Homework 5

Solutions

1. (Babysitter)

(a)

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

(b)

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

(c) We have:

$$p_{\tilde{x} \,|\, \tilde{w}_1,\tilde{b}_1}(1 \,|\, 1, 1) = \frac{p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(1, 1, 1)}{p_{\tilde{w}_1,\tilde{b}_1}(1, 1)} \tag{9}$$
$$= \frac{p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(1, 1, 1)}{\sum_{x=0}^{1} p_{\tilde{x},\tilde{b}_1,\tilde{w}_1}(x, 1, 1)} \tag{10}$$
$$= \frac{p_{\tilde{x}}(1)\, p_{\tilde{b}_1}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 1, 1)}{p_{\tilde{b}_1}(1) \left(p_{\tilde{x}}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 1, 1) + p_{\tilde{x}}(0)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 0, 1) \right)} \tag{11}$$
$$= \frac{p_{\tilde{x}}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 1, 1)}{p_{\tilde{x}}(1)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 1, 1) + p_{\tilde{x}}(0)\, p_{\tilde{w}_1 \,|\, \tilde{x},\tilde{b}_1}(1 \,|\, 0, 1)} \tag{12}$$
$$= \frac{0.1}{0.1 + 0.9 \cdot 0.1} \tag{13}$$
$$= 0.526 \tag{14}$$

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. (Earthquake)

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

$$
\begin{aligned}
p_{\tilde{s}}(1) &= \sum_{e=0}^{1}\sum_{v=0}^{2} p_{\tilde{e},\tilde{v}}(e,v)p_{\tilde{s}\,|\,\tilde{e},\tilde{v}}(1\,|\,e,v) && (15)\\
&= \sum_{e=0}^{1}\sum_{v=0}^{2} p_{\tilde{e},\tilde{v}}(e,v)p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,v) && (16)\\
&= p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,1)(p_{\tilde{e},\tilde{v}}(0,1)+p_{\tilde{e},\tilde{v}}(1,1)) + p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,2)(p_{\tilde{e},\tilde{v}}(0,2)+p_{\tilde{e},\tilde{v}}(1,2)) && (17)\\
&= 0.5(0.05+0.05)+0.1 && (18)\\
&= 0.15 && (19)
\end{aligned}
$$

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

   (b)

$$
\begin{aligned}
p_{\tilde{e}\,|\,\tilde{s}}(1\,|\,1) &= \frac{p_{\tilde{e},\tilde{s}}(1,1)}{p_{\tilde{s}}(1)} && (20)\\
&= \frac{\sum_{v=0}^{2} p_{\tilde{e},\tilde{v},\tilde{s}}(1,v,1)}{p_{\tilde{s}}(1)} && (21)\\
&= \frac{\sum_{v=0}^{2} p_{\tilde{e},\tilde{v},\tilde{s}}(1,v)p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,v)}{p_{\tilde{s}}(1)} && (22)\\
&= \frac{p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,1)p_{\tilde{e},\tilde{v}}(1,1) + p_{\tilde{s}\,|\,\tilde{v}}(1\,|\,2)p_{\tilde{e},\tilde{v}}(1,2)}{p_{\tilde{s}}(1)} && (23)\\
&= \frac{0.5\cdot 0.05+0.1}{0.15} && (24)\\
&= 0.833. && (25)
\end{aligned}
$$

   (c) We have

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

   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.

3. (Surgery)

   (a) Let us define the potential outcomes $\widetilde{\mathrm{po}}_A$ and $\widetilde{\mathrm{po}}_B$. $\widetilde{\mathrm{po}}_A = 1$ and $\widetilde{\mathrm{po}}_B = 1$ indicate recovery after procedure A and B respectively. $\widetilde{\mathrm{po}}_A = 0$ and $\widetilde{\mathrm{po}}_B = 0$ indicate non-recovery. The random variable $\tilde{y}$ represents the observed outcome and $\tilde{t}$ the treatment, so that

$$\tilde{y} := \begin{cases} \widetilde{\mathrm{po}}_A & \text{if } \tilde{t} = A, \\ \widetilde{\mathrm{po}}_B & \text{if } \tilde{t} = B. \end{cases} \tag{29}$$

We also define a random variable $\tilde{d}$ to indicate whether each case is mild ($\tilde{d} = m$) or serious ($\tilde{d} = s$). We denote the probability that a case is mild if the treatment is A or B by α_A and α_B respectively. The observed probability of recovery for patients undergoing procedure A is

$$\begin{aligned}
\mathrm{P}\left(\tilde{y} = 1 \,|\, \tilde{t} = A\right) &= \mathrm{P}\left(\widetilde{\mathrm{po}}_A = 1 \,|\, \tilde{t} = A\right) &(30)\\
&= \mathrm{P}\left(\widetilde{\mathrm{po}}_A = 1, \tilde{d} = m \,|\, \tilde{t} = A\right) + \mathrm{P}\left(\widetilde{\mathrm{po}}_A = 1, \tilde{d} = s \,|\, \tilde{t} = A\right) &(31)\\
&= \mathrm{P}\left(\tilde{d} = m \,|\, \tilde{t} = A\right) \mathrm{P}\left(\widetilde{\mathrm{po}}_A = 1 \,|\, \tilde{t} = A, \tilde{d} = m\right) &(32)\\
&\quad + \mathrm{P}\left(\tilde{d} = s \,|\, \tilde{t} = A\right) \mathrm{P}\left(\widetilde{\mathrm{po}}_A = 1 \,|\, \tilde{t} = A, \tilde{d} = s\right) &(33)\\
&= 0.9\alpha_A + 0.5(1 - \alpha_A). &(34)
\end{aligned}$$

Setting this equal to 0.58 yields $\alpha_A = 0.2$.

Similarly,

$$\begin{aligned}
\mathrm{P}\left(\tilde{y} = 1 \,|\, \tilde{t} = B\right) &= \mathrm{P}\left(\widetilde{\mathrm{po}}_B = 1 \,|\, \tilde{t} = B\right) &(35)\\
&= \mathrm{P}\left(\widetilde{\mathrm{po}}_B = 1, \tilde{d} = m \,|\, \tilde{t} = B\right) + \mathrm{P}\left(\widetilde{\mathrm{po}}_B = 1, \tilde{d} = s \,|\, \tilde{t} = B\right) &(36)\\
&= \mathrm{P}\left(\tilde{d} = m \,|\, \tilde{t} = B\right) \mathrm{P}\left(\widetilde{\mathrm{po}}_B = 1 \,|\, \tilde{t} = B, \tilde{d} = m\right) &(37)\\
&\quad + \mathrm{P}\left(\tilde{d} = s \,|\, \tilde{t} = B\right) \mathrm{P}\left(\widetilde{\mathrm{po}}_B = 1 \,|\, \tilde{t} = B, \tilde{d} = s\right) &(38)\\
&= 0.8\alpha_B + 0.2(1 - \alpha_B). &(39)
\end{aligned}$$

Setting this equal to 0.68 yields $\alpha_A = 0.8$.

What is happening is that the fraction of patients with mild cases is much lower for procedure A (20%) than for procedure B (80%). Since those patients are more likely to recover, irrespective of the procedure, this inflates the recovery rate for procedure B.

(b) We can correct for this by taking into account whether each case is mild or serious. Under the assumption that the treatment $\tilde{t}$ and the potential outcomes $\widetilde{\mathrm{po}}_A$ and $\widetilde{\mathrm{po}}_B$ are conditionally independent given the degree of severity $\tilde{d}$, then

$$\begin{aligned}
\mathrm{P}\left(\tilde{y} \,|\, \tilde{t} = A, \tilde{d} = m\right) &= \mathrm{P}\left(\widetilde{\mathrm{po}}_A \,|\, \tilde{t} = A, \tilde{d} = m\right) &(40)\\
&= \mathrm{P}\left(\widetilde{\mathrm{po}}_A \,|\, \tilde{d} = m\right). &(41)
\end{aligned}$$

By the same argument,

$$\mathrm{P}\left(\tilde{y}=1 \,|\, \tilde{t}=A, \tilde{d}=s\right)=\mathrm{P}\left(\widetilde{\mathrm{po}}_A=1 \,|\, \tilde{d}=s\right), \tag{42}$$

$$\mathrm{P}\left(\tilde{y}=1 \,|\, \tilde{t}=B, \tilde{d}=m\right)=\mathrm{P}\left(\widetilde{\mathrm{po}}_B=1 \,|\, \tilde{d}=m\right), \tag{43}$$

$$\mathrm{P}\left(\tilde{y}=1 \,|\, \tilde{t}=B, \tilde{d}=s\right)=\mathrm{P}\left(\widetilde{\mathrm{po}}_B=1 \,|\, \tilde{d}=s\right). \tag{44}$$

Consequently, we can compute the *true* efficacy as follows,

$$\mathrm{P}\left(\widetilde{\mathrm{po}}_A\right)=\sum_{d\in\{m,s\}} \mathrm{P}\left(\widetilde{\mathrm{po}}_A=1 \,|\, \tilde{d}=d\right)\mathrm{P}\left(\tilde{d}=d\right) \tag{45}$$

$$=\sum_{d\in\{m,s\}} \mathrm{P}\left(\tilde{y}=1 \,|\, \tilde{t}=A, \tilde{d}=d\right)\mathrm{P}\left(\tilde{d}=d\right), \tag{46}$$

$$\mathrm{P}\left(\widetilde{\mathrm{po}}_B\right)=\sum_{d\in\{m,s\}} \mathrm{P}\left(\widetilde{\mathrm{po}}_B=1 \,|\, \tilde{d}=d\right)\mathrm{P}\left(\tilde{d}=d\right) \tag{47}$$

$$=\sum_{d\in\{m,s\}} \mathrm{P}\left(\tilde{y}=1 \,|\, \tilde{t}=B, \tilde{d}=d\right)\mathrm{P}\left(\tilde{d}=d\right). \tag{48}$$

This only works if the conditional independence assumption holds. Intuitively, we are assuming that once we control for the severity, there are no other systematic differences between the patients that undergo each of the procedures.

(c) Randomizing what patient undergoes each procedure would enable us to neutralize all confounding factors, even if we don't know what they are explicitly.

4. (Stock) Denote three Bernoulli random variables indicating whether the price raises (1) or drop (0) for amzn, jpm, spy as $\tilde{a}, \tilde{j}, \tilde{s}$.

   (a) $p_{\tilde{a},\tilde{j},\tilde{s}}(a, j, s)$

```
columns = ['amzn', 'jpm', 'spy']
data = pd.read_csv("pricedelta.csv").drop("Dates", axis=1) > 0
data = data[columns].astype('int')

def compute_joint_pmf(stock_1, stock_2, stock_3, data_matrix):
    submatrix = data_matrix[[stock_1, stock_2, stock_3]]
    submatrix = np.array(submatrix)
    total_count = 0.
    counts = np.zeros((2,2,2))
    for ind in range(submatrix.shape[0]):
        rain_1 = int(submatrix[ind,0] > 0.)
        rain_2 = int(submatrix[ind,1] > 0.)
        rain_3 = int(submatrix[ind,2] > 0.)
        total_count += 1
```

```
            counts[rain_1,rain_2,rain_3] += 1
        joint_pmf = counts/total_count
        return counts,joint_pmf

counts,joint_pmf = compute_joint_pmf('amzn','jpm', 'spy', data)

[[[0.22453704 0.03472222]
  [0.07407407 0.1087963 ]]

 [[0.11574074 0.10416667]
  [0.03703704 0.30092593]]]
```

(b) $p_{\tilde{a},\tilde{j}}(a,j)$, $p_{\tilde{a},\tilde{s}}(a,s)$, $p_{\tilde{s},\tilde{j}}(s,j)$

```
def marginal_2_stocks(joint_pmf):
    marginal_pmf_12 = np.zeros((2,2))
    for ind_3 in range(2):
        marginal_pmf_12 += joint_pmf[:,:,ind_3]
    marginal_pmf_13 = np.zeros((2,2))
    for ind_2 in range(2):
        marginal_pmf_13 += joint_pmf[:,ind_2,:]
    marginal_pmf_23 = np.zeros((2,2))
    for ind_1 in range(2):
        marginal_pmf_23 += joint_pmf[ind_1,:,:]
    return marginal_pmf_12,marginal_pmf_13,marginal_pmf_23

marginal_pmf_12,marginal_pmf_13,marginal_pmf_23 = marginal_2_stocks(joint_pmf)

print(marginal_pmf_12)
print(marginal_pmf_13)
print(marginal_pmf_23)


[[0.25925926 0.18287037]
 [0.21990741 0.33796296]]

[[0.29861111 0.14351852]
 [0.15277778 0.40509259]]

[[0.34027778 0.13888889]
 [0.11111111 0.40972222]]
```

(c) $p_{\tilde{a},\tilde{j}|\tilde{s}}(a,j,s)$, $p_{\tilde{a},\tilde{s}|\tilde{j}}(a,s,j)$, $p_{\tilde{s},\tilde{j}|\tilde{a}}(s,j,a)$

```
def conditional_2_stock_given_1(joint_pmf):
    cond_12_given_3 = np.zeros((2,2,2))
    for ind_3 in range(2):
```

```
        cond_12_given_3[:, :, ind_3] = joint_pmf[:,:,ind_3] \
                                        / np.sum(joint_pmf[:,:,ind_3])
    cond_23_given_1 = np.zeros((2,2,2))
    for ind_1 in range(2):
        cond_23_given_1[:, :, ind_1] = joint_pmf[ind_1,:,:] \
                                        / np.sum(joint_pmf[ind_1,:,:])
    cond_13_given_2 = np.zeros((2,2,2))
    for ind_2 in range(2):
        cond_13_given_2[:, :, ind_2] = joint_pmf[:,ind_2,:] \
                                        / np.sum(joint_pmf[:,ind_2,:])
    return cond_12_given_3,cond_23_given_1,cond_13_given_2

cond_12_given_3,cond_23_given_1,cond_13_given_2 = \
                    conditional_2_stock_given_1(joint_pmf)

print(cond_12_given_3)
print(cond_23_given_1)
print(cond_13_given_2)

[[[0.4974359  0.06329114]
  [0.16410256 0.19831224]]
 [[0.25641026 0.18987342]
  [0.08205128 0.54852321]]]

[[[0.5078534  0.20746888]
  [0.07853403 0.18672199]]
 [[0.16753927 0.06639004]
  [0.2460733  0.53941909]]]

[[[0.46859903 0.14222222]
  [0.07246377 0.20888889]]
 [[0.24154589 0.07111111]
  [0.2173913  0.57777778]]]
```