Midterm practice problems

Solutions

1. (Babysitter)

   (a)

$$
\begin{aligned}
p_{\tilde{w}_1}(1) &= \sum_{x=0}^{1}\sum_{b=0}^{1} p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(x,b,1) && (1)\\
&= \sum_{x=0}^{1}\sum_{b=0}^{1} p_{\tilde{x}}(x)\, p_{\tilde{b}_1}(b_1)\, p_{\tilde{w}_1\,|\,\tilde{x},\tilde{b}_1}(1\,|\,x,b) && (2)\\
&= 0.1 + 0.9\cdot(0.6\cdot 0.1 + 0.4\cdot 0.8) && (3)\\
&= 0.442. && (4)
\end{aligned}
$$

   (b)

$$
\begin{aligned}
p_{\tilde{x}\,|\,\tilde{w}_1}(1\,|\,1) &= \frac{p_{\tilde{x},\tilde{w}_1}(1,1)}{p_{\tilde{w}_1}(1)} && (5)\\
&= \frac{p_{\tilde{x}}(1)\, p_{\tilde{w}_1\,|\,\tilde{x}}(1\,|\,1)}{0.442} && (6)\\
&= \frac{0.1}{0.442} && (7)\\
&= 0.226. && (8)
\end{aligned}
$$

   (c) We have:

$$
\begin{aligned}
p_{\tilde{x}\,|\,\tilde{w}_1,\tilde{b}_1}(1\,|\,1,1) &= \frac{p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(1,1,1)}{p_{\tilde{w}_1,\tilde{b}_1}(1,1)} && (9)\\
&= \frac{p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(1,1,1)}{\sum_{x=0}^{1} p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(x,1,1)} && (10)\\
&= \frac{p_{\tilde{x}}(1)\, p_{\tilde{b}_1}(1)\, p_{\tilde{w}_1\,|\,\tilde{x},\tilde{b}_1}(1\,|\,1,1)}{p_{\tilde{b}_1}(1)\Big(p_{\tilde{x}}(1)\, p_{\tilde{w}_1\,|\,\tilde{x},\tilde{b}_1}(1\,|\,1,1) + p_{\tilde{x}}(0)\, p_{\tilde{w}_1\,|\,\tilde{x},\tilde{b}_1}(1\,|\,0,1)\Big)} && (11)\\
&= \frac{p_{\tilde{x}}(1)\, p_{\tilde{w}_1\,|\,\tilde{x},\tilde{b}_1}(1\,|\,1,1)}{p_{\tilde{x}}(1)\, p_{\tilde{w}_1\,|\,\tilde{x},\tilde{b}_1}(1\,|\,1,1) + p_{\tilde{x}}(0)\, p_{\tilde{w}_1\,|\,\tilde{x},\tilde{b}_1}(1\,|\,0,1)} && (12)\\
&= \frac{0.1}{0.1 + 0.9\cdot 0.1} && (13)\\
&= 0.526 && (14)
\end{aligned}
$$

   so $\tilde{x}$ and $\tilde{b}_1$ are not conditionally independent given $\tilde{w}_1$ because if they were this would equal $p_{\tilde{x}\,|\,\tilde{w}_1}(1\,|\,1)$. This makes sense because if we know that the baby has woken up, whether the food is bad or not provides information about whether the baby is a good sleeper (and vice versa). In particular, conditioned on $\tilde{w}_1 = 1$, if the baby is good, then the food is more likely to be bad.

2. (Nuclear power plant)

(a) The pdf should integrate to one. We have

$$\int_{-\infty}^{\infty} f_{\tilde{t}}(t)\ \mathrm{d}t = \int_{-1}^{0} \alpha\, \mathrm{d}t + \int_{0}^{\infty} \alpha \exp(t)\ \mathrm{d}t \tag{15}$$

$$= \alpha(0 - (-1)) + \alpha(\exp(0) - \exp -\infty) \tag{16}$$

$$= 2\alpha, \tag{17}$$

so $\alpha = 1/2$.

(b) We compute the cdf of $\tilde{t}$ conditioned on the event $\{\tilde{t} < 0\}$:

$$F_{\tilde{t}\,|\,\tilde{t}<0}(t) := \mathrm{P}(\tilde{t} \leq t \,|\, \tilde{t} < 0) \tag{18}$$

$$= \frac{\mathrm{P}(0 < \tilde{t} \leq \min\{0, t\})}{\mathrm{P}(\tilde{t} < 0)} \tag{19}$$

$$= \begin{cases} 0 & \text{if } t < -1, \\ \frac{\int_{a=-1}^{t} f_{\tilde{t}}(a)\,\mathrm{d}a}{\int_{a=-1}^{0} f_{\tilde{t}}(a)\,\mathrm{d}a} & \text{if } -1 \leq t < 0, \\ 1 & \text{if } t \geq 0. \end{cases} \tag{20}$$

For $-1 \leq t < 0$, we have

$$\int_{a=-1}^{t} f_{\tilde{t}}(a)\,\mathrm{d}a = \frac{t+1}{2}, \tag{21}$$

so

$$F_{\tilde{t}\,|\,\tilde{t}<0}(t) = \begin{cases} 0 & \text{if } t < -1, \\ 1+t & \text{if } -1 \leq t < 0, \\ 1 & \text{if } t \geq 0. \end{cases} \tag{22}$$

Differentiating, we obtain

$$f_{\tilde{t}\,|\,\tilde{t}<0}(t) = \begin{cases} 0 & \text{if } t < -1, \\ 1 & \text{if } -1 \leq t < 0, \\ 0 & \text{if } t \geq 0. \end{cases} \tag{23}$$

(c) We apply the inverse-transform method. The cdf equals

$$F_{\tilde{t}}(t) = \int_{-\infty}^{t} f_{\tilde{t}}(t)\ \mathrm{d}t = \begin{cases} 0 & \text{if } t < -1, \\ \frac{t+1}{2} & \text{if } -1 \leq t \leq 0, \\ \frac{1}{2} + \frac{1-\exp(-t)}{2} & \text{if } 0 \leq t \leq \infty. \end{cases} \tag{24}$$

The inverse is

$$F_{\tilde{t}}^{-1}(u) = \begin{cases} 2u-1 & \text{if } 0 \leq u \leq \frac{1}{2}, \\ \log\left(\frac{1}{2(1-u)}\right) & \text{if } \frac{1}{2} \leq u \leq 1. \end{cases} \tag{25}$$

We obtain $F_{\tilde{t}}^{-1}(0.3) = -0.4$ and $F_{\tilde{t}}^{-1}(0.8) = \log 2 = 0.693$.

3. (Earthquake)

(a) By conditional independence of $\tilde{s}$ and $\tilde{e}$ given $\tilde{v}$

$$
\begin{align}
p_{\tilde{s}}(1) &= \sum_{e=0}^{1} \sum_{v=0}^{2} p_{\tilde{e},\tilde{v}}(e,v) p_{\tilde{s}\,|\,\tilde{e},\tilde{v}}(1\,|\,e,v) \tag{26}\\
&= \sum_{e=0}^{1} \sum_{v=0}^{2} p_{\tilde{e},\tilde{v}}(e,v) p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,v) \tag{27}\\
&= p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,1)(p_{\tilde{e},\tilde{v}}(0,1) + p_{\tilde{e},\tilde{v}}(1,1)) + p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,2)(p_{\tilde{e},\tilde{v}}(0,2) + p_{\tilde{e},\tilde{v}}(1,2)) \tag{28}\\
&= 0.5(0.05 + 0.05) + 0.1 \tag{29}\\
&= 0.15 \tag{30}
\end{align}
$$

and consequently $p_{\tilde{s}}(0) = 1 - p_{\tilde{s}}(1) = 0.85$.

(b)

$$
\begin{align}
p_{\tilde{e}\,|\,\tilde{s}}(1\,|\,1) &= \frac{p_{\tilde{e},\tilde{s}}(1,1)}{p_{\tilde{s}}(1)} \tag{31}\\
&= \frac{\sum_{v=0}^{2} p_{\tilde{e},\tilde{v},\tilde{s}}(1,v,1)}{p_{\tilde{s}}(1)} \tag{32}\\
&= \frac{\sum_{v=0}^{2} p_{\tilde{e},\tilde{v},\tilde{s}}(1,v) p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,v)}{p_{\tilde{s}}(1)} \tag{33}\\
&= \frac{p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,1) p_{\tilde{e},\tilde{v}}(1,1) + p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,2) p_{\tilde{e},\tilde{v}}(1,2)}{p_{\tilde{s}}(1)} \tag{34}\\
&= \frac{0.5 \cdot 0.05 + 0.1}{0.15} \tag{35}\\
&= 0.833. \tag{36}
\end{align}
$$

(c) We have

$$
\begin{align}
p_{\tilde{e}}(1) &= \sum_{v=0}^{2} p_{\tilde{e},\tilde{v}}(1,v) \tag{37}\\
&= 0.05 + 0.1 \tag{38}\\
&= 0.15 \neq p_{\tilde{e}\,|\,\tilde{s}}(1\,|\,1), \tag{39}
\end{align}
$$

so they are not independent. This makes sense, because the sensor reading is more likely to be one if there are vibrations, which is more likely if there is an earthquake.

4. (Rat)

(a) The pdf needs to be nonnegative, which requires $0 \leq \alpha \leq 1$. We also need the pdf to integrate to one, which it does if α is in that range:

$$\int_{a=0}^{1} f_{\tilde{a}}(a)\,\mathrm{d}a = \int_{a=0}^{0.5} 2\alpha\,\mathrm{d}a + \int_{a=0.5}^{1} 2(1-\alpha)\,\mathrm{d}a \tag{40}$$
$$= 2\alpha \cdot 0.5 + 2(1-\alpha)\cdot 0.5 \tag{41}$$
$$= 1. \tag{42}$$

(b) Expressing the pdf as a function of α the likelihood of each data point is equal to 2α if the point is between 0 and 0.5, and to $2(1-\alpha)$ if it is between 0.5 and 1. Let n be the number of data, $n_{[0,0.5]}$ the number of data between 0 and 0.5, and $n_{(0.5,1]}$ the number of points between 0.5 and 1. We have,

$$\log \mathcal{L}(X) = \sum_{i=1}^{n} \log f_{\alpha}(x_i) \tag{43}$$
$$= n_{[0,0.5]} \log 2\alpha + n_{(0.5,1]} \log(2(1-\alpha)). \tag{44}$$

The first and second derivatives of the log likelihood equal

$$(\log \mathcal{L}(X))' = \frac{n_{[0,0.5]}}{\alpha} - \frac{n_{(0.5,1]}}{1-\alpha}, \tag{45}$$
$$(\log \mathcal{L}(X))'' = -\frac{n_{[0,0.5]}}{\alpha^2} - \frac{n_{(0.5,1]}}{(1-\alpha)^2}. \tag{46}$$

The function is concave, so we can set the first derivative to zero to find the ML estimate, it equals

$$\alpha_{ML} = \frac{n_{[0,0.5]}}{n} \tag{47}$$
$$= \frac{2}{5}. \tag{48}$$

(c) The probability of the rat being in the first half is

$$\mathrm{P}(0 \leq \tilde{a} \leq 0.5) = \int_{a=0}^{0.5} 2\alpha\,\mathrm{d}a \tag{49}$$
$$= \alpha. \tag{50}$$

Estimating this probability using empirical probabilities yields exactly the same estimate for α as the ML estimate,

$$\hat{\alpha_{ML}} = \frac{n_{[0,0.5]}}{n}. \tag{51}$$

(d) We compute the cdf and then invert it. The cdf equals 0 for $a < 0$,

$$F_{\tilde{a}}(a) = \int_{0}^{a} 0.4\,\mathrm{d}a \tag{52}$$
$$= 0.4a \tag{53}$$

for $0 \leq a \leq 0.5$,

$$F_{\tilde{a}}(a) = \int_0^{0.5} 0.4 \, \mathrm{d}a + \int_{0.5}^{a} 1.6 \, \mathrm{d}a = \tag{54}$$
$$= 0.2 + 1.6(a - 0.5) \tag{55}$$
$$= 1.6a - 0.6 \tag{56}$$

for $0.5 \leq a \leq 1$, and 1 for $a \geq 1$. Setting

$$F_{\tilde{a}}(a) = 0.4 \tag{57}$$
$$= 1.6a - 0.6, \tag{58}$$

yields a sample equal to $1/1.6 = 0.625$.

5. (Election)

(a)

$$p_{\tilde{o}}(1) = \mathrm{P}(\text{A wins all states}) + \sum_{i=1}^{3} \mathrm{P}(\text{A wins all states except } i) \tag{59}$$
$$= 0.6^3 + 3 \cdot 0.4 \cdot 0.6^2 \tag{60}$$
$$= 0.648. \tag{61}$$

(b) We have

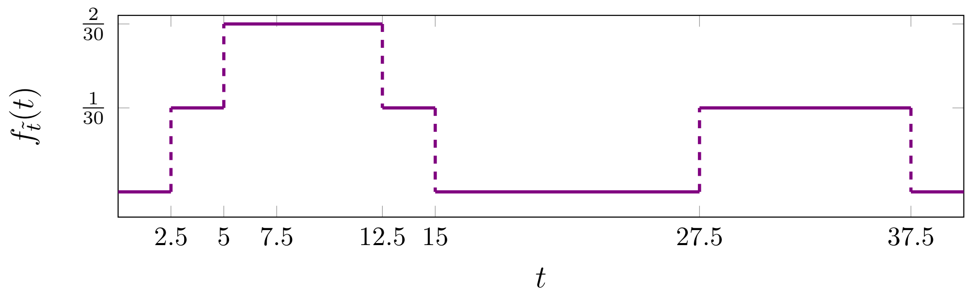
$$p_{\tilde{s}_1 \,|\, \tilde{o}}(1 \,|\, 1) \tag{62}$$
$$= \frac{p_{\tilde{s}_1, \tilde{o}}(1, 1)}{p_{\tilde{o}}(1)} \tag{63}$$
$$= \frac{\mathrm{P}(\text{A wins all states}) + \mathrm{P}(\text{A wins 1 and 2 but not 3}) + \mathrm{P}(\text{A wins 1 and 3 but not 2})}{p_{\tilde{o}}(1)}$$
$$= \frac{0.6^3 + 2 \cdot 0.4 \cdot 0.6^2}{0.648} \tag{64}$$
$$= 0.778. \tag{65}$$

However, $p_{\tilde{s}_1 \,|\, \tilde{o}, \tilde{s}_2}(1 \,|\, 1, 0) = 1$ because if A wins the election then they cannot lose state 1 and state 2. Intuitively, even if the state results are independent, they both determine the result of the election, so revealing who wins the election *connects* them. For example, if candidate A has won the election but lost state 2, then this completely determines the result of state 1.

(c) To compute the probability, we need to aggregate all the possible ways in which a candidate can win. This requires considering the 2^{50} possible outcomes, which may become infeasible computationally. Instead, we can apply the Monte Carlo methods. We sample the result from each state and aggregate them to get a sample from the overall result of the election. We then repeat the process many times and use the fraction of times a candidate wins to estimate the corresponding probability.

6. (Another earthquake)

(a) The KDE estimate looks like this:



The probability is equal to

$$P(\tilde{t} > 10) = \frac{2 \cdot 2.5}{30} + \frac{2.5}{30} + \frac{10}{30} \tag{66}$$
$$= \frac{17.5}{30} \tag{67}$$
$$= 0.583, \tag{68}$$

which is the area under the pdf to the right of 10.

(b) From the notes, the ML estimate for the parameter of the exponential is

$$\lambda_{\text{ML}} = \frac{3}{7.5 + 10 + 32.5} \tag{69}$$
$$= \frac{3}{50}. \tag{70}$$

The probability equals

$$P(\tilde{t} > 10) = \int_{10}^{\infty} f_{\tilde{t}}(t)\,\mathrm{d}t \tag{71}$$
$$= \int_{10}^{\infty} \frac{3\exp(-3t/50)}{50}\,\mathrm{d}t \tag{72}$$
$$= -\exp(-3t/50)]_{10}^{\infty} \tag{73}$$
$$= 0.549. \tag{74}$$

(c) The parametric method requires less data but makes a stronger assumption about the distribution. The nonparametric method is more flexible but requires more data.

7. (Scaling random variables)

(a) False. Recall that the cdf of $\tilde{a}$ equals $F_{\tilde{a}}(a) = 1 - \exp(-\lambda a)$ for $a \geq 0$ and 0 for

$a < 0$. The cdf of $\tilde{b}$ equals

$$F_{\tilde{b}}(b) := \mathrm{P}(\tilde{b} \leq b) \tag{75}$$
$$= \mathrm{P}(\alpha\tilde{a} \leq b) \tag{76}$$
$$= \mathrm{P}\left(\tilde{a} \leq \frac{b}{\alpha}\right) \tag{77}$$
$$= F_{\tilde{a}}\left(\frac{b}{\alpha}\right) \tag{78}$$
$$= 1 - \exp\left(-\frac{\lambda}{\alpha}b\right), \tag{79}$$

so $\tilde{b}$ is actually an exponential random variable with parameter λ/α.

(b) False. If $\alpha = 2$ then $\tilde{b}$ only takes even values, and therefore is not a geometric random variable.

(c) True. If the median of a $\tilde{a}$ is m, then

$$\mathrm{P}(\tilde{a} \leq m) = \frac{1}{2}, \tag{80}$$

which implies

$$\mathrm{P}(\tilde{b} \leq \alpha m + \beta) = \mathrm{P}(\alpha\tilde{a} + \beta \leq \alpha m + \beta) \tag{81}$$
$$= \mathrm{P}(\tilde{a} \leq m). \tag{82}$$

8. (Spam detector)

(a) The problem is that there are $2^4 = 16$ different possible values for the entries of $\tilde{x}$, and hence 16 different conditional distributions. However, we only have 10 data points.

(b) The email corresponds to $\tilde{x}[1] = 1$, $\tilde{x}[2] = 0$, $\tilde{x}[3] = 1$, $\tilde{x}[4] = 1$. By the naive Bayes assumption combined with empirical estimates of $p_{\tilde{y}}$ and $p_{\tilde{x}[i]\,|\,\tilde{y}}$

$$p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,1,0,1,1) = p_{\tilde{y}}(1)p_{\tilde{x}[1]\,|\,\tilde{y}}(1\,|\,1)p_{\tilde{x}[2]\,|\,\tilde{y}}(0\,|\,1)p_{\tilde{x}[3]\,|\,\tilde{y}}(1\,|\,1)p_{\tilde{x}[4]\,|\,\tilde{y}}(1\,|\,1)$$
$$= \frac{5}{10}\frac{4}{5}\frac{3}{5}\frac{2}{5}\frac{1}{5}, \tag{83}$$
$$p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(0,1,0,1,1) = p_{\tilde{y}}(0)p_{\tilde{x}[1]\,|\,\tilde{y}}(1\,|\,0)p_{\tilde{x}[2]\,|\,\tilde{y}}(0\,|\,0)p_{\tilde{x}[3]\,|\,\tilde{y}}(1\,|\,0)p_{\tilde{x}[4]\,|\,\tilde{y}}(1\,|\,0)$$
$$= \frac{5}{10}\frac{1}{5}\frac{4}{5}\frac{2}{5}\frac{4}{5}. \tag{84}$$

As a result,

$$
\begin{aligned}
& p_{\tilde{y} \mid \tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1 \mid 1,0,1,1) && (85)\\
&= \frac{p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,1,0,1,1)}{p_{\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,0,1,1)} && (86)\\
&= \frac{p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,1,0,1,1)}{p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,1,0,1,1) + p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(0,1,0,1,1)} && (87)\\
&= \frac{4 \cdot 3 \cdot 2}{4 \cdot 3 \cdot 2 + 4 \cdot 2 \cdot 4} && (88)\\
&= \frac{3}{7}, && (89)
\end{aligned}
$$

so $p_{\tilde{y} \mid \tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(0 \mid 1,0,1,1) = \frac{4}{7}$ and we classify the email as not spam.

(c) The email corresponds to $\tilde{x}[1] = 0$, $\tilde{x}[2] = 1$, $\tilde{x}[3] = 1$, $\tilde{x}[4] = 0$. By the naive Bayes assumption combined with empirical estimates of $p_{\tilde{y}}$ and $p_{\tilde{x}[i] \mid \tilde{y}}$

$$
\begin{aligned}
p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,0,1,1,0) &= p_{\tilde{y}}(1) p_{\tilde{x}[1] \mid \tilde{y}}(0 \mid 1) p_{\tilde{x}[2] \mid \tilde{y}}(1 \mid 1) p_{\tilde{x}[3] \mid \tilde{y}}(1 \mid 1) p_{\tilde{x}[4] \mid \tilde{y}}(0 \mid 1)\\
&= \frac{5}{10} \frac{1}{5} \frac{2}{5} \frac{2}{5} \frac{4}{5}, && (90)\\
p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(0,0,1,1,0) &= p_{\tilde{y}}(0) p_{\tilde{x}[1] \mid \tilde{y}}(0 \mid 0) p_{\tilde{x}[2] \mid \tilde{y}}(1 \mid 0) p_{\tilde{x}[3] \mid \tilde{y}}(1 \mid 0) p_{\tilde{x}[4] \mid \tilde{y}}(0 \mid 0)\\
&= \frac{5}{10} \frac{4}{5} \frac{3}{5} \frac{2}{5} \frac{1}{5}. && (91)
\end{aligned}
$$

As a result,

$$
\begin{aligned}
& p_{\tilde{y} \mid \tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1 \mid 0,1,1,0) && (92)\\
&= \frac{p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,0,1,1,0)}{p_{\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(0,1,1,0)} && (93)\\
&= \frac{p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,0,1,1,0)}{p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(1,0,1,1,0) + p_{\tilde{y},\tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(0,1,1,0)} && (94)\\
&= \frac{2 \cdot 2 \cdot 4}{4 \cdot 3 \cdot 2 + 2 \cdot 2 \cdot 4} && (95)\\
&= \frac{2}{5}, && (96)
\end{aligned}
$$

so $p_{\tilde{y} \mid \tilde{x}[1],\tilde{x}[2],\tilde{x}[3],\tilde{x}[4]}(0 \mid 1,0,1,1) = \frac{3}{5}$ and we classify the email as not spam.

In our data the words *alternative* and *medicine* only appear together in spam emails, but the naive Bayes classifier ignores these dependencies.

9. (Noisy data)

(a)

$$
\begin{aligned}
&\mathrm{P}(\tilde{z}=1 \mid \tilde{x}=1) && (97)\\
&= \frac{\mathrm{P}(\tilde{z}=1, \tilde{x}=1)}{\mathrm{P}(\tilde{x}=1)} && (98)\\
&= \frac{\mathrm{P}(\tilde{z}=1, \tilde{x}=1, \tilde{y}=0) + \mathrm{P}(\tilde{z}=1, \tilde{x}=1, \tilde{y}=1)}{\mathrm{P}(\tilde{x}=1)} && (99)\\
&= \frac{\mathrm{P}(\tilde{x}=1)\,\mathrm{P}(\tilde{y}=0 \mid \tilde{x}=1)\,\mathrm{P}(\tilde{z}=1 \mid \tilde{x}=1, \tilde{y}=0) + \mathrm{P}(\tilde{x}=1)\,\mathrm{P}(\tilde{y}=1 \mid \tilde{x}=1)\,\mathrm{P}(\tilde{z}=1 \mid \tilde{x}=1, \tilde{y}=1)}{\mathrm{P}(\tilde{x}=1)}\\
&= \mathrm{P}(\tilde{y}=0 \mid \tilde{x}=1)\,\mathrm{P}(\tilde{z}=1 \mid \tilde{y}=0) + \mathrm{P}(\tilde{y}=1 \mid \tilde{x}=1)\,\mathrm{P}(\tilde{z}=1 \mid \tilde{x}=1, \tilde{y}=1) && (100)\\
&= 0.1 \cdot 0.1 + 0.9 \cdot 0.9 && (101)\\
&= 0.82. && (102)
\end{aligned}
$$

(b) We have $\mathrm{P}(\tilde{z}=1 \mid \tilde{x}=1, \tilde{y}=1) = \mathrm{P}(\tilde{z}=1 \mid \tilde{y}=1) = 0.9 \neq \mathrm{P}(\tilde{z}=1 \mid \tilde{x}=1)$ so they are not conditionally independent given $\tilde{x}$.

(c) By definition of the problem $\mathrm{P}(\tilde{z}=z \mid \tilde{x}=x, \tilde{y}=y) = \mathrm{P}(\tilde{z}=z \mid \tilde{y}=y)$ for any values of x, y and z so $\tilde{z}$ and $\tilde{x}$ are conditionally independent given $\tilde{y}$.

10. (Self-driving car)

(a) It doesn't make sense for the Markov chain to be time-homogeneous, as the probability of people using the cars will vary during the day (being much higher at rush hour for example) and also at different locations.

(b) The only way to go from zero to two occupants in two stops is by having one person come in at each stop. As a result,

$$
\begin{aligned}
p_{\tilde{x}[0] \mid \tilde{x}[2]}(0 \mid 2) &= \frac{p_{\tilde{x}[0],\tilde{x}[2]}(0,2)}{p_{\tilde{x}[2]}(2)} && (103)\\
&= \frac{p_{\tilde{x}[0],\tilde{x}[1],\tilde{x}[2]}(0,1,2)}{\sum_{a=1}^{2}\sum_{b=1}^{2} p_{\tilde{x}[0],\tilde{x}[1]\tilde{x}[2]}(a,b,2)} && (104)\\
&= \frac{p_{\tilde{x}[0],\tilde{x}[1],\tilde{x}[2]}(0,1,2)}{p_{\tilde{x}[0],\tilde{x}[1],\tilde{x}[2]}(0,1,2) + p_{\tilde{x}[0],\tilde{x}[1],\tilde{x}[2]}(1,1,2) + p_{\tilde{x}[0],\tilde{x}[1],\tilde{x}[2]}(1,2,2)} && (105)\\
&= \frac{\frac{1}{2}\frac{1}{2}\frac{1}{4}}{\frac{1}{2}\frac{1}{2}\frac{1}{4} + \frac{1}{2}\frac{1}{2}\frac{1}{4} + \frac{1}{2}\frac{1}{4}\frac{3}{4}} && (106)\\
&= \frac{2}{7}. && (107)
\end{aligned}
$$

The probability is 1/8.

(c) The transition matrix is

$$
T := \begin{bmatrix} 1 - p_{\text{in}} & (1 - p_{\text{in}})\,p_{\text{out}} & 0 \\ p_{\text{in}} & (1 - p_{\text{in}})(1 - p_{\text{out}}) + p_{\text{in}}p_{\text{out}} & (1 - p_{\text{in}})\,p_{\text{out}} \\ 0 & (1 - p_{\text{out}})\,p_{\text{in}} & 1 - p_{\text{out}} + p_{\text{in}}p_{\text{out}} \end{bmatrix}. \tag{108}
$$

(d) For $p_{\text{in}} = 0$ and $p_{\text{out}} = 0$

$$T_{\widetilde{X}} := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{109}$$

In this case customers don't get in or out so the Markov chain just stays in whichever state it starts in.