Assignment 4

Benjamin Jakubowski

October 9, 2015

1. Basketball

a. Probability Curry makes his 14^{th} shot

Let X_i represent Curry's i^{th} shot. Then

$$X_i = \begin{cases} 1 & \text{if he makes it} \\ 0 & \text{if he misses} \end{cases}$$

Now,

$$P_{X_{i+1}}(1) = \begin{cases} .6 & \text{if } X_i = 1 \\ .3 & \text{if } X_i = 0 \end{cases}$$

Thus, the only shot that matters when predicting the $i+1^{st}$ shot is the outcome of the i^{th} shot. Therefore, when predicting the outcome of the 14^{th} shot given the outcomes of the 2^{nd} and 12^{th}, the outcome of the 12^{th} shot provides all necessary information- knowing the outcome of the second shot gives no additional information. (More formally, let $m, n \in \{1, 2, ...j-1\}$ and $m > n$. Then $P(X_j|X_m, X_n) = P(X_j|X_m)$.)
Now, since $X_{12} = 1$,

$$P_{X_{13}|X_{12}=1}(x) = \begin{cases} .6 & \text{for } x = 1 \\ .4 & \text{for } x = 0 \end{cases}$$

Moreover,

$$P_{X_{14}|X_{13}=1}(1) = .6$$
$$P_{X_{14}|X_{13}=0}(1) = .3$$

Therefore,

$$\begin{aligned} P_{X_{14}}(1) &= P_{X_{14}|X_{13}=0}(1) \cdot P_{X_{13}|X_{12}=1}(0) + P_{X_{14}|X_{13}=1}(1) \cdot P_{X_{13}|X_{12}=1}(1) \\ &= .3 \cdot .4 + .6 \cdot .6 = .12 + .36 = .48 \end{aligned}$$

b. Probability Curry made his 1^{st} shot

First, note $P(X_1 = 1|X_3 = 0, X_9 = 0) = P(X_1 = 1|X_3 = 0)$. Then $P(X_1 = 1|X_3 = 0)$

$$
\begin{aligned}
&= \frac{P(X_3 = 0|X_1 = 1)P(X_1 = 1)}{P(X_3 = 0)} \\
&= \frac{(P(X_3 = 0|X_2 = 0)P(X_2 = 0|X_1 = 1) + P(X_3 = 0|X_2 = 1)P(X_2 = 1|X_1 = 1))\,P(X_1 = 1)}{P(X_3 = 0)}
\end{aligned}
$$

And

$$
\begin{aligned}
P(X_3 = 0) = P(X_3 = 0|X_2 = 0)&P(X_2 = 0|X_1 = 0)P(X_1 = 0) \\
&+ P(X_3 = 0|X_2 = 1)P(X_2 = 1|X_1 = 0)P(X_1 = 0) \\
&+ P(X_3 = 0|X_2 = 0)P(X_2 = 0|X_1 = 1)P(X_1 = 1) \\
&+ P(X_3 = 0|X_2 = 1)P(X_2 = 1|X_0 = 1)P(X_1 = 1) \\
= .7 \cdot .7 \cdot .6 &+ .4 \cdot .3 \cdot .6 + .7 \cdot .4 \cdot .4 + .4 \cdot .6 \cdot .4 \\
= .294 + .072 &+ .112 + .096 \\
= .574&
\end{aligned}
$$

Thus, $P(X_1 = 1|X_3 = 0)$

$$
\begin{aligned}
&= \frac{(P(X_3 = 0|X_2 = 0)P(X_2 = 0|X_1 = 1) + P(X_3 = 0|X_2 = 1)P(X_2 = 1|X_1 = 1))\,P(X_1 = 1)}{P(X_3 = 0)} \\
&= \frac{(.7 \cdot .4 + .4 \cdot .6) \cdot .4}{.574} \\
&= \frac{.52 \cdot .4}{.574} = .36237
\end{aligned}
$$

c. Expected number of consecutive baskets

Let Y be the number of shots made in a row following the first shot. Then

$$
\begin{aligned}
E(Y) &= E(E(Y|X_1)) \\
&= E(Y|X_1 = 0) \cdot P(X_1 = 0) + E(Y|X_1 = 1) \cdot P(X_1 = 1) \\
&= E(Y|X_1 = 0) \cdot .6 + E(Y|X_1 = 1) \cdot .4
\end{aligned}
$$

Next, note

$$E(Y|X_1 = 0) = \sum_{y=0}^{\infty} y \cdot P(Y = y|X_1 = 0)$$

and

$$P_{Y|X_1=0}(y) = \begin{cases} .7 & \text{for } y = 0 \\ .3 \cdot .6^{y-1} \cdot .4 & \text{for } y \geq 1 \end{cases}$$

Therefore,

$$\begin{aligned}
E(Y|X_1 = 0) &= \sum_{y=0}^{\infty} y \cdot P(Y = y|X_1 = 0) \\
&= 0 \cdot .7 + \sum_{y=1}^{\infty} (y \cdot .3 \cdot .6^{y-1} \cdot .4) \\
&= .3 \cdot .4 \cdot \sum_{y=1}^{\infty} (y \cdot .6^{y-1}) \\
&= .3 \cdot .4 \cdot \frac{1}{.6} \cdot \sum_{y=1}^{\infty} (y \cdot .6^{y}) \\
&= .3 \cdot .4 \cdot \frac{1}{.6} \cdot \frac{.6}{(1-.6)^{\cdot 2}} \\
&= .75
\end{aligned}$$

Similarly, note

$$E(Y|X_1 = 1) = \sum_{y=0}^{\infty} y \cdot P(Y = y|X_1 = 1)$$

and

$$P_{Y|X_1=1}(y) = .6^{y} \cdot .4 \text{ for } y \geq 0$$

Therefore

$$\begin{aligned}
E(Y|X_1 = 1) &= \sum_{y=0}^{\infty} y \cdot P(Y = y|X_1 = 1) \\
&= \sum_{y=0}^{\infty} (y \cdot .6^{y} \cdot .4) \\
&= .4 \cdot \sum_{y=0}^{\infty} (y \cdot .6^{y}) \\
&= .4 \cdot \frac{.6}{(1-.6)^{\cdot 2}} \\
&= 1.5
\end{aligned}$$

Finally,

$$E(Y) = E(Y|X_1 = 0) \cdot .6 + E(Y|X_1 = 1) \cdot .4 = .75 \cdot .6 + 1.5 \cdot .4 = 1.05$$

Therefore, we expect Curry to make 1.05 shots in a row regardless of whether he makes the first shot or not.

2. Model

a. Finding P_N and P_C

First, we find P_N:

$$\begin{aligned}
P_N(n) &= \sum_{k=0}^{n} \sum_{c \in \{1/4, 4/5\}} P_{N,C,K}(n,c,k) \\
&= \sum_{k=0}^{n} \left[\frac{1}{30} \binom{n}{k} (1/4)^k (3/4)^{n-k} + \frac{1}{60} \binom{n}{k} (4/5)^k (1/5)^{n-k} \right] \\
&= \frac{1}{30} \sum_{k=0}^{n} \left[\binom{n}{k} (1/4)^k (3/4)^{n-k} \right] + \frac{1}{60} \sum_{k=0}^{n} \left[\binom{n}{k} (4/5)^k (1/5)^{n-k} \right] \\
&= \frac{1}{30} \cdot 1 + \frac{1}{60} \cdot 1 = 1/20
\end{aligned}$$

Next, we find P_C:

$$\begin{aligned}
P_C(c) &= \begin{cases} \displaystyle\sum_{n=1}^{20} \left[\sum_{k=0}^{n} \left(\frac{1}{30} \binom{n}{k} (1/4)^k (3/4)^{n-k} \right) \right] & \text{for } c = 1/4 \\ \displaystyle\sum_{n=1}^{20} \left[\sum_{k=0}^{n} \left(\frac{1}{60} \binom{n}{k} (4/5)^k (1/5)^{n-k} \right) \right] & \text{for } c = 4/5 \end{cases} \\
&= \begin{cases} \displaystyle\sum_{n=1}^{20} \frac{1}{30} = \frac{20}{30} = \frac{2}{3} \text{ for } c = 1/4 \\ \displaystyle\sum_{n=1}^{20} \frac{1}{60} = \frac{20}{60} = \frac{1}{3} \text{ for } c = 4/5 \end{cases}
\end{aligned}$$

Finally, we will show $P_{N,C}(n,c) = P_N(n) \cdot P_C(c)$ (and, as such, C and N are independent):

$$\begin{aligned}
P_{N,C}(n,c) &= \sum_{k=0}^{n} P_{N,C,K}(n,c,k) \\
&= \begin{cases} \displaystyle\sum_{k=0}^{n} \frac{1}{30} \binom{n}{k} (1/4)^k (3/4)^{n-k} \text{ for } c = 1/4 \\ \displaystyle\sum_{k=0}^{n} \frac{1}{60} \binom{n}{k} (4/5)^k (1/5)^{n-k} \text{ for } c = 4/5 \end{cases} \\
&= \begin{cases} \dfrac{1}{30} \text{ for } c = 1/4 \\ \dfrac{1}{60} \text{ for } c = 4/5 \end{cases}
\end{aligned}$$

But then note:

$$P_{N,C}(n,c) = \begin{cases} \dfrac{1}{30} = \dfrac{1}{20}\cdot\dfrac{2}{3} = P_N(n)\cdot P_C(c) \text{ for } c = 1/4, n \in \{1,2,...20\} \\ \dfrac{1}{60} = \dfrac{1}{20}\cdot\dfrac{1}{3} = P_N(n)\cdot P_C(c) \text{ for } c = 4/5, n \in \{1,2,...20\} \end{cases}$$

Thus, C and N are independent

B. Real-life example

This model could correspond to the following scenario:

1. C: Pick one of two biased coins. Specifically, pick:
   - Coin 1 with probability 2/3. This coin flips heads with probability 1/4.
   - Coin 2 with probability 1/3. This coin flips heads with probability 4/5.
2. N: Spin a fair (i.e. random uniform) spinner, numbered 1 to 20, to determine how many times to flip said coin.
3. K: Count the number of heads in your N flips.

This scenario would be well modeled by the joint pmf given in the problem.

C. Condition PMF $P_{N|K,C}$

First, note

$$P(N=n|C=c,K=k) = \frac{P(N=n,C=c,K=k)}{P(C=c,K=k)}$$

Thus, to proceed we first determine:

$$\begin{aligned} P_{C,K}(c,k) &= \sum_{n=0}^{20} P_{N,C,K}(n,c,k) \\ &= \begin{cases} \displaystyle\sum_{n=1}^{20} 1/30\binom{n}{k}(1/4)^k(3/4)^{n-k} \text{ for } c = 1/4 \\ \displaystyle\sum_{n=1}^{20} 1/60\binom{n}{k}(4/5)^k(1/5)^{n-k} \text{ for } c = 4/5 \end{cases} \end{aligned}$$

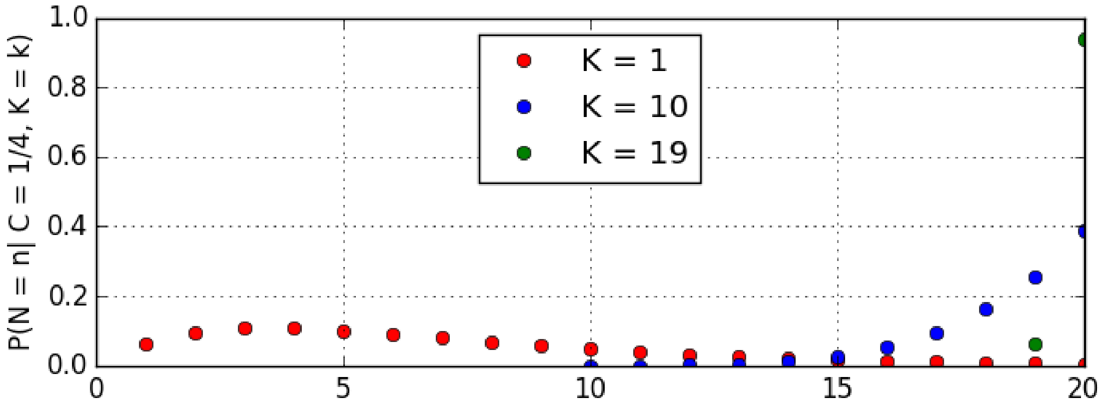
Therefore:

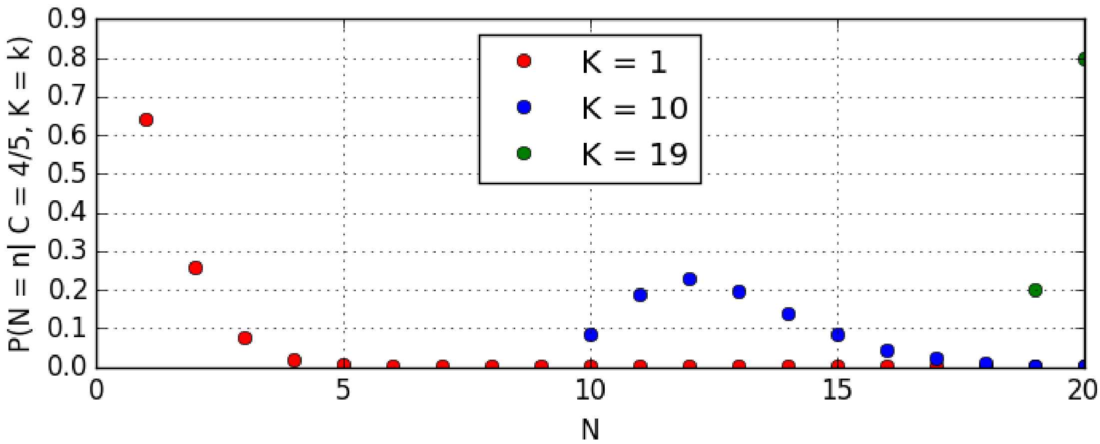
$$P_{N|C,K}(n|c,k) = \begin{cases} \dfrac{1/30\binom{n}{k}(1/4)^k(3/4)^{n-k}}{1/30(1/4)^k\sum_{n=1}^{20}\binom{n}{k}(3/4)^{n-k}} & \text{for } c = 1/4 \\ \dfrac{1/60\binom{n}{k}(4/5)^k(1/5)^{n-k}}{1/60(4/5)^k\sum_{n=1}^{20}\binom{n}{k}(1/5)^{n-k}} & \text{for } c = 4/5 \end{cases}$$

$$= \begin{cases} \dfrac{\binom{n}{k}(3/4)^{n-k}}{\sum_{n=1}^{20}\binom{n}{k}(3/4)^{n-k}} & \text{for } c = 1/4 \\ \dfrac{\binom{n}{k}(1/5)^{n-k}}{\sum_{n=1}^{20}\binom{n}{k}(1/5)^{n-k}} & \text{for } c = 4/5 \end{cases}$$

D. Plotting the distribution $P_{N|K,C}$

Below are plots of the distribution of N given $K = 1, K = 10, K = 19$ and $C = 1/4$ on one graph and $K = 1, K = 10, K = 19$ and $C = 4/5$ on another graph.





These results make intuitive sense in the context of the example posed in question b. First, comparing the PMFs when $C = 4/5$ and $C = 1/4$, it is clear that for any $K = k$ lower values of N are more likely when $C = 4/5$ than when $C = 1/4$. More specifically, it is apparent from the plots that given any $K = k$, for any value x such that $k \leq x < 20$

$$P(N \leq x | C = 1/4) < P(N \leq x | C = 4/5)$$

In the context of the problem, that makes sense- if you get heads with probability 4/5, you'd expect to need fewer flips to get k heads than you would if you had the coin with 1/4 probability of heads.

3. Router

a. Mean and variance of number of packets

Let N be the number of packets that arrive at the router in a second. Note $N \sim \text{Pois}(\lambda)$.
Let X be the number of packets routed through connection 1. Then $X \sim \text{B}(N, p)$.
Now we can find the mean and variance of X:

$$\begin{aligned}
E(X) &= E(E(X|N)) \\
&= E(N \cdot p) = p \cdot E(N) = p \cdot \lambda \\
Var(X) &= E(Var(X|N)) + Var(E(X|N)) \\
&= E(N \cdot p \cdot (1-p)) + Var(N \cdot p) \\
&= p \cdot (1-p) \cdot E(N) + p^2 \cdot Var(N) \\
&= p \cdot (1-p) \cdot \lambda + p^2 \cdot \lambda \\
&= p \cdot \lambda
\end{aligned}$$

Thus, $E(X) = Var(X) = p \cdot \lambda$.

b. PMF of number of packets

The PMF $P_X(x)$ is

$$P_X(x) = \sum_{n=0}^{\infty} P(X = x | N = n) \cdot P(N = n)$$

Note since $N \geq X$ by definition, this is equivalent to

$$
\begin{aligned}
P_X(x) &= \sum_{n=x}^{\infty} \left[\binom{n}{x} p^x (1-p)^{n-x} \cdot \frac{\lambda^n}{n!} e^{-\lambda} \right] \\
&= p^x e^{-\lambda} \cdot \sum_{n=x}^{\infty} \left[\frac{n!}{(n-x)!x!} \cdot (1-p)^{n-x} \cdot \frac{\lambda^n}{n!} \right] \\
&= \frac{p^x e^{-\lambda}}{x!} \cdot \sum_{n=x}^{\infty} \left[\frac{(1-p)^{n-x}}{(n-x)!} \cdot \lambda^n \right] \\
&= \frac{p^x e^{-\lambda}}{x!} \cdot \sum_{n=x}^{\infty} \left[\frac{(1-p)^{n-x}}{(n-x)!} \cdot \lambda^{n-x} \cdot \lambda^x \right] \\
&= \frac{p^x e^{-\lambda} \lambda^x}{x!} \cdot \sum_{n=x}^{\infty} \left[\frac{[(1-p)\lambda]^{n-x}}{(n-x)!} \right] \\
&= \frac{p^x e^{-\lambda} \lambda^x}{x!} \cdot \sum_{n=0}^{\infty} \left[\frac{[(1-p)\lambda]^{n}}{n!} \right] \\
&= \frac{p^x e^{-\lambda} \lambda^x}{x!} e^{(1-p)\cdot\lambda} \\
&= \frac{(p\lambda)^x}{x!} e^{\lambda - p\lambda - \lambda} = \frac{(p\lambda)^x}{x!} e^{-p\lambda}
\end{aligned}
$$

Thus, $P_X(x) = \frac{(p\lambda)^x}{x!} e^{-p\lambda}$ and $X \sim \text{Pois}(p\lambda)$.

4. Cheap GPS

a. Finding precision Δ_1

First, recall Jennifer has defined the precision of the location estimate as the smallest Δ such that the probability of the error being larger than Δ is smaller than 1%.
Thus, using X_i as an estimator of d_i, we want to find Δ such that:

$$
\begin{aligned}
P(|X_i - d_i| > \Delta) < .01 \\
P(|Z_i| > \Delta) < .01
\end{aligned}
$$

Using Chebyshev's inequality, we know

$$P(|Z_i - E(Z_i)| > a) \leq \frac{Var(Z_i)}{a^2}$$

Thus

$$P(|Z_i - E(Z_i)| = |Z_i - 0| = |Z_i| > \Delta) < \frac{1}{\Delta^2}$$

So $P(|Z_i| > \Delta) < 0.1$ when $\Delta \geq 10$. Hence, the precision Δ_1 of using X_i as an estimate of d_i is $\Delta_1 = 10$.

B. Finding precision Δ_2

First,

$$
\begin{aligned}
Y_i &= \frac{1}{m} \sum_{j=i-(m-1)}^{i} X_i \\
&= \frac{1}{m} \sum_{j=i-(m-1)}^{i} (d_j + Z_j) \\
&= \frac{1}{m} \sum_{j=i-(m-1)}^{i} (d_j) + \frac{1}{m} \sum_{j=i-(m-1)}^{i} (Z_j)
\end{aligned}
$$

Let's consider just

$$
\sum_{j=i-(m-1)}^{i} (d_j)
$$

First, since Mary never moves backward and has a maximum speed of 2 meters per second,

$$
d_i - 2(n) \leq d_{i-n} \leq d_i
$$

Thus, we can place the following bound on the sum of the d_js:

$$
\begin{aligned}
\sum_{j=i-(m-1)}^{i} (d_i - 2(i-j)) &\leq \sum_{j=i-(m-1)}^{i} d_j \leq \sum_{j=i-(m-1)}^{i} d_i \\
m \cdot d_i - 2 \sum_{j=i-(m-1)}^{i} (i-j) &\leq \sum_{j=i-(m-1)}^{i} d_j \leq \sum_{j=i-(m-1)}^{i} d_i \\
m \cdot d_i - 2 \sum_{j=0}^{m-1} (j) &\leq \sum_{j=i-(m-1)}^{i} d_j \leq \sum_{j=i-(m-1)}^{i} d_i \\
m \cdot d_i - 2\frac{(m-1)m}{2} &\leq \sum_{j=i-(m-1)}^{i} d_j \leq \sum_{j=i-(m-1)}^{i} d_i \\
m \cdot d_i - (m-1)m &\leq \sum_{j=i-(m-1)}^{i} d_j \leq \sum_{j=i-(m-1)}^{i} d_i
\end{aligned}
$$

Now recall

$$
Y_i = \frac{1}{m} \sum_{j=i-(m-1)}^{i} d_j + \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j
$$

Therefore, multiplying by $\frac{1}{m}$ and adding $\frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j$ throughout bounds Y_i:

$$
\frac{1}{m}(m \cdot d_i - (m-1)m) + \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \leq \; Y_i \; \leq \frac{1}{m} \sum_{j=i-(m-1)}^{i} d_i + \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j
$$

Simplifying, we get

$$d_i - (m-1) + \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \leq \; Y_i \; \leq \frac{1}{m} \cdot m \cdot d_i + \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j$$

$$-(m-1) + \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \leq \; Y_i - d_i \; \leq \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j$$

This implies

$$|Y_i - d_i| \leq \max \left\{ \left| -(m-1) + \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \right|, \left| \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \right| \right\}$$

But, by the triangle inequality,

$$\left| -(m-1) + \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \right| \leq |m-1| + \left| \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \right|$$

Thus, we have

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

That implies (noting $m \geq 1$)

$$P(|Y_i - d_i| > \Delta) \leq P\left(|m-1| + \left| \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \right| > \Delta \right) = P\left(\left| \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j \right| > \Delta - (m-1) \right)$$

Now let's let

$$X = \frac{1}{m} \sum_{j=i-(m-1)}^{i} Z_j$$

Then $E(X) = 0$ (since $E(Z_j) = 0$ and expectation is a linear operator), and (by the independence of the Z_js)

$$Var(X) = \sum_{j=i-(m-1)}^{i} Var\left(\frac{1}{m} Z_j\right) = \sum_{j=i-(m-1)}^{i} \frac{1}{m^2} Var(Z_j) = 0$$

Now, setting

$$P\left(|X| > \Delta - (m-1)\right) \leq 0.01$$

Chebyshev's inequality implies

$$0.01 \geq \frac{Var(X)}{(\Delta - (m-1))^2}$$

Thus

$$(\Delta_2 - (m-1))^2 = \frac{100}{m} \implies \Delta_2 = 10 \cdot m^{-1/2} + m - 1$$

c. Evaluating Δ_2 for $m = 4$

When $m = 4$,

$$\Delta_2 = 10 \cdot m^{-1/2} + m - 1 = 10 \cdot (4)^{-1/2} + 4 - 1 = 8$$

That's why Jennifer is using the running mean- it's an improved estimator of d_i (in the sense it is more precise, though it is unfortunately biased assuming Mary's speed is non-zero).

d. Minimizing Δ_2 fusing the running mean

First, let's express Δ_2 as a function of m:

$$\Delta_2(m) = 10 \cdot m^{-1/2} + m - 1$$

Then

$$\begin{aligned}
\Delta_2(1) &= 10 \\
\Delta_2(2) &\approx 8.071 \\
\Delta_2(3) &\approx 7.7735 \\
\Delta_2(4) &= 8
\end{aligned}$$

Finally, note $\mathrm{d}\Delta_2/\mathrm{d}m = -5m^{-3/2} + 1 > 0$ for all $n \in \mathbb{N}, n \geq 4$. Thus, Δ_2 is strictly increasing from 4 to ∞. Hence, the best precision Jennifer can achieve using a running mean is $\Delta_2(3) \approx 7.7735$.

5. Iterated Expectation for Random Vectors

a. Proof for disjoint random vectors

We will show that for any disjoint subvectors indexed by $\mathcal{I}, \mathcal{J} \subseteq \{1, 2, ..., n\}, \mathcal{I} \cap \mathcal{J} = \emptyset$,

$$E(E(\mathbf{X}_\mathcal{I} | \mathbf{X}_\mathcal{J})) = E(\mathbf{X}_\mathcal{I})$$

Note we will demonstrate this in the continuous case- the discrete case is similar (replacing integrals with sums).
First, let

$$\begin{aligned}
h(\vec{x_\mathcal{J}}) &:= E(\mathbf{X}_\mathcal{I} | \mathbf{X}_\mathcal{J} = \vec{x_\mathcal{J}}) \\
&= \int_\mathcal{I} \vec{x_\mathcal{J}} \cdot f_{\mathbf{X}_\mathcal{I}|\mathbf{X}_\mathcal{J}}(\vec{x_\mathcal{I}} | \vec{x_\mathcal{J}}) \mathbf{d}\vec{x_\mathcal{I}}
\end{aligned}$$

Before we proceed, let's clarify this abuse of notation:

$$\int_{\mathcal{I}} \text{ represents } \int_{x_{\mathcal{I}_1}=-\infty}^{\infty}\int_{x_{\mathcal{I}_2}=-\infty}^{\infty}...\int_{x_{\mathcal{I}_i}=-\infty}^{\infty}$$

$$\mathbf{d}\vec{x_{\mathcal{I}}} \text{ represents } \mathrm{d}x_{\mathcal{I}_i}\mathrm{d}x_{\mathcal{I}_{i-1}}...\ \mathrm{d}x_{\mathcal{I}_1}$$

Additionally, note that while $h(\vec{x_{\mathcal{J}}}) = E(\mathbf{X}_{\mathcal{I}}|\mathbf{X}_{\mathcal{J}} = \vec{x_{\mathcal{J}}})$ is a function mapping $\mathbb{R}^j \to \mathbb{R}^j$ (where j is the length of $\vec{x_{\mathcal{J}}}$), $h(\mathbf{X}_{\mathcal{J}}) = E(\mathbf{X}_{\mathcal{I}}|\mathbf{X}_{\mathcal{J}})$ is a random vector (since a function of a random vector is itself a random vector). Then

$$\begin{aligned}
E(E(\mathbf{X}_{\mathcal{I}}|\mathbf{X}_{\mathcal{J}})) &= E(h(\mathbf{X}_{\mathcal{J}}))\\
&= \int_{\mathcal{J}} h(\vec{x_{\mathcal{J}}}) \cdot f_{\mathbf{X}_{\mathcal{J}}}(\vec{x_{\mathcal{J}}})\mathbf{d}\vec{x_{\mathcal{J}}}\\
&= \int_{\mathcal{J}} \left(\int_{\mathcal{I}} \vec{x_{\mathcal{I}}} f_{\mathbf{X}_{\mathcal{I}}|\mathbf{X}_{\mathcal{J}}}(\vec{x_{\mathcal{I}}}|\vec{x_{\mathcal{J}}})\mathbf{d}\vec{x_{\mathcal{I}}}\right) \cdot f_{\mathbf{X}_{\mathcal{J}}}(\vec{x_{\mathcal{J}}})\mathbf{d}\vec{x_{\mathcal{J}}}\\
&= \int_{\mathcal{I}} \vec{x_{\mathcal{I}}} \left(\int_{\mathcal{J}} f_{\mathbf{X}_{\mathcal{I}}|\mathbf{X}_{\mathcal{J}}}(\vec{x_{\mathcal{I}}}|\vec{x_{\mathcal{J}}}) \cdot f_{\mathbf{X}_{\mathcal{J}}}(\vec{x_{\mathcal{J}}})\mathbf{d}\vec{x_{\mathcal{J}}}\right)\mathbf{d}\vec{x_{\mathcal{I}}}\\
&= \int_{\mathcal{I}} \vec{x_{\mathcal{I}}} f_{\mathbf{X}_{\mathcal{I}}}(\vec{x_t})\mathbf{d}\vec{x_{\mathcal{I}}}\\
&= E(\mathbf{X}_{\mathcal{I}})
\end{aligned}$$

B. Mean and Variance of K in Problem 2

First, from question 2 we know

$$\begin{aligned}
P_{K|N,C} &= \frac{P_{N,C,K}}{P_{N,C,}}\\
&= \binom{N}{k} C^k (1-C)^{N-k}
\end{aligned}$$

So

$$\begin{aligned}
E(K|N,C) &= \sum_{k=0}^{N} k\binom{N}{k} C^k (1-C)^{N-k}\\
&= C \cdot N
\end{aligned}$$

But then

$$
\begin{aligned}
E(K) &= E(E(K|N,C)) = E(CN) \\
&= \sum_{n=1}^{20} \sum_{c \in \{1/4, 4/5\}} c \cdot n \cdot P_{N,C}(n,c) \\
&= \sum_{n=1}^{20} \left[\frac{1}{4} \cdot n \cdot \frac{1}{30} + \frac{4}{5} \cdot n \cdot \frac{1}{60} \right] \\
&= \sum_{n=1}^{20} \frac{13}{600} \cdot n \\
&= \frac{13 \cdot 20 \cdot 21}{600 \cdot 2} = 4.55
\end{aligned}
$$

Next, let's find the variance of K. First, note

$$Var(K) = E(K^2) - (E(K))^2$$

Well, $E(K^2) = E(E(K^2|N,C))$, and

$$E(K^2|N,C) = \sum_{k=0}^{N} k^2 \binom{N}{k} C^k (1-C)^{N-k}$$

But this is just the second moment of a binomial (though, significantly, a binomial parameterized by random variables). Recall the variance of $B(N,C)$ is $NC(1-C)$. So

$$
\begin{aligned}
E(K^2|N,C) &= Var(K|N,C) + (E(K|N,C))^2 \\
&= NC(1-C) + (NC)^2
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E(K^2) &= E(E(K^2|N,C)) \\
&= \sum_{n=1}^{20} \sum_{c \in \{1/4, 4/5\}} \big((nc(1-c) + (nc)^2) P_{N,C}(n,c) \big) \\
&= \sum_{n=1}^{20} \left[\left(n \cdot \frac{1}{4} \cdot \frac{3}{4} + \left(n \cdot \frac{1}{4} \right)^2 \right) \frac{1}{30} + \left(n \cdot \frac{4}{5} \cdot \frac{1}{5} + \left(n \cdot \frac{4}{5} \right)^2 \right) \frac{1}{60} \right] \\
&= \sum_{n=1}^{20} \left[\left(\frac{3}{16} n + \frac{1}{16} n^2 \right) \frac{1}{30} + \left(\frac{4}{25} n + \frac{16}{25} n^2 \right) \frac{1}{60} \right] \\
&= 38.465
\end{aligned}
$$

So $Var(K) = E(K^2) - (E(K))^2 = 38.465 - (4.55)^2 = 17.7625$.