Midterm practice problems

Solutions

1. (Short questions)

   (a) No.

   (b) Yes, because for fixed values of y and z, $f_{\tilde{x}\,|\,\tilde{y},\tilde{z}}\left(x\,|\,y,z\right)$ is a valid pdf.

   (c) Yes. If $\tilde{u}$ is uniform between 0 and 1 and if conditioned on $\tilde{u}=u$, $\tilde{x}$ is binomial with parameters $n:=10$ and $p:=u$, then the expression is equal to $\int_0^1 f_{\tilde{u}}\left(u\right)\sum_{k=1}^{\infty} p_{\tilde{x}\,|\,\tilde{u}}\left(x\,|\,u\right)\,\mathrm{d}u$. For any fixed u

$$\sum_{k=0}^{20} p_{\tilde{x}\,|\,\tilde{u}}\left(x\,|\,u\right)=1 \tag{1}$$

   so the expression is equal to

$$\begin{aligned}\int_0^1 f_{\tilde{u}}\left(u\right)\sum_{k=0}^{20} p_{\tilde{x}\,|\,\tilde{u}}\left(x\,|\,u\right)\,\mathrm{d}u &= \int_0^1 f_{\tilde{u}}\left(u\right)\,\mathrm{d}u & (2)\\ &=1. & (3)\end{aligned}$$

   (d) To specify the joint pmf we need $3^3-1=26$ parameters. Under the independence assumption we only need $3-1=2$ for $\tilde{x}$ and $\tilde{y}$ and then for $\tilde{z}$ an additional $3-1=2$ for each of the $3^2=9$ possible values of $(\tilde{x},\tilde{y})$. This results in 22 parameters.

   (e)

$$\begin{aligned}F_{\tilde{u}\,|\,\{\tilde{u}\geq 0.5\}}(u) &= \mathrm{P}(\tilde{u}\leq u\,|\;\{\tilde{u}\geq 0.5\}) & (4)\\ &= \frac{\mathrm{P}(0.5\leq\tilde{u}\leq u)}{\mathrm{P}(\tilde{u}\geq 0.5)} & (5)\\ &= 2\mathrm{P}(0.5\leq\tilde{u}\leq u). & (6)\end{aligned}$$

   If $u\leq 0.5$, the numerator is zero, so $F_{\tilde{u}\,|\,\{\tilde{u}\geq 0.5\}}(u)=0$. If $0.5\leq u\leq 1$

$$\begin{aligned}F_{\tilde{u}\,|\,\{\tilde{u}\geq 0.5\}}(u) &= 2\int_{0.5}^{u}\mathrm{d}t & (7)\\ &= 2\left(u-0.5\right). & (8)\end{aligned}$$

   If $u\geq 1$ then the numerator is $1/2$, so $F_{\tilde{u}\,|\,\{\tilde{u}\geq 0.5\}}(u)=1$. Differentiating, we conclude that the density of $\tilde{u}$ conditioned on $\tilde{u}\geq 1/2$ is uniform between $1/2$ and 1.

   (f) In the region $\{(x,y):x\leq 0\}\cup\{(x,y):y\leq 0\}$ $P(\tilde{x}\leq x,\tilde{y}\leq y)=0$ so $F_{\tilde{x},\tilde{y}}(x,y)$ equals 0.
   In the region $\{(x,y):x\geq 10,y\geq 10\}$ $P(\tilde{x}\leq x,\tilde{y}\leq y)=1$ so $F_{\tilde{x},\tilde{y}}(x,y)$ equals 1.

   (g) It is impossible. The joint pmf has $2^{100}-1$ degrees of freedom, which is more than 10^{30}!

(h) We will probably underestimate it because the price of flights and hotels should have positive covariance, and the variance of the sum equals the sum of the variances plus two times the covariance.

2. (Nuclear power plant)

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

$$\int_{-\infty}^{\infty} f_{\tilde{t}}(t)\, \mathrm{d}t = \int_{-1}^{0} \alpha \, \mathrm{d}t + \int_{0}^{\infty} \alpha \exp(t)\, \mathrm{d}t \tag{9}$$

$$= \alpha(0 - (-1)) + \alpha(\exp(0) - \exp -\infty) \tag{10}$$

$$= 2\alpha, \tag{11}$$

so $\alpha = 1/2$.

(b) The cdf equals

$$F_{\tilde{t}}(t) = \int_{-\infty}^{t} f_{\tilde{t}}(t)\, \mathrm{d}t = \begin{cases} 0 & \text{if } t < -1, \\ \frac{t+1}{2} & \text{if } -1 \leq t \leq 0, \\ \frac{1}{2} + \frac{1-\exp(-t)}{2} & \text{if } 0 \leq t \leq \infty. \end{cases} \tag{12}$$

(c) Setting $F_{\tilde{t}}(t) = \frac{t+1}{2} = 0.3$ ($F_{\tilde{t}}$ is smaller than 0.5 when $t \leq 0$, and larger when $t \geq 0$), we obtain $t = -0.4$.

(d) We define $\tilde{s}$ to be Bernoulli with parameter 1/2. We let the conditional pdf of $\tilde{t}$ given $\tilde{s} = 1$ be uniform between -1 and 0, and the conditional pdf given $\tilde{s} = 0$ to be exponential with constant 1. We confirm that

$$f_{\tilde{t}} t = p_{\tilde{s}}(1) f_{\tilde{t}\,|\,\tilde{s}}(t\,|\,1) + p_{\tilde{s}}(0) f_{\tilde{t}\,|\,\tilde{s}}(t\,|\,0) = \begin{cases} 0 & t < -1, \\ \frac{1}{2} \quad \text{if} & -1 \leq t < 0, \\ \frac{1}{2} \exp(-t) & t > 0. \end{cases} \tag{13}$$

(e) By iterated expectation

$$\mathrm{E}(\tilde{t}) = \mathrm{E}\left(\mathrm{E}(\tilde{t}\,|\,\tilde{s})\right) \tag{14}$$

$$= \frac{\mathrm{E}(\tilde{t}\,|\,\tilde{s} = 0) + \mathrm{E}(\tilde{t}\,|\,\tilde{s} = 1)}{2} \tag{15}$$

$$= \frac{-0.5 + 1}{2} \tag{16}$$

$$= 0.25. \tag{17}$$

(f) By iterated expectation

$$\mathrm{E}(\tilde{t}^2) = \mathrm{E}\left(\mathrm{E}(\tilde{t}^2\,|\,\tilde{s})\right) \tag{18}$$

$$= \frac{\mathrm{E}(\tilde{t}^2\,|\,\tilde{s} = 0) + \mathrm{E}(\tilde{t}^2\,|\,\tilde{s} = 1)}{2} \tag{19}$$

$$= \frac{\frac{1}{12} + \left(\frac{1}{2}\right)^2 + 1 + 1}{2} \tag{20}$$

$$= \frac{7}{6} \tag{21}$$

so the variance equals $\mathrm{Var}\left(\tilde{t}\right) = \mathrm{E}(\tilde{t}^2) - \mathrm{E}^2(\tilde{t}) = 1.104$.

3. (Noisy data)

   (a)

$$
\begin{aligned}
&\mathrm{P}\left(\tilde{z}=1 \,|\, \tilde{x}=1\right) && (22)\\
&= \frac{\mathrm{P}\left(\tilde{z}=1, \tilde{x}=1\right)}{\mathrm{P}\left(\tilde{x}=1\right)} && (23)\\
&= \frac{\mathrm{P}\left(\tilde{z}=1, \tilde{x}=1, \tilde{y}=0\right) + \mathrm{P}\left(\tilde{z}=1, \tilde{x}=1, \tilde{y}=1\right)}{\mathrm{P}\left(\tilde{x}=1\right)} && (24)\\
&= \frac{\mathrm{P}\left(\tilde{x}=1\right)\mathrm{P}\left(\tilde{y}=0 \,|\, \tilde{x}=1\right)\mathrm{P}\left(\tilde{z}=1 \,|\, \tilde{x}=1, \tilde{y}=0\right) + \mathrm{P}\left(\tilde{x}=1\right)\mathrm{P}\left(\tilde{y}=1 \,|\, \tilde{x}=1\right)\mathrm{P}\left(\tilde{z}=1 \,|\, \tilde{x}=1, \tilde{y}=1\right)}{\mathrm{P}\left(\tilde{x}=1\right)} \\
&= \mathrm{P}\left(\tilde{y}=0 \,|\, \tilde{x}=1\right)\mathrm{P}\left(\tilde{z}=1 \,|\, \tilde{y}=0\right) + \mathrm{P}\left(\tilde{y}=1 \,|\, \tilde{x}=1\right)\mathrm{P}\left(\tilde{z}=1 \,|\, \tilde{x}=1, \tilde{y}=1\right) && (25)\\
&= 0.1 \cdot 0.1 + 0.9 \cdot 0.9 && (26)\\
&= 0.82. && (27)
\end{aligned}
$$

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

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

4. (Dead fish)

   (a) The position is a uniform random variable since

$$
f_{\tilde{x}} x = p_{\tilde{s}}(a) f_{\tilde{x}|\tilde{s}}(x|a) + p_{\tilde{s}}(b) f_{\tilde{x}|\tilde{s}}(x|b) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{28}
$$

   (b) By the chain rule for discrete and continuous random variables

$$
\begin{aligned}
p_{\tilde{s}\,|\,\tilde{x}}(b \,|\, 0.25) &= \frac{p_{\tilde{s}}\left(b\right) f_{\tilde{x}\,|\,\tilde{s}}(0.25 \,|\, b)}{f_{\tilde{x}}\left(0.25\right)} && (29)\\
&= 0.25. && (30)
\end{aligned}
$$

   (c) Integrating the conditional pdfs we obtain

$$
F_{\tilde{x}\,|\,\tilde{s}}\left(x \,|\, a\right) = \begin{cases} 0 & \text{if } x < 0, \\ 2x - x^2 & \text{if } 0 \leq x \leq 1, \\ 1 & \text{if } x > 1. \end{cases} \tag{31}
$$

$$
F_{\tilde{x}\,|\,\tilde{s}}\left(x \,|\, b\right) = \begin{cases} 0 & \text{if } x < 0, \\ x^2 & \text{if } 0 \leq x \leq 1, \\ 1 & \text{if } x > 1, \end{cases} \tag{32}
$$

(d) First we sample from $\tilde{s}$ by assigning $\tilde{s} = a$ if the uniform sample is below 0.5, and $\tilde{s} = b$ otherwise. The sample consequently equals b. Then we sample from the conditional distribution of $\tilde{x}$ given $\tilde{s} = b$ using the inverse transform method. We have $F_{\tilde{x}\,|\,b}(0.8\,|\,b) = 0.64$ (computed by solving for $x^2 = 0.64$), so the sample of $\tilde{x}$ equals 0.8.

(e)

$$\begin{aligned}
\mathrm{E}\left((\tilde{x}_2 - \tilde{x}_1)^2\right) &= \mathrm{E}\left(\tilde{x}_1^2 + \tilde{x}_2^2 - 2\tilde{x}_1\tilde{x}_2\right) && (33)\\
&= \mathrm{E}\left(\tilde{x}_1^2\right) + \mathrm{E}\left(\tilde{x}_2^2\right) - 2\mathrm{E}\left(\tilde{x}_1\tilde{x}_2\right) \quad \text{by linearity} && (34)\\
&= \mathrm{E}\left(\tilde{x}_1^2\right) + \mathrm{E}\left(\tilde{x}_2^2\right) - 2\mathrm{E}\left(\tilde{x}_1\right)\mathrm{E}\left(\tilde{x}_2\right) \quad \text{by independence} && (35)\\
&= \mathrm{E}\left(\tilde{x}_1^2\right) + \mathrm{E}\left(\tilde{x}_2^2\right) - \mathrm{E}^2\left(\tilde{x}_1\right) - \mathrm{E}^2\left(\tilde{x}_2\right) \quad \text{same distribution} && (36)\\
&= 2\sigma_{\tilde{x}}^2. && (37)
\end{aligned}$$

5. (Interview)

(a) We have

$$\begin{aligned}
p_{\tilde{i}_1,\tilde{i}_2}(1,1) &= \mathrm{P}(\tilde{e}_1\tilde{q} = 1, \tilde{e}_2\tilde{q} = 1) && (38)\\
&= \mathrm{P}(\tilde{e}_1 = 1, \tilde{e}_2 = 1, \tilde{q} = 1) + \mathrm{P}(\tilde{e}_1 = -1, \tilde{e}_2 = -1, \tilde{q} = -1) && (39)\\
&= \mathrm{P}(\tilde{e}_1 = 1)\mathrm{P}(\tilde{e}_2 = 1)\mathrm{P}(\tilde{q} = 1) + \mathrm{P}(\tilde{e}_1 = -1)\mathrm{P}(\tilde{e}_2 = -1)\mathrm{P}(\tilde{q} = -1) && (40)\\
&= 0.19. && (41)
\end{aligned}$$

(b) We have

$$\begin{aligned}
p_{\tilde{i}_1}(1) &= \mathrm{P}(\tilde{e}_1\tilde{q} = 1) && (42)\\
&= \mathrm{P}(\tilde{e}_1 = 1, \tilde{q} = 1) + \mathrm{P}(\tilde{e}_1 = -1, \tilde{q} = -1) && (43)\\
&= \mathrm{P}(\tilde{e}_1 = 1)\mathrm{P}(\tilde{q} = 1) + \mathrm{P}(\tilde{e}_1 = -1)\mathrm{P}(\tilde{q} = -1) && (44)\\
&= 0.35. && (45)\\
p_{\tilde{i}_2}(1) &= 0.35. && (46)
\end{aligned}$$

Since $0.19 \neq 0.35^2$, the random variables are not independent.

(c) For any x_1, x_2 and q in $\{-1, 1\}$, we have

$$\begin{aligned}
p_{\tilde{i}_1,\tilde{i}_2\,|\,\tilde{q}}(x_1, x_2\,|\,q) &= \mathrm{P}(\tilde{e}_1\tilde{q} = x_1, \tilde{e}_2\tilde{q} = x_2\,|\,\tilde{q} = q) && (47)\\
&= \mathrm{P}(\tilde{e}_1 = x_1/q, \tilde{e}_2 = x_2/q\,|\,\tilde{q} = q) && (48)\\
&= \mathrm{P}(\tilde{e}_1 = x_1/q)\mathrm{P}(\tilde{e}_2 = x_2/q) \quad \text{by independence} && (49)\\
&= \mathrm{P}(\tilde{e}_1 = x_1/q\,|\,\tilde{q} = q)\mathrm{P}(\tilde{e}_2 = x_2/q\,|\,\tilde{q} = q) \quad \text{by independence} && (50)\\
&= \mathrm{P}(\tilde{e}_1\tilde{q} = x_1\,|\,\tilde{q} = q)\mathrm{P}(\tilde{e}_2\tilde{q} = x_2\,|\,\tilde{q} = q) && (51)\\
&= p_{\tilde{i}_1\,|\,\tilde{q}}(x_1\,|\,q)p_{\tilde{i}_2\,|\,\tilde{q}}(x_2\,|\,q), && (52)
\end{aligned}$$

so $\tilde{i}_1$ and $\tilde{i}_2$ are conditionally independent given $\tilde{q}$.

6. (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) After two stops we have

$$q_2 = T^2 q_0 = T \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{8} \\ \frac{1}{2} \\ \frac{1}{8} \end{bmatrix}. \tag{53}$$

   The probability is 1/8.

   (c) The transition matrix is

$$T := \begin{bmatrix} 1 - p_{\text{in}} & (1 - p_{\text{in}})\, p_{\text{out}} & 0 \\ p_{\text{in}} & (1 - p_{\text{in}})(1 - p_{\text{out}}) + p_{\text{in}} p_{\text{out}} & (1 - p_{\text{in}})\, p_{\text{out}} \\ 0 & (1 - p_{\text{out}})\, p_{\text{in}} & 1 - p_{\text{out}} + p_{\text{in}} p_{\text{out}} \end{bmatrix}. \tag{54}$$

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

$$T_{\widetilde{X}} := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{55}$$

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

7. (Potatoes)

   (a) If we fit a kernel density estimate with a narrow width the estimated pdf would be zero for most values, which would make it difficult to use the model in order to do inference on new data. However, if we have a lot of data, applying this nonparametric method could allow to learn irregular structure that may not be well described by parametric models.

   (b) We use the empirical pmfs

$$p_{\tilde{w}}(1) = \frac{1}{3}, \qquad p_{\tilde{w}}(0) = \frac{2}{3}, \tag{56}$$

$$p_{\tilde{b}}(1) = \frac{2}{3}, \qquad p_{\tilde{b}}(0) = \frac{1}{3}. \tag{57}$$

   (c) The random variables are independent, the joint pmf factorizes into the product of the marginal pmfs

$$p_{\tilde{b},\tilde{w}}(0,0) = \frac{2}{9} = p_{\tilde{w}}(0)\, p_{\tilde{b}}(0), \tag{58}$$

$$p_{\tilde{b},\tilde{w}}(0,1) = \frac{1}{9} = p_{\tilde{w}}(1)\, p_{\tilde{b}}(0), \tag{59}$$

$$p_{\tilde{b},\tilde{w}}(1,0) = \frac{4}{9} = p_{\tilde{w}}(0)\, p_{\tilde{b}}(1), \tag{60}$$

$$p_{\tilde{b},\tilde{w}}(1,1) = \frac{2}{9} = p_{\tilde{w}}(1)\, p_{\tilde{b}}(1). \tag{61}$$

(d) The best MSE estimate is $\mathrm{E}(\tilde{x} \,|\, \tilde{b} = 0, \tilde{w} = 1) = 50$ tons (from the graph).

(e) We compare

$$p_{\tilde{b}|\tilde{x},\tilde{w}}(1|40,1) = \frac{p_{\tilde{b}}(1)\, p_{\tilde{w}}(1)\, f_{\tilde{x}|\tilde{b},\tilde{w}}(40|1,1)}{p_{\tilde{w}}(1)\, f_{\tilde{x}|\tilde{w}}(40)} \tag{62}$$

$$= \frac{\frac{2}{3} \cdot \frac{1}{3} \cdot 0.02}{p_{\tilde{w}}(1)\, f_{\tilde{x}|\tilde{w}}(40)} \tag{63}$$

with

$$p_{\tilde{b}|\tilde{x},\tilde{w}}(0|40,1) = \frac{p_{\tilde{b}}(0)\, p_{\tilde{w}}(1)\, f_{\tilde{x}|\tilde{b},\tilde{w}}(40|0,1)}{p_{\tilde{w}}(1)\, f_{\tilde{x}|\tilde{w}}(40)} \tag{64}$$

$$= \frac{\frac{1}{3} \cdot \frac{1}{3} \cdot 0.01}{p_{\tilde{w}}(1)\, f_{\tilde{x}|\tilde{w}}(40)}. \tag{65}$$

The estimate is $\tilde{b} = 1$, i.e. that the beetle was present that year.

(f) $\tilde{b}$ and $\tilde{w}$ are not independent given $\tilde{x}$. If we know the production, then knowing the weather provides information about the presence of the beetle. For example, if the production was high, then knowing that $\tilde{b} = 0$ makes it more likely that the weather was good.

8. (Chad)

(a) The kernel density estimate is of the form

$$f_{\tilde{t}|\tilde{c}}(t|0) = \frac{1}{n_0} \sum_{i=1}^{n_0} \frac{1}{2}\, \Pi\left(\frac{t - d_{0,i}}{2}\right), \tag{66}$$

$$f_{\tilde{t}|\tilde{c}}(t|1) = \frac{1}{n_1} \sum_{i=1}^{n_1} \frac{1}{2}\, \Pi\left(\frac{t - d_{1,i}}{2}\right), \tag{67}$$

$$\tag{68}$$

where Π is a rectangular kernel with unit width, $d_{1,0}, \ldots, d_{n_1,0}$ the temperatures when Chad is not there and $d_{1,1}, \ldots, d_{n_1,1}$ the temperatures when he is there. The estimate is shown in Figure 1.

(b) We have

$$f_{\tilde{t}|\tilde{c}}(68|0) = 0.2 > 0 = f_{\tilde{t}|\tilde{c}}(68|1), \tag{69}$$

so the ML estimate is that Chad is not at the office.

(c) The empirical pmf is

$$p_{\tilde{c}}(0) = \frac{5}{15} = \frac{1}{3}, \tag{70}$$

$$p_{\tilde{c}}(1) = \frac{10}{15} = \frac{2}{3}. \tag{71}$$

(d) Applying Bayes' rule,

$$\begin{aligned} p_{\tilde{c}|\tilde{t}}(0|64) &= \frac{p_{\tilde{c}}(0)\, f_{\tilde{t}|\tilde{c}}(64|0)}{p_{\tilde{c}}(0)\, f_{\tilde{t}|\tilde{c}}(64|0) + p_{\tilde{c}}(1)\, f_{\tilde{t}|\tilde{c}}(64|1)} && (72)\\ &= \frac{\frac{1}{3}0.2}{\frac{1}{3}0.2 + \frac{2}{3}0.1} && (73)\\ &= \frac{1}{2}, && (74)\\ p_{\tilde{c}|\tilde{t}}(1|64) &= 1 - p_{\tilde{c}|\tilde{t}}(0|64) && (75)\\ &= \frac{1}{2}. && (76) \end{aligned}$$

According to the posterior pmf there is a 50 % chance that Chad is there.

(e) Both $f_{\tilde{t}|\tilde{c}}(57|0)$ and $f_{\tilde{t}|\tilde{c}}(68|1)$ are zero so both the ML estiamate and Bayesian posterior are inconclusive. If we use a parametric distribution such as a Gaussian to fit the data, then $f_{\tilde{t}|\tilde{c}}(57|0)$ and $f_{\tilde{t}|\tilde{c}}(68|1)$ would not be set to zero as long as the distribution has nonzero values on all of the real line (as is the case for a Gaussian pdf). This would allow us to apply MAP or ML estimation. A nonparametric solution would be to use a kernel with a larger width.

9. (Rufus)

(a) By assumption, the joint probability density is equal to a constant c. It must integrate to one, so

$$\int_{\text{garden}} c \,\mathrm{d}x\,\mathrm{d}y = c\,\text{Area}\,(\text{garden}) = 1. \tag{77}$$

The area of the garden is equal to $100^2 - 40^2 = 10000 - 1600 = 8400$, which implies

$$f_{\tilde{x},\tilde{y}}(x,y) = \begin{cases} \frac{1}{8400} & \text{if } \{(x,y) \mid -50 \leq x \leq 50, -50 \leq y \leq 50\} / \{(x,y) \mid -20 \leq x \leq 20, -20 \leq \\ 0 & \text{otherwise.} \end{cases} \tag{78}$$

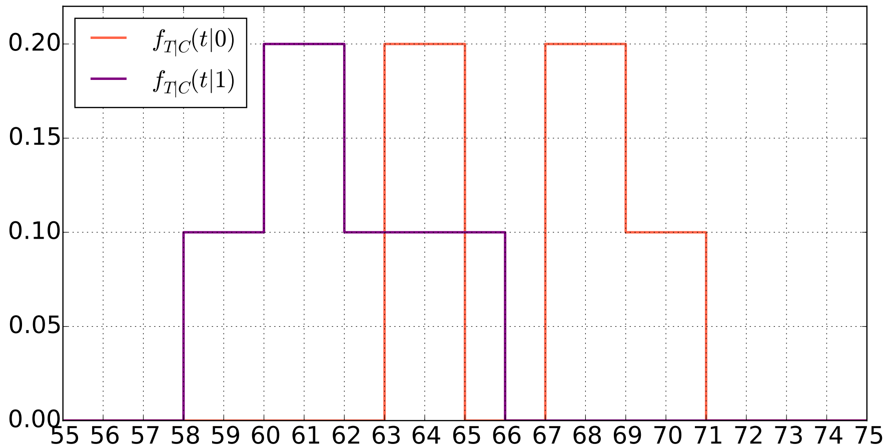


Figure 1: Kernel density estimate for Problem (Chad).

(b) Let $c := 1/8400$, by applying a change of variables we obtain

$$\mathrm{E}\left(\tilde{x}\right) = \int_{y=-50}^{50} \left(\int_{x=-50}^{-20} cx \,\mathrm{d}x + \int_{x=20}^{50} cx \,\mathrm{d}x \right) \mathrm{d}y \tag{79}$$

$$+ \int_{x=-20}^{20} cx \,\mathrm{d}x \left(\int_{y=-50}^{-20} \mathrm{d}y + \int_{y=20}^{50} cx \,\mathrm{d}y \right). \tag{80}$$

Since by a change of variables $t = -x$ $\int_{x=-50}^{-20} cx \,\mathrm{d}x = -\int_{x=20}^{50} cx \,\mathrm{d}x$, and $\int_{x=-20}^{20} cx \,\mathrm{d}x = 0$, we have $\mathrm{E}\left(\tilde{x}\right) = 0$.

$$\mathrm{Var}\left(\tilde{x}\right) = \mathrm{E}\left(\tilde{x}^2\right) \tag{81}$$

$$= \int_{y=-50}^{50} \left(\int_{x=-50}^{-20} cx^2 \,\mathrm{d}x + \int_{x=20}^{50} cx^2 \,\mathrm{d}x \right) \mathrm{d}y \tag{82}$$

$$+ \int_{x=-20}^{20} cx^2 \,\mathrm{d}x \left(\int_{y=-50}^{-20} \mathrm{d}y + \int_{y=20}^{50} \mathrm{d}y \right) \tag{83}$$

$$= 100c(\frac{(-20)^3 - (-50)^3}{3} + \frac{50^3 - 20^3}{3}) + c\frac{0^3 - (-20)^3}{3}(30 + 30) \tag{84}$$

$$= 966.6 \tag{85}$$

The standard deviation is 31.09.

(c) For y in $[-50, -20]$ or $[20, 50]$

$$f_{\tilde{y}}\left(y\right) = \int_{x=-\infty}^{\infty} f_{\tilde{x},\tilde{y}}\left(x, y\right) \mathrm{d}x \tag{86}$$

$$= \int_{x=-50}^{50} \frac{1}{8400} \left(x, y\right) \mathrm{d}x \tag{87}$$

$$= \frac{50 + 50}{8400} = \frac{1}{84}. \tag{88}$$

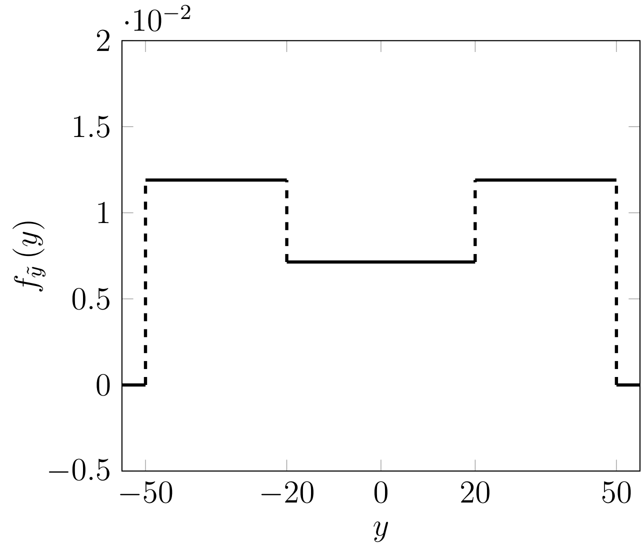


Figure 2: Pdf of the vertical position $\tilde{y}$ of Rufus.

For y in $[-20, 20]$

$$f_{\tilde{y}}(y) = \int_{x=-\infty}^{\infty} f_{\tilde{x},\tilde{y}}(x, y)\, \mathrm{d}x \tag{89}$$

$$= \int_{x=-50}^{-20} \frac{1}{8400}(x, y)\, \mathrm{d}x + \int_{x=20}^{50} \frac{1}{8400}(x, y)\, \mathrm{d}x \tag{90}$$

$$= 2 \cdot \frac{50 - 20}{8400} = \frac{1}{140}. \tag{91}$$

For other values of y $f_{\tilde{y}}(y) = 0$. The pdf is shown in Figure 2.

(d) In order to compute the conditional pdf of $\tilde{x}$ given $\tilde{y}$ we need to compute the marginal pdf of $\tilde{y}$ by marginalizing over $\tilde{x}$. The conditional density is only defined for values of y such that $f_{\tilde{y}}(y) \neq 0$. For y in $[-50, -20]$ or $[20, 50]$

$$f_{\tilde{x}|\tilde{y}}(x|y) = \frac{f_{\tilde{x},\tilde{y}}(x, y)}{f_{\tilde{y}}(y)} = \begin{cases} \frac{1}{100} & \text{if } -50 \leq x \leq 50, \\ 0 & \text{otherwise.} \end{cases} \tag{92}$$

For y in $[-20, 20]$

$$f_{\tilde{x}|\tilde{y}}(x|y) = \begin{cases} \frac{1}{60} & \text{if } -50 \leq x \leq -20 \text{ or } 20 \leq x \leq 50, \\ 0 & \text{otherwise.} \end{cases} \tag{93}$$

The conditional pdfs are plotted in Figure 3.

(e) $\tilde{x}$ and $\tilde{y}$ are not independent. By (91) $f_{\tilde{y}}(0) = \frac{60}{8400} \neq 0$ and by symmetry $f_{\tilde{x}}(0) = \frac{60}{8400} \neq 0$. However $(0, 0)$ is not in the garden, so

$$f_{\tilde{x},\tilde{y}}(0, 0) = 0 \neq f_{\tilde{x}}(0)\, f_{\tilde{y}}(0). \tag{94}$$

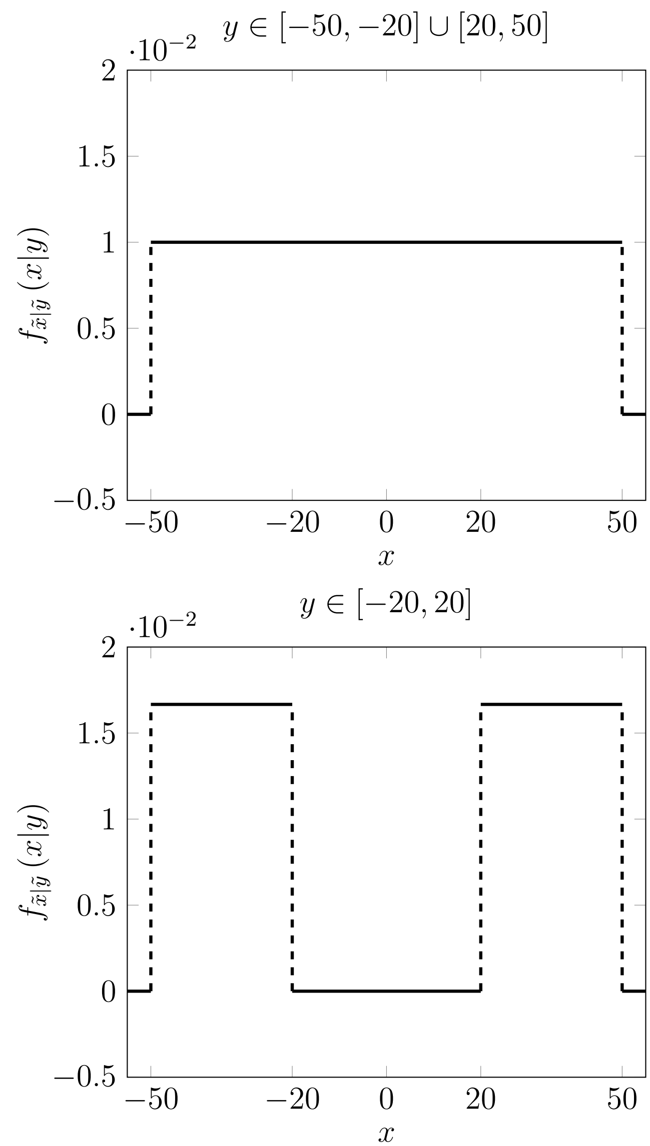


Figure 3: Conditional pdf of the position $\tilde{x}$ of Rufus given his $\tilde{y}$ position.

(f) The mean of $\tilde{x}$ is zero, so to compute the covariance, we only need to compute $\mathrm{E}(\tilde{x}\tilde{y})$. By iterated expectation, we can take the mean of $\mathrm{E}(\tilde{x}\tilde{y} \mid \tilde{y})$, which is computed by first deriving the function $h(y) := \mathrm{E}(\tilde{x}\tilde{y} \mid \tilde{y} = y)$ for all possible values of y and plugging in $\tilde{y}$. For y in $[-50, -20]$ or $[20, 50]$, we have

$$\begin{aligned} \mathrm{E}(\tilde{x}\tilde{y} \mid \tilde{y} = y) &= \int_{x=-50}^{50} \frac{yx}{100} \, \mathrm{d}x && (95) \\ &= \frac{yx}{100} \int_{x=-50}^{50} x \, \mathrm{d}x && (96) \\ &= 0. && (97) \end{aligned}$$

For y in $[-20, 20]$, similarly

$$\begin{aligned} \mathrm{E}(\tilde{x}\tilde{y} \mid \tilde{y} = y) &= \frac{1}{60} \left(\int_{x=-50}^{-20} x \, \mathrm{d}x + \int_{x=20}^{50} x \, \mathrm{d}x \right) && (98) \\ &= 0. && (99) \end{aligned}$$

The conditional mean of $\tilde{x}\tilde{y}$ is always zero for values of $\tilde{y}$ that have nonzero probability. By iterated expectation this implies that the covariance is zero, so the variables are uncorrelated.

10. (Frog)

(a) From the assumptions we know that the density is zero outside of the ponds, equal to a constant c_L in the large pond and to another constant c_S in the small pond. The probability of the frog being in the large pond is 1/4 so

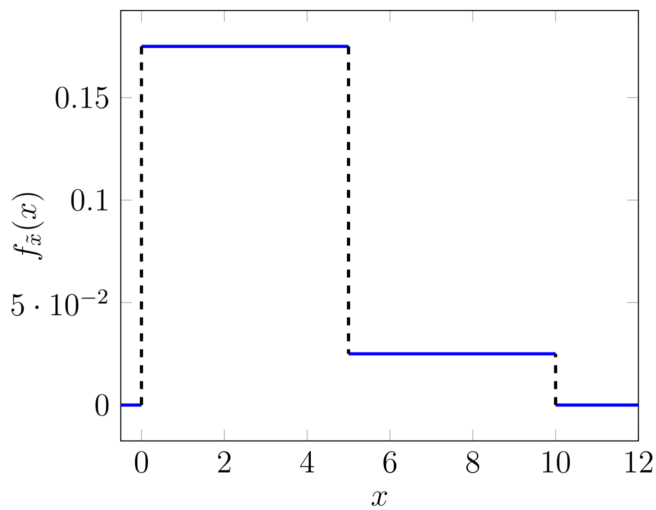
$$\begin{aligned} P\,(\text{large pond}) &= \int_0^{10} \int_0^{10} c_L \, \mathrm{d}x \, \mathrm{d}y && (100) \\ &= 100 c_L = \frac{1}{4}, && (101) \end{aligned}$$

which implies $c_L = \frac{1}{400}$. Similarly

$$\begin{aligned} P\,(\text{small pond}) &= \int_{12}^{17} \int_0^{5} c_S \, \mathrm{d}x \, \mathrm{d}y && (102) \\ &= 25 c_S = \frac{3}{4}, && (103) \end{aligned}$$

which implies $c_S = \frac{3}{100}$. Therefore,

$$f_{\tilde{x},\tilde{y}}(x, y) = \begin{cases} \frac{1}{400} & \text{if } 0 \le x, y \le 10 \\ \frac{3}{100} & \text{if } 0 \le x \le 5, 12 \le y \le 17 \\ 0 & \text{otherwise.} \end{cases}$$



(b) The marginal pdf of the horizontal position of the frog is given by

$$f_{\tilde{x}}(x) = \int_0^{17} f_{\tilde{x},\tilde{y}}(x,y)dy = \begin{cases} \int_0^{10} \frac{1}{400}dy + \int_{12}^{17} \frac{3}{100}dy = \frac{7}{40} & \text{if } 0 \leq x \leq 5 \\ \int_0^{10} \frac{1}{400}dy = \frac{1}{40} & \text{if } 5 \leq x \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

(c)

$$f_{\tilde{y}|\tilde{x}=3}(y) = \frac{f(3,y)}{f_{\tilde{x}}(3)} = \frac{40}{7} \begin{cases} \frac{1}{400} & \text{if } 0 \leq y \leq 10 \\ \frac{3}{100} & \text{if } 12 \leq y \leq 17 \\ 0 & \text{otherwise} \end{cases} \tag{104}$$

$$= \begin{cases} \frac{1}{70} & \text{if } 0 \leq y \leq 10 \\ \frac{6}{35} & \text{if } 12 \leq y \leq 17 \\ 0 & \text{otherwise} \end{cases} \tag{105}$$

(d) The vertical position of the frog is not independent from the horizontal position since

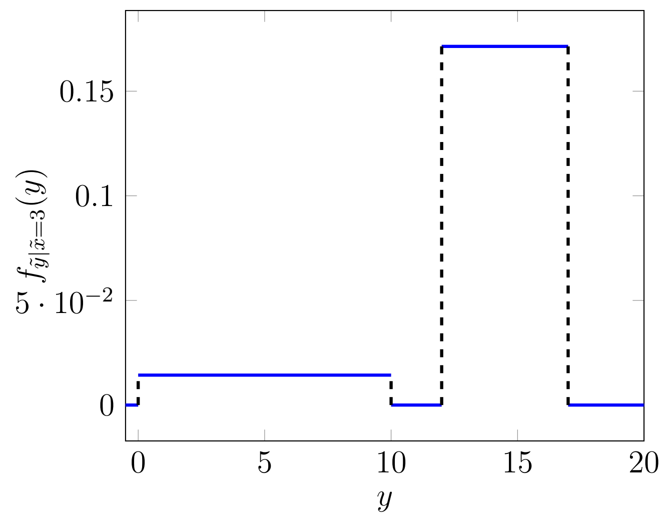
$$f_{\tilde{y}}(13) = \int_0^5 f_{\tilde{x},\tilde{y}}(x,y)\, \mathrm{d}x \tag{106}$$

$$= 5\frac{3}{100} = \frac{3}{20} \neq \frac{6}{35} = f_{\tilde{y}|\tilde{x}=3}(13). \tag{107}$$

(e) Conditioned on the frog being in the small pond the joint pdf is a constant c such that

$$P\,(\text{small pond}) = \int_{12}^{17} \int_0^5 c\, \mathrm{d}x\, \mathrm{d}y \tag{108}$$

$$= 25c = 1, \tag{109}$$



so $c = \frac{1}{25}$. The joint pdf is consequently

$$f_{\tilde{x},\tilde{y}|\text{small pond}}(x, y) = \begin{cases} \frac{1}{25} & \text{if } 0 \leq x \leq 5 \text{ and } 12 \leq y \leq 17, \\ 0 & \text{otherwise.} \end{cases} \tag{110}$$

We marginalize to find the pdfs of $\tilde{x}$ and $\tilde{y}$.

$$f_{\tilde{x}|\text{small pond}}(x) = \int_{12}^{17} f_{\tilde{x},\tilde{y}|\text{small pond}}(x, y)\, \mathrm{d}y = \begin{cases} \frac{1}{5} & \text{if } 0 \leq x \leq 5, \\ 0 & \text{otherwise.} \end{cases} \tag{111}$$

$$f_{\tilde{y}|\text{small pond}}(y) = \int_{0}^{5} f_{\tilde{x},\tilde{y}|\text{small pond}}(x, y)\, \mathrm{d}x = \begin{cases} \frac{1}{5} & \text{if } 12 \leq y \leq 17, \\ 0 & \text{otherwise.} \end{cases} \tag{112}$$

The two variables are conditionally independent because $f_{\tilde{x}|\text{small pond}}(x)\, f_{\tilde{y}|\text{small pond}}(y) = f_{\tilde{x},\tilde{y}|\text{small pond}}(x, y)$ for any values of x and y.

(f)

$$\begin{aligned} \mathrm{E}(\tilde{y}) &= \int_{x=0}^{10} \int_{y=0}^{10} c_L y \,\mathrm{d}x\, \mathrm{d}y + \int_{x=0}^{5} \int_{y=12}^{17} c_S y \,\mathrm{d}x\, \mathrm{d}y && (113) \\ &= \frac{1}{400}(10 - 0)\frac{10^2}{2} + \frac{3}{100}(5 - 0)\frac{17^2 - 12^2}{2} && (114) \\ &= 12.125. && (115) \end{aligned}$$

Alternatively, we could have used iterated expectation.

11. (Babysitter)

(a)

$$
\begin{aligned}
p_{\tilde{w}_1}(1) &= \sum_{x=0}^{1} \sum_{b=0}^{1} p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(x,b,1) && (116)\\
&= \sum_{x=0}^{1} \sum_{b=0}^{1} p_{\tilde{x}}(x)\, p_{\tilde{b}_1}(b_1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, x,b) && (117)\\
&= 0.1 + 0.9 \cdot (0.6 \cdot 0.1 + 0.4 \cdot 0.8) && (118)\\
&= 0.442. && (119)
\end{aligned}
$$

(b)

$$
\begin{aligned}
p_{\tilde{x} \,|\, \tilde{w}_1}(1 \,|\, 1) &= \frac{p_{\tilde{x},\tilde{w}_1}(1,1)}{p_{\tilde{w}_1}(1)} && (120)\\
&= \frac{p_{\tilde{x}}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x}}(1 \,|\, 1)}{0.442} && (121)\\
&= \frac{0.1}{0.442} && (122)\\
&= 0.226. && (123)
\end{aligned}
$$

(c) We have:

$$
\begin{aligned}
p_{\tilde{x} \,|\, \tilde{w}_1,\tilde{b}_1}(1 \,|\, 1,1) &= \frac{p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(1,1,1)}{p_{\tilde{w}_1,\tilde{b}_1}(1,1)} && (124)\\
&= \frac{p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(1,1,1)}{\sum_{x=0}^{1} p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(x,1,1)} && (125)\\
&= \frac{p_{\tilde{x}}(1)\, p_{\tilde{b}_1}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 1,1)}{p_{\tilde{b}_1}(1) \left(p_{\tilde{x}}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 1,1) + p_{\tilde{x}}(0)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 0,1) \right)} && (126)\\
&= \frac{p_{\tilde{x}}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 1,1)}{p_{\tilde{x}}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 1,1) + p_{\tilde{x}}(0)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 0,1)} && (127)\\
&= \frac{0.1}{0.1 + 0.9 \cdot 0.1} && (128)\\
&= 0.526 && (129)
\end{aligned}
$$

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.

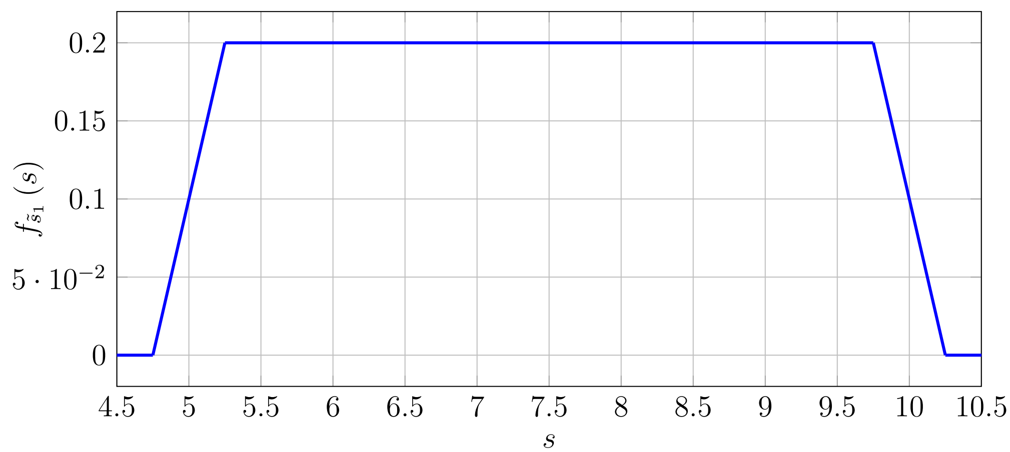
12. (Sonar)

(a) Note that for fixed x $f_{\tilde{s}_1 \,|\, \tilde{x}}(s \,|\, x) = 0$ if $x < s - 0.25$ or $x > s + 0.25$. This implies

$$f_{\tilde{s}_1}(s) = \int_5^{10} f_{\tilde{x}}(x) f_{\tilde{s}_1 \,|\, \tilde{x}}(s \,|\, x) \, \mathrm{d}x \tag{130}$$

$$= \int_{\max\{5, s-0.25\}}^{\min\{10, s+0.25\}} \frac{2}{5} \, \mathrm{d}x \tag{131}$$

$$= \begin{cases} 0 & \text{if } s < 4.75 \\ \frac{2(s-4.75)}{5} & \text{if } 4.75 \le s \le 5.25 \\ \frac{1}{5} & \text{if } 5.25 \le s \le 9.75 \\ \frac{2(10.25-s)}{5} & \text{if } 9.75 \le s \le 10.25 \\ 0 & \text{if } x > 10.25. \end{cases} \tag{132}$$



(b) For fixed s_1 $f_{\tilde{s}_1 \,|\, \tilde{x}}(s_1 \,|\, x) = 0$ if $x < s_1 - 0.25$ or $x > s_1 + 0.25$. For fixed s_2 $f_{\tilde{s}_2 \,|\, \tilde{x}}(s_2 \,|\, x) = 0$ if $x < s_2 - 0.25$ or $x > s_2 + 0.25$, so

$$f_{\tilde{s}_1, \tilde{s}_2}(7, 7.1) = \int_5^{10} f_{\tilde{x}}(x) f_{\tilde{s}_1 \,|\, \tilde{x}}(7 \,|\, x) f_{\tilde{s}_2 \,|\, \tilde{x}}(7.1 \,|\, x) \, \mathrm{d}x \tag{133}$$

$$= \int_{6.85}^{7.25} \frac{4}{5} \, \mathrm{d}x \tag{134}$$

$$= 0.32. \tag{135}$$

$$f_{\tilde{x} \,|\, \tilde{s}_1, \tilde{s}_2}(x \,|\, 7, 7.1) = \frac{f_{\tilde{x}, \tilde{s}_1, \tilde{s}_2}(x, 7, 7.1)}{f_{\tilde{s}_1, \tilde{s}_2}(7, 7.1)} \tag{136}$$

$$= \begin{cases} \frac{4/5}{0.32} = 2.5 & \text{if } 6.85 \le x \le 7.25, \\ 0 & \text{otherwise.} \end{cases} \tag{137}$$

(c) For fixed s_1 $f_{\tilde{s}_1 \,|\, \tilde{x}}(s_1 \,|\, x) = 0$ if $x < s_1 - 0.25$ or $x > s_1 + 0.25$. For fixed s_2

$f_{\tilde{s}_2 \,|\, \tilde{x}}(s_2 \,|\, x) = 0$ if $x < s_2 - 0.25$ or $x > s_2 + 0.25$, so

$$f_{\tilde{s}_1,\tilde{s}_2}(s_1, s_2) = \int_5^{10} f_{\tilde{x}}(x)\, f_{\tilde{s}_1 \,|\, \tilde{x}}(s_1 \,|\, x)\, f_{\tilde{s}_2 \,|\, \tilde{x}}(s_2 \,|\, x)\, \mathrm{d}x \tag{138}$$

$$= \int_{\max\{5, s_1 - 0.25, s_2 - 0.25\}}^{\min\{10, s_1 + 0.25, s_2 + 0.25\}} \frac{4}{5}\, \mathrm{d}x \tag{139}$$

$$= 0.8 \left(\min\{10, s_1 + 0.25, s_2 + 0.25\} - \max\{5, s_1 - 0.25, s_2 - 0.25\}\right)$$

if $4.75 \leq s_1, s_2 \leq 10.25$ and $|s_2 - s_1| \leq 0.5$, and 0 otherwise.

The two measurements are not independent. For example $f_{\tilde{s}_1}(5) \neq 0$ and $f_{\tilde{s}_2}(10) \neq 0$ but $f_{\tilde{s}_1,\tilde{s}_2}(5, 10) = 0$ so the joint pdf is not the product of the marginals. This makes sense since $\tilde{s}_1$ provides information about the location of $\tilde{x}$, which in turn provides information about the location of $\tilde{s}_2$.

(d) We apply iterated expectation, we have

$$\mathrm{E}\left(\left(\frac{\tilde{s}_1 + \tilde{s}_2}{2} - \tilde{x}\right)^2 \,|\, \tilde{x} = x\right) = \mathrm{E}\left(\left(\frac{\tilde{u}_1 + \tilde{u}_2}{2} + x - x\right)^2 \,|\, \tilde{x} = x\right) \tag{140}$$

$$= \mathrm{E}\left(\left(\frac{\tilde{u}_1 + \tilde{u}_2}{2} + x - x\right)^2 \,|\, \tilde{x} = x\right) \tag{141}$$

$$= \mathrm{E}\left(\left(\frac{\tilde{u}_1 + \tilde{u}_2}{2}\right)^2 \,|\, \tilde{x} = x\right) \tag{142}$$

$$= \mathrm{E}\left(\left(\frac{\tilde{u}_1 + \tilde{u}_2}{2}\right)^2\right) \qquad \text{by independence of } \tilde{u}_1,\ \tilde{u}_2 \text{ and } \tilde{x}$$

$$= \frac{1}{4}\left(\mathrm{E}\left(\tilde{u}_1^2\right) + \mathrm{E}\left(\tilde{u}_2^2\right) + 2\mathrm{E}(\tilde{u}_1)\mathrm{E}(\tilde{u}_2)\right) \tag{143}$$

$$= \frac{1}{4}\left(\mathrm{E}\left(\tilde{u}_1^2\right) + \mathrm{E}\left(\tilde{u}_2^2\right)\right) \tag{144}$$

$$= \frac{2}{4}\int_{-0.25}^{0.25} 2u^2\, \mathrm{d}u \tag{145}$$

$$= 2\left(5.21 \cdot 10^{-3}\right). \tag{146}$$

Since this is a constant, taking the integral with respect to $\tilde{x}$ gives the constant itself. The MSE equals $2\,(5.21 \cdot 10^{-3})$.