the r.h.s. of (2) gives

$$P(x_0, x_1, \ldots, x_n) = P(x_0) P(x_1 \mid x_0) P(x_2 \mid x_1) \cdots P(x_n \mid x_{n-1}). \quad (3)$$

From (3) we obtain 12.0.2 by

$$\begin{aligned} P(x_{n+1} \mid x_0, \ldots, x_n) &= \frac{P(x_0, \ldots, x_n, x_{n+1})}{P(x_0, \ldots, x_n)} \\ &= \frac{P(x_0) P(x_1 \mid x_0) P(x_2 \mid x_1) \cdots P(x_n \mid x_{n-1}) P(x_{n+1} \mid x_n).}{P(x_0) P(x_1 \mid x_0) P(x_2 \mid x_1) \cdots P(x_n \mid x_{n-1})} \\ &= P(x_{n+1} \mid x_n). \end{aligned}$$

To show that 12.0.2 implies 12.0.1, we first show that 12.0.2 implies (3). This follows directly from the factorization formula for conditional probabilities:

$$\begin{aligned} & P(x_0, x_1, \ldots, x_n) = \\ &= P(x_0) P(x_1 \mid x_0) P(x_2 \mid x_0, x_1) \cdots P(x_2 \mid x_0, x_1, \ldots, x_{n-1}) \\ &= P(x_0) P(x_1 \mid x_0) P(x_2 \mid x_1) \cdots P(x_n \mid x_{n-1}). \end{aligned}$$

Next we show an auxiliary formula

$$P(x_0, \ldots, x_{m-1} \mid x_m) = P(x_0 \mid x_1) P(x_1 \mid x_2) \cdots P(x_{m-1} \mid x_m) \quad (4)$$

by

$$
\begin{aligned}
P(x_0,\ldots,x_{m-1} \mid x_m) &= P(x_m \mid x_0,\ldots,x_{m-1})\frac{P(x_0,\ldots,x_{m-1})}{P(x_m)} \\
&=^2 P(x_m \mid x_{m-1})\frac{P(x_0,\ldots,x_{m-1})}{P(x_m)} \\
&=^3 P(x_m \mid x_{m-1})\frac{P(x_0)P(x_1 \mid x_0)P(x_2 \mid x_1)\cdots P(x_{m-1} \mid x_{m-2})}{P(x_m)} \\
&= P(x_m \mid x_{m-1})\frac{P(x_0,x_1)P(x_2 \mid x_1)\cdots P(x_{m-1} \mid x_{m-2})}{P(x_m)} \\
&= P(x_m \mid x_{m-1})\frac{P(x_0 \mid x_1)P(x_1)P(x_2 \mid x_1)\cdots P(x_{m-1} \mid x_{m-2})}{P(x_m)} \\
&= \ldots \\
&= P(x_m \mid x_{m-1})\frac{P(x_0 \mid x_1)P(x_1 \mid x_2)\cdots P(x_{m-2} \mid x_{m-1})\ P(x_{m-1})}{P(x_m)} \\
&= P(x_0 \mid x_1)P(x_1 \mid x_2)\cdots P(x_{m-1} \mid x_m),
\end{aligned}
$$

where in $=^2$ we used 12.0.2 and in $=^3$ we used (3). Using (4) we now conclude

$$
\begin{aligned}
&P(x_0,x_1,\ldots,x_{m-1},\ x_{m+1},\ldots,x_n \mid x_m) = \\
&= \frac{P(x_0,\ldots,x_n)}{P(x_m)} =^4 \frac{P(x_0)P(x_1 \mid x_0)\cdots P(x_n \mid x_{n-1})}{P(x_m)} \\
&= \ldots \text{(similar as in the "..." above)} \\
&= \frac{P(x_0 \mid x_1)\cdots P(x_{m-1} \mid x_m)\ P(x_m)\ P(x_{m+1} \mid x_m)\cdots P(x_n \mid x_{n-1})}{P(x_m)} \\
&= P(x_0 \mid x_1)\cdots P(x_{m-1} \mid x_m)\ P((x_{m+1}\cdots x_n \mid x_m) \\
&=^5 P(x_0,\ldots,x_{m-1} \mid x_m)\ P(x_{m+1}\cdots x_n \mid x_m),
\end{aligned}
$$

where $=^5$ inserts (4).