Homework 4 Solutions

1. (10 points) *Basketball*

a. Let X_i is the outcome of the i^{th} shot, and it is 0 if the shot is missed, and 1 if it is made. Since we know that the 12^{th} shot is made the information on the 2^{nd} shot is irrelevant. More formally, $\mathrm{P}(X_{14}|X_{12},X_2) = \frac{\mathrm{P}(X_{14},X_{12},X_2)}{\mathrm{P}(X_{12},X_2)} = \frac{\mathrm{P}(X_1)\mathrm{P}(X_2|X_1)...\mathrm{P}(X_{14}|X_{13},...,X_1)}{\mathrm{P}(X_1)\mathrm{P}(X_2|X_1)...\mathrm{P}(X_{12}|X_{11},...,X_1)} = \mathrm{P}(X_{14}|X_{13})\mathrm{P}(X_{13}|X_{12})$, here we use chain rule and the fact that $\mathrm{P}(X_n|X_{n-1},X_{n-2}) = \mathrm{P}(X_n|X_{n-1})$, as described. It is also independent of the future shots. We don't know what happens to the 13^{th} shot so we take that into account and consider all possible cases for that:

$$\mathrm{P}(X_{14}=1|X_2=1,X_{12}=1) = \mathrm{P}(X_{14}=1|X_{12}=1) \tag{1}$$
$$= \mathrm{P}(X_{14}=1,X_{13}=1|X_{12}=1) + \mathrm{P}(X_{14}=1,X_{13}=0|X_{12}=1) \tag{2}$$
$$= \mathrm{P}(X_{14}=1|X_{13}=1)P(X_{13}=1|X_{12}=1) + \mathrm{P}(X_{14}=1|X_{13}=0)P(X_{13}=0|X_{12}=1) \tag{3}$$
$$= 0.6\cdot 0.6 + 0.3\cdot 0.4 = 0.48 \tag{4}$$

b. We can reverse the order of influence by Bayes' theorem:

$$\mathrm{P}(X_{n-1}|X_n) = \mathrm{P}(X_n|X_{n-1})\mathrm{P}(X_{n-1})/\mathrm{P}(X_n) \tag{5}$$

From this we can get $\mathrm{P}(X_{n-1}|X_n,X_{n+1}) = \mathrm{P}(X_{n-1}|X_n)$ and use the chain formula to get independence beyond 3^{rd} shot. Therefore in the reverse relation, given the outcome of the 3^{rd} shot, the outcome of the 1^{st} is independent of the 9^{th}. We have,

$$\mathrm{P}(X_1=1|X_3=0,X_9=0) = \mathrm{P}(X_1=1|X_3=0) \tag{6}$$
$$= \mathrm{P}(X_1=1,X_2=1|X_3=0) + \mathrm{P}(X_1=1,X_2=0|X_3=0) \tag{7}$$
$$= \mathrm{P}(X_1=1|X_2=1)\mathrm{P}(X_2=1|X_3=0) + \mathrm{P}(X_1=1|X_2=0)\mathrm{P}(X_2=0|X_3=0) \tag{8}$$
$$= \mathrm{P}(X_3=0|X_2=1)\mathrm{P}(X_2=1|X_1=1)\frac{P(X_1=1)}{P(X_3=0)} + \mathrm{P}(X_3=0|X_2=0)\mathrm{P}(X_2=0|X_1=1)\frac{P(X_1=1)}{P(X_3=0)} \tag{9}$$
$$= (0.4\cdot 0.6\cdot 0.4 + 0.7\cdot 0.4\cdot 0.6)/0.574 = 0.362 \tag{10}$$

c. *Solving directly.* Given $\mathrm{P}(X_1=1) = 0.4$. Let K be the number of shots he makes in a row after the first shot no matter what the outcome of the first shot is. For example the set $K=2$ contains the following two events $\{1110, 0110\}$. Therefore $P(K=k) = P(K=k, X_1=1) + P(K=k, X_1=0) = 0.4\cdot(0.6)^k 0.4 + 0.6\cdot 0.3\cdot(0.6)^{k-1}0.4 = 0.28\cdot(0.6)^k$:

$$\mathrm{E}(K) = \sum_{k=1}^{\infty} kP(K=k) = \sum_{k=1}^{\infty} k\cdot 0.28\cdot(0.6)^k \tag{11}$$
$$= 0.6\cdot 0.28\sum_{k=1}^{\infty} k\cdot(0.6)^{k-1} = 0.168\cdot 1/(1-0.6)^2 = 1.05 \tag{12}$$

Solving through iterated expectations. Note that given $X_1 = 1$ & $X_2 = 1$, and $X_1 = 0$ &

$X_2 = 1$, K is a geometric random variable with pmf $q(k) = (0.6)^k 0.4$.

$$
\begin{aligned}
\mathrm{E}(K) &= \mathrm{E}(\mathrm{E}(K|X_1, X_2)) && (13)\\
&= \mathrm{E}(K|X_1 = 1, X_2 = 1)\mathrm{P}(X_1 = 1)\mathrm{P}(X_2 = 1|X_1 = 1) && (14)\\
&+ \mathrm{E}(K|X_1 = 0, X_2 = 1)\mathrm{P}(X_1 = 0)\mathrm{P}(X_2 = 1|X_1 = 0) && (15)\\
&= \frac{1}{0.4}0.4 \cdot 0.6 + \frac{1}{0.4}0.6 \cdot 0.3 = 1.05 && (16)
\end{aligned}
$$

Remark: For both parts a and b, it is easier to visualize the solution by drawing a tree diagram of possible cases and multiple over the desired paths.

2. (10 points) *Model*

a. Note that for a fixed n and c

$$
\sum_{k=1}^{n} \binom{n}{k} c^k (1-c)^{n-k} = 1, \tag{17}
$$

since it is the sum of a binomial pmf over all its possible values.

To compute the marginal of N we sum over k and c. For $n \in \{0, 1, \ldots, 20\}$,

$$
\begin{aligned}
p_N(n) &= \sum_{c \in \left\{\frac{1}{4}, \frac{4}{5}\right\}} \sum_{k=1}^{n} p_{N,C,K}(n, c, k) && (18)\\
&= \left(\frac{1}{60} + \frac{1}{30}\right) \sum_{k=1}^{n} \binom{n}{k} c^k (1-c)^{n-k} && (19)\\
&= \frac{1}{60} + \frac{1}{30} \quad \text{by (17)} && (20)\\
&= \frac{1}{20}. && (21)
\end{aligned}
$$

Otherwise the pmf is zero. N is a discrete uniform random variable between 1 and 20.

To compute the marginal of C we sum over k and c. For $c = 1/4$,

$$
\begin{aligned}
p_C\left(\frac{1}{4}\right) &= \sum_{n=1}^{20} \sum_{k=1}^{n} p_{N,C,K}\left(n, \frac{1}{4}, k\right) && (22)\\
&= \sum_{n=1}^{20} \frac{1}{30} \sum_{k=1}^{n} \binom{n}{k} \left(\frac{1}{4}\right)^k \left(1 - \frac{1}{4}\right)^{n-k} && (23)\\
&= \sum_{n=1}^{20} \frac{1}{30} \quad \text{by (17)} && (24)\\
&= \frac{2}{3}, && (25)\\
p_C\left(\frac{4}{5}\right) &= \frac{1}{3} \quad \text{by the same argument.} && (26)
\end{aligned}
$$

The pmf is zero otherwise.

To check whether the random variables are independent, we compute their joint pmf and show that it factors into the marginal pmfs. For $n \in \{0, 1, \ldots, 20\}$,

$$p_{N,C}\left(n, \frac{1}{4}\right) = \sum_{k=1}^{n} p_{N,C,K}\left(n, \frac{1}{4}, k\right) \tag{27}$$

$$= \frac{1}{30} \sum_{k=1}^{n} \binom{n}{k} \left(\frac{1}{4}\right)^{k} \left(1 - \frac{1}{4}\right)^{n-k} \tag{28}$$

$$= \frac{1}{30} \quad \text{by (17)} \tag{29}$$

$$= p_N(n)\, p_C\left(\frac{1}{4}\right), \tag{30}$$

$$p_{N,C}\left(n, \frac{4}{5}\right) = \frac{1}{60} \quad \text{by the same argument} \tag{31}$$

$$= p_N(n)\, p_C\left(\frac{4}{5}\right). \tag{32}$$

Otherwise the joint pmf is zero, and so are the marginal pmfs. This establishes that N and C are independent.

b. Given N and C, K is a binomial random variable with parameters N and C. A simple example would be a situation in which you flip a coin a random number of times between 1 and 20. You choose a coin that has bias 1/3 with probability 1/3 and a coin that has bias 4/5 with probability 2/3. This choice is independent of the number of flips and each flip is independent of the other flips given the number of times you flip it and the coin you choose.

A more creative example could involve a basketball player. N would model the number of shots she takes in a game (it is equally likely to be any number between 1 and 20). C would be the precision, determined by whether her family are at the game (if they are then she is more motivated and makes 80% of her shots, if they are not then she just makes 25% of her shots; they come to see her 1 out of every 3 games). Finally, K would be the number of shots she makes. The independence assumptions are the same as in the coin-flip example above.

c. Applying the definition of conditional pmf, if $k \in \{0, 20\}$ and $n \in \{k, \}$

$$p_{N|C,K}\left(n | \frac{1}{4}, k\right) = \frac{p_{N,C,K}\left(n, \frac{1}{4}, k\right)}{p_{C,K}\left(\frac{1}{4}, k\right)} \tag{33}$$

$$= \frac{p_{N,C,K}\left(n, \frac{1}{4}, k\right)}{\sum_{n=k}^{20} p_{N,C,K}\left(n, \frac{1}{4}, k\right)} \tag{34}$$

$$= \frac{\frac{1}{60}\binom{n}{k}\left(\frac{1}{4}\right)^{k}\left(1-\frac{1}{4}\right)^{n-k}}{\frac{1}{60}\sum_{m=k}^{20}\binom{m}{k}\left(\frac{1}{4}\right)^{k}\left(1-\frac{1}{4}\right)^{m-k}} \tag{35}$$

$$= \frac{\frac{n!}{(n-k)!}\left(\frac{3}{4}\right)^{n}}{\sum_{m=k}^{20}\frac{m!}{(m-k)!}\left(\frac{3}{4}\right)^{m}}, \tag{36}$$

$$p_{N|C,K}\left(n | \frac{4}{5}, k\right) = \frac{\frac{m!}{(m-k)!}\left(\frac{1}{5}\right)^{n}}{\sum_{m=k}^{20}\frac{m!}{(m-k)!}\left(\frac{1}{5}\right)^{m}} \quad \text{by the same argument.} \tag{37}$$

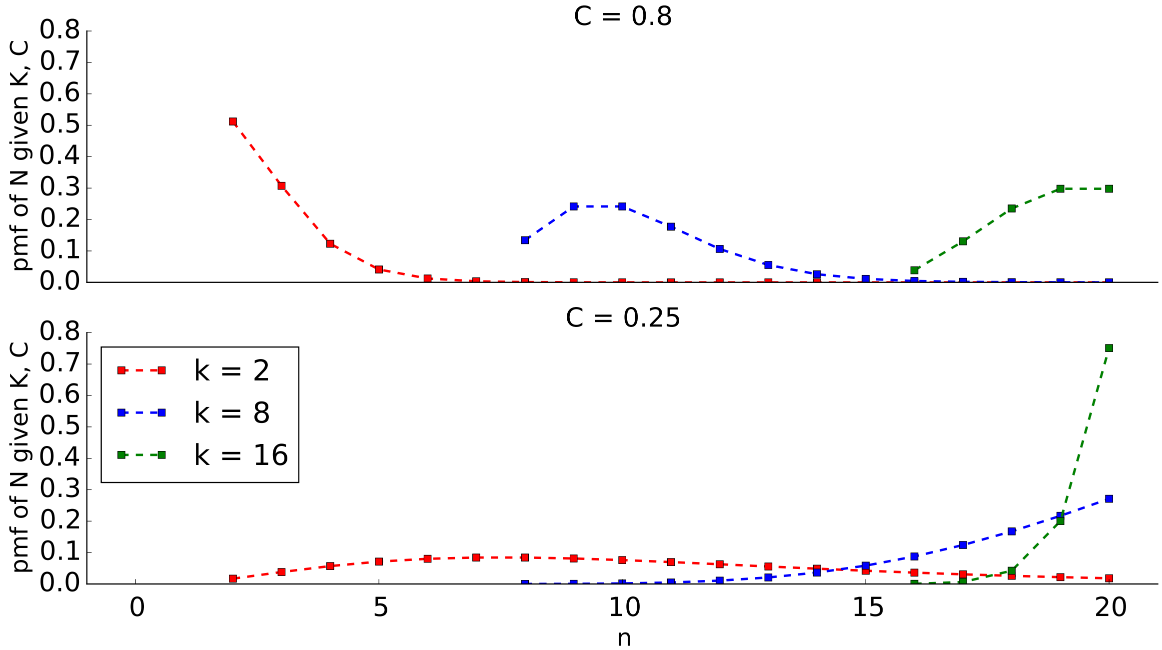

Figure 1: $p_{N|C,K}(n|c,k)$ in Problem 2.

d. The graphs are shown on Figure 1. For both values of C, higher values of K skew the distribution towards higher values of C (the distribution has non-zero values between K and 20). In terms of the example, a higher number of made shots increases the probability of higher numbers of taken shots. $C = 4/5$ skews the distribution towards smaller values of N for every fixed K with respect to $C = 1/4$. For a given number of made shots, if the probability of making each shot is higher then it is more likely that you have needed less attempts to make k shots.

```
import math
import matplotlib.pyplot as plt
import numpy as np

n_max = 20

def p_N_given_C_K(n, c, k):
  numerator = np.math.factorial(n) * np.math.pow(c,n) / np.math.factorial(n - k)
  denominator = 0
  for m in range(k, n_max + 1):
    aux = np.math.factorial(m) / np.math.factorial(m - k)
    denominator = denominator + np.math.factorial(m) * np.math.pow(c,m) / np.math.factorial(m - k)
  return numerator/denominator

c_val = [1./4, 4./5]
k_val = [2, 8, 16]
pmf_list = []
for i_c in range(len(c_val)):
  c = c_val[i_c]
  pmf_list.append([])
```

```
for i_k in range(len(k_val)):
  k = k_val[i_k]
  n_range = range(k, n_max + 1)
  pmf = np.zeros(len(n_range))
  for i_n in range(len(n_range)):
    n = n_range[i_n]
    pmf[i_n] = p_N_given_C_K(n, c, k)
  pmf_list[i_c].append(pmf)
```

3. *Router.*

   a. Let N be the total number of packets and N_1 the packets routed to connection 1. To find the mean of N_1 we use iterated expectation. Conditioned on $N = n$ N_1 is binomial with parameters n and p,

$$\mathrm{E}(N_1) = \mathrm{E}\left(\mathrm{E}\left(N_1|N\right)\right) = \mathrm{E}\left(Np\right) = p\,\lambda. \tag{38}$$

   We also use iterated expectation to find the mean square:

$$\mathrm{E}(N_1^2) = \mathrm{E}\left(\mathrm{E}\,N_1^2|N\right) = \mathrm{E}\left(N^2p^2 + Np\,(1-p)\right) = 2p^2\lambda^2 + p\,(1-p)\,\lambda. \tag{39}$$

   Finally,

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{E}(N_1^2) - \mathrm{E}^2(N_1) &(40)\\ &= 2\lambda^2p^2 + \lambda p\,(1-p) - \lambda^2p^2 = \lambda p. &(41) \end{aligned}$$

   b. From part (a), both the mean and variance are λp, so we suspect that N_1 is Poisson. To verify this we use the law of total probability. For any integer $n_1 \geq 0$,

$$\begin{aligned} p_{N_1}(n_1) &= \sum_{k=n_1}^{\infty} \mathrm{P}\{N_1 = n_1|N = k\}\mathrm{P}\{N = k\} \\ &= \sum_{k=n_1}^{\infty} \binom{k}{n_1} p_1^n (1-p)^{k-n_1} \frac{\lambda^k}{k!} e^{-\lambda} \\ &= \frac{\lambda^{n_1} p_1^n e^{-\lambda}}{n_1!} \sum_{k=n_1}^{\infty} \frac{k!(1-p)^{k-n_1}\lambda^{k-n_1}}{(k-n_1)!} \\ &= \frac{(\lambda p)_1^n}{n_1!} e^{-\lambda} \sum_{j=0}^{\infty} \frac{((1-p)\lambda)^j}{j!} \\ &= \frac{(\lambda p)_1^n}{n_1!} e^{-\lambda p}. \end{aligned}$$

   Thus N_1 is Poisson with parameter λp. Similarly, $N - N_1$ is Poisson with parameter $\lambda(1-p)$. The Poisson input traffic is divided into two Poisson output traffics.

4. (10 points) *Cheap GPS*

   a. Applying Chebyshev's inequality

$$\mathrm{P}\left(|X_i - d_i| > b\right) = \mathrm{P}\left(|Z_i| > \Delta_1\right) \leq \frac{\sigma^2}{\Delta_1^2} = \frac{1}{\Delta_1^2}. \tag{42}$$

$1/\Delta^2 = 0.01$ implies

$$\Delta_1 = 10. \tag{43}$$

b. We have

$$|Y_i - d_i| = \left| \sum_{j=i-m}^{i} Z_i + d_j - md_i \right| \tag{44}$$

$$= \left| \frac{1}{m} \sum_{j=i-m}^{i} Z_i \right| + \left| \frac{1}{m} \sum_{j=i-m}^{i} d_j - md_i \right| \quad \text{by the triangle inequality.} \tag{45}$$

In a second Mary can only run 2 meters, so

$$|d_j - d_{j-1}| < 2i \tag{46}$$

for any j. Using this,

$$\left| \frac{1}{m} \sum_{j=i-m}^{i} d_j - md_i \right| \leq \frac{1}{m} \sum_{j=i-m}^{i-1} |d_j - d_i| \quad \text{by the triangle inequality} \tag{47}$$

$$\leq \frac{1}{m} \sum_{j=1}^{m-1} 2j \quad \text{by (46)} \tag{48}$$

$$= m - 1 \quad \text{by the suggested identity.} \tag{49}$$

Since $|Y_i - d_i| \leq m - 1 + \left| \frac{1}{m} \sum_{j=i-m}^{i} Z_i \right|$

$$\mathrm{P}\left(|Y_i - d_i| > \Delta_2\right) \leq \mathrm{P}\left(m - 1 + \left| \frac{1}{m} \sum_{j=i-m}^{i} Z_i \right| > \Delta_2 \right) = \mathrm{P}\left(\left| \frac{1}{m} \sum_{j=i-m}^{i} Z_i \right| > \Delta_2 + 1 - m \right). \tag{50}$$

Because the Z_i are independent, the variance of their sum is the sum of the individual variances,

$$\mathrm{Var}\left(\frac{1}{m} \sum_{j=i-m}^{i} Z_i \right) = \frac{1}{m^2} \sum_{j=i-m}^{i} \mathrm{Var}\left(Z_i\right) \tag{51}$$

$$= \frac{\sigma^2}{m}. \tag{52}$$

By Chebyshev's inequality,

$$\mathrm{P}\left(\left| \sum_{j=i-m}^{i} Z_i > \Delta_2 + 1 - m \right| \right) \leq \frac{\sigma^2}{m\left(\Delta_2 + 1 - m\right)^2} = \frac{1}{m\left(\Delta_2 + 1 - m\right)^2}. \tag{53}$$

$\frac{1}{m(\Delta_2+1-m)^2} = 0.01$ implies

$$\Delta_2 = \frac{10}{\sqrt{m}} + m - 1. \tag{54}$$

c. Jennifer wants to increase the precision. For some values of m, $\Delta_2 < \Delta_1$. Taking $m = 4$

$$\Delta_2 = 8 < \Delta_1. \tag{55}$$

d. The best m should minimize Δ_2. We compute

$$\frac{\mathrm{d}\Delta_2}{\mathrm{d}m} = -\frac{5}{m\sqrt{m}} + 1, \tag{56}$$

$$\frac{\mathrm{d}^2\Delta_2}{\mathrm{d}m^2} = \frac{7.5}{m^2\sqrt{m}}. \tag{57}$$

The second derivative is positive so the function is convex. Setting the first derivative to zero we determine that the best m is

$$m^* = (5)^{\frac{2}{3}} \approx 2.92. \tag{58}$$

The precision for $m = 2$ is $\Delta_2 = 8.07$, whereas for $m = 3$ it is $\Delta_2 = 7.77$. The best precision that can be achieved is consequently $\Delta_2 = 7.77$ for $m = 3$.

5. (10 points) *Iterated expectation for random vectors*

a. $\mathrm{E}(\mathbf{X}_{\mathcal{I}}|\mathbf{X}_{\mathcal{J}})$ is a random vector, for $i \in \mathcal{I}$ its i^{th} component is $[\mathrm{E}(\mathbf{X}_{\mathcal{I}}|\mathbf{X}_{\mathcal{J}})]_i = \mathrm{E}(X_i|\mathbf{X}_{\mathcal{J}}) - \int_{\text{Range of } x_i} x_i f_{X_i|\mathbf{X}_{\mathcal{J}}}(x_i)\, dx_i$. From this, combined with the known version of the iterated expectations, it is easy to deduce:

$$\mathrm{E}(\mathrm{E}(X_i|\mathbf{X}_{\mathcal{J}})) = \mathrm{E}(X_i) \tag{59}$$

$$\implies \mathrm{E}(\mathrm{E}(\mathbf{X}_{\mathcal{I}}|\mathbf{X}_{\mathcal{J}})) = \mathrm{E}(\mathbf{X}_{\mathcal{I}}) \tag{60}$$

b. Given N and C, K is a binomial random variable with parameters N and C. And N and C are independent.

$$\mathrm{E}(K) = \mathrm{E}(\mathrm{E}(K|C,N)) = \mathrm{E}(NC) = \mathrm{E}(N)\,\mathrm{E}(C) \tag{61}$$

$$= [\sum_{n=1}^{20} n \cdot \frac{1}{20}] \cdot [\frac{1}{4} \cdot \frac{2}{3} + \frac{4}{5} \cdot \frac{1}{3}] = 4.55 \tag{62}$$

$$\mathrm{E}(K^2) = \mathrm{E}(\mathrm{E}(K^2|C,N)) = \mathrm{E}(N^2C^2 + NC(1-C)) \tag{63}$$

$$= \mathrm{E}(N^2)\,\mathrm{E}(C^2) + \mathrm{E}(N)[\mathrm{E}(C) - \mathrm{E}(C^2)] \tag{64}$$

$$= [\sum_{n=1}^{20} n^2 \cdot \frac{1}{20}] \cdot [\frac{1}{4^2} \cdot \frac{2}{3} + \frac{4^2}{5^2} \cdot \frac{1}{3}] + 10.5[\frac{13}{30} + 0.255] = 43.82 \tag{65}$$

$$\implies \mathrm{Var}(K) = 43.82 - (4.55)^2 \approx 23.12 \tag{66}$$