Probability and Statistics Homework #10

Ari Feiglin

Question 10.1:

A fair die is rolled X times, let Y be the sum of all the results.

(1) Compute $\mathbb{E}[Y]$

(2) Compute $\mathbb{E}[Y \mid X = 5]$

(3) Compute $\mathbb{E}[X \mid Y = 12]$

(1) We know that

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid X]]$$

Let X_i be the result of the ith roll, then:

$$Y = \sum_{i=1}^{X} X_i$$

This means that:

$$\mathbb{E}[Y \mid X] = \mathbb{E}\left[\sum_{i=1}^{X} X_i \,\middle|\, X\right] = \sum_{i=1}^{X} \mathbb{E}[X_i]$$

Since the die is fair, $X_i \sim \text{Unif}[6] \implies \mathbb{E}[X_i] = 3.5$.
Thus:

$$\mathbb{E}[Y] = \mathbb{E}\left[\sum_{i=1}^{X} 3.5\right] = 3.5 \cdot \mathbb{E}[X]$$

(2) As discussed before:

$$\mathbb{E}[Y \mid X] = \sum_{i=1}^{X} \mathbb{E}[X_i] = 3.5 \cdot X$$

This means that:

$$\mathbb{E}[Y \mid X = 5] = 3.5 \cdot 5 = \boxed{17.25}$$

Question 10.2:

A series of random variables, $\{X_n\}_{n\in\mathbb{N}}$, is defined as follows:

- $X_0 \coloneqq \lambda$
- $X_{n+1} \sim \text{Poi}\left(X_n\right)$

Answer the following:

(1) Compute the expected value and variance of X_n for every n,

(2) For every n and m, compute $\text{Cov}\left(X_n, X_m\right)$.

(1) We know by the law of total expectation:

$$\mathbb{E}\left[X_{n+1}\right] = \mathbb{E}\left[\mathbb{E}\left[X_{n+1} \mid X_n\right]\right]$$

And since $X_{n+1} \sim \text{Poi}\left(X_n\right)$, we know that $\mathbb{E}\left[X_{n+1} \mid X_n\right] = X_n$, so:

$$\mathbb{E}\left[X_{n+1}\right] = \mathbb{E}\left[X_n\right]$$

This means that the expected value of every X_n is equal. Specifically, this means that:

$$\mathbb{E}\left[X_n\right] = \mathbb{E}\left[X_0\right] = \mathbb{E}\left[\lambda\right]$$

So:

$$\boxed{\mathbb{E}\left[X_n\right] = \lambda}$$

Furthermore, we know that:

$$\text{Var}\left(X_n\right) = \mathbb{E}\left[\text{Var}\left(X_n \mid X_{n-1}\right)\right] + \text{Var}\left(\mathbb{E}\left[X_n \mid X_{n-1}\right]\right)$$

And since $X_n \sim \text{Poi}\left(X_{n-1}\right)$, we know that $\text{Var}\left(X_n \mid X_{n-1}\right) = X_{n-1}$, and as discussed above:

$$\mathbb{E}\left[X_n \mid X_{n-1}\right] = X_{n-1}$$

So:

$$\text{Var}\left(X_n\right) = \mathbb{E}\left[X_{n-1}\right] + \text{Var}\left(X_{n-1}\right) = \lambda + \text{Var}\left(X_{n-1}\right)$$

Which means that:

$$\text{Var}\left(X_n\right) = n \cdot \lambda + \text{Var}\left(X_0\right) = n \cdot \lambda + \text{Var}\left(\lambda\right)$$

So:

$$\boxed{\text{Var}\left(X_n\right) = n \cdot \lambda}$$

(2) We can assume without loss of generality that $n \leq m$ since covariance is commutative, so we can just swap n and m. So suppose $m = n + a$ where $a \geq 0$.
Notice that:

$$\begin{aligned}
\text{Cov}\left(X_n, \mathbb{E}\left[X_{n+a} \mid X_n\right]\right) &= \mathbb{E}\left[X_n \mathbb{E}\left[X_{n+a}[X_n]\right]\right] - \mathbb{E}\left[X_n\right]\mathbb{E}\left[\mathbb{E}\left[X_{n+a} \mid X_n\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[X_n \cdot X_{n+a}[X_n]\right]\right] - \mathbb{E}\left[X_n\right]\mathbb{E}\left[X_{n+a}\right] \\
&= \mathbb{E}\left[X_n \cdot X_{n+a}\right] - \mathbb{E}\left[X_n\right]\mathbb{E}\left[X_{n+a}\right] \\
&= \text{Cov}\left(X_n, X_{n+a}\right)
\end{aligned}$$

We can show inductively that $\mathbb{E}\left[X_{n+a} \mid X_n\right] = X_n$ for $a \geq 0$.

Base case: for $a = 0$, we need to show $\mathbb{E}\left[X_n \mid X_n\right] = X_n$, which is true trivially.
Inductive step: Suppose $\omega \in \Omega$, and let $x = X(\omega)$, then:

$$\begin{aligned}
\mathbb{E}\left[X_{n+a+1} \mid X_n\right](\omega) &= \mathbb{E}\left[X_{n+a+1} \mid X_n = x\right] \\
&= \sum_{y\in\mathbb{R}} \mathbb{P}\left(X_{n+a} = y \mid X_n = x\right) \cdot \mathbb{E}\left[X_{n+a+1} \mid X_{n+a} = y, X_n = x\right] \\
&= \sum_{y\in\mathbb{R}} \mathbb{P}\left(X_{n+a} = y \mid X_n = x\right) \cdot y \\
&= \mathbb{E}\left[X_{n+a} \mid X_n = x\right] \\
&= \mathbb{E}\left[X_{n+a} \mid X_n\right](\omega)
\end{aligned}$$

And by our inductive hypothesis, $\mathbb{E}\left[X_{n+a} \mid X_n\right] = X_n$, so:

$$\mathbb{E}\left[X_{n+a+1} \mid X_n\right] = X_n$$

As required.

So:

$$\mathrm{Cov}\left(X_n, X_{n+a}\right) = \mathrm{Cov}\left(X_n, \mathbb{E}\left[X_{n+a} \mid X_n\right]\right) = \mathrm{Cov}\left(X_n, X_n\right) = \mathrm{Var}\left(X_n\right) = n \cdot \lambda$$

So in general:

$$\boxed{\mathrm{Cov}\left(X_n, X_m\right) = \min\left\{n, m\right\} \cdot \lambda}$$

Question 10.3:

Suppose a conversation at a public phone distributes uniformly between 1 and 10 minutes, and is independent of other conversations. Alpha is currently at the public phone and speaks to a friend of theirs for X minutes. Beta arrives at the public phone in a uniform manner between 1 and X minutes. As soon as Alpha finishes, they leave the booth immediately and Beta begins to talk to their friend. What is the expected length of time Beta will be required to wait?

Instead of 10 minutes, suppose the number of minutes one can spend at the phone distributes uniformly in $[n]$. This means $X \text{Unif}[n]$.

Let Y be the time Beta arrives at the phone, which means $Y \sim \text{Unif}[X]$. The time Beta waits is equal to $X - Y$, since this represents the minutes between when Beta arrived and Alpha finishes. So the expected time Beta has to wait is:

$$\mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y]$$

Since X distributes uniformly over n, it has an expected value of $\frac{n+1}{2}$.

And we know since $Y \sim \text{Unif}[X]$

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}\left[\frac{X+1}{2}\right] = \frac{1}{2}\mathbb{E}[X] + \frac{1}{2} = \frac{n+3}{4}$$

So:

$$\mathbb{E}[X - Y] = \frac{n+1}{2} - \frac{n+3}{4} = \frac{n-1}{4}$$

So the expected length of time Beta will have to wait in our case where $n = 10$ is:

$$\boxed{\frac{9}{4}}$$

Question 10.4:

In a building there are 80 stairs. A person rolls a fair die 20 times and climbs up the same number of stairs as the sum of the results of the rolls. Find an upper bound to the probability that the person reaches the top of the building using Markov's and Chebyshev's inequalities.

Let X be the sum of all the dice. Let X_i be the result of the ith roll, so $X_i \text{Unif}\left[6\right]$. This means that:

$$X = \sum_{i=1}^{20} X_i$$

So:

$$\mathbb{E}\left[X\right] = \sum_{i=1}^{20} \mathbb{E}\left[X_i\right] = 20 \cdot 3.5 = 70$$

And since each roll is independent:

$$\text{Var}\left(X\right) = \sum_{i=1}^{20} \text{Var}\left(X_i\right) = 20 \cdot \frac{36-1}{12} = 58\frac{1}{3}$$

We want to find the probability that $X \geq 80$, which by Markov's identity has an upper bound of:

$$\mathbb{P}\left(X \geq 80\right) \leq \frac{\mathbb{E}\left[X\right]}{80} = \frac{70}{80} = \frac{7}{8}$$

That is, by Markov's inequality, we have a bound:

$$\boxed{\mathbb{P}\left(X \geq 80\right) \leq \frac{7}{8}}$$

And in preparation for Chebyshev:

$$\mathbb{P}\left(\left|X - \mathbb{E}\left[X\right]\right| \geq c\right) \geq \mathbb{P}\left(X \geq c + \mathbb{E}\left[X\right]\right) = \mathbb{P}\left(X \geq 70 + c\right)$$

So if we let $c = 10$, we have that by Chebyshev's inequality:

$$\mathbb{P}\left(X \geq 80\right) \leq \mathbb{P}\left(\left|X - \mathbb{E}\left[X\right]\right| \geq 10\right) \leq \frac{\text{Var}\left(X\right)}{10^2} = \frac{58\frac{1}{3}}{100} = \frac{7}{12}$$

That is, by Chebyshev's inequality, we have a bound:

$$\boxed{\mathbb{P}\left(X \geq 80\right) \leq \frac{7}{12}}$$

Question 10.5:

We wish to order n people in a circle. Each person can request people to not sit next to. Show that if the total number of requests is less than $\binom{n-1}{2}$, then there is a distribution of these requests which allows it to be possible to order the people according to their requests.

Note:

The original question just asks

"Show that if the total number of requests is less than $\binom{n-1}{2}$, then it is possible to order the people according to their requests."

But this would be incorrect if it meant for any distribution of requests which sums to $\binom{n-1}{2}$. This is because we could distribute the requests such that one person has $n-1$ requests (since $n-1$ is less than $\binom{n-1}{2}$ for large enough ns), and then distribute the rest of the requests. Since a person has requested to not sit next to anyone else, it is impossible to seat the people in a circle.

We will show this combinatorically.
Firstly, if $n < 3$, then $\binom{n-1}{2} = 0$, so there are no requests and it is possible to seat the people in any order.
Otherwise, we can seat the people in a circle, then have them request to not sit next to anyone they're not already sitting next to. Thus everyone has $n-3$ requests (there are $n-2$ people not sitting next to them, and they count as a person, so there are $n-3$ remaining people), and therefore $n \cdot (n-3)$ requests total. And since this ordering satisfies the requests, since the requests are based on this ordering, there exists an ordering with $n \cdot (n-3)$ requests.
Notice that $n \cdot (n-3) \geq \binom{n-1}{2} - 1$:

$$\begin{aligned}\binom{n-1}{2} - 1 = \frac{(n-1)(n-2)}{2} - 1 &\leq n(n-3) \\ \iff \quad \frac{n^2-3n+2-2}{2} &\leq n^2 - 3n \\ \iff \quad \frac{n^2-3n}{2} &\leq n^2 - 3n\end{aligned}$$

Which is true.
And note that for any number less than or equal to $n \cdot (n-3)$, there must be an ordering which allows for this number of requests. This is because we can do the same process as above and just remove a few requests.
So this implies that for any number of requests less than or equal to $\binom{n-1}{2} - 1$, there exists some ordering which allows for this number of requests.
And since $a \leq b - 1 \iff a < b$, this means that for any number of requests strictly less than $\binom{n-1}{2}$, there exists some distribution of requests and an ordering which satisfies the requests, as required.

Question 10.6:

X is a function with the following distribution:

$$\mathbb{P}(X = k) = 6 \cdot \left(\frac{1}{\pi \cdot k}\right)^2$$

(1) Compute $\mathbb{E}[X \mid X > 5]$

(2) Compute $\mathbb{E}[X \mid X < 5]$

(1) We know that:

$$\mathbb{P}(X = x \mid X > 5) = \frac{\mathbb{P}(X = x)}{\mathbb{P}(X > 5)}$$

If $x > 5$ and 0 otherwise. And so:

$$\mathbb{E}[X \mid X > 5] = \frac{1}{\mathbb{P}(X > 5)} \cdot \sum_{x=6}^{\infty} x\mathbb{P}(X = x)$$

If we let $\alpha := \mathbb{P}(X > 5)$, we know $\alpha > 0$ since the probability $X = 6$ for example is nonzero. So:

$$\mathbb{E}[X \mid X > 5] = \frac{1}{\alpha} \cdot \left(\mathbb{E}[X] - \sum_{x=1}^{5} x\mathbb{P}(X = x)\right)$$

But:

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} x \cdot 6 \cdot \left(\frac{1}{\pi x}\right)^2 = \frac{6}{\pi^2} \cdot \sum_{x=1}^{\infty} \frac{1}{x}$$

Which diverges to infinity (since $\sum \frac{1}{x}$ is the harmonic series which diverges).
Therefore so does $\mathbb{E}[X \mid X > 5]$ (ie. it has no expected value).
So $(X \mid X > 5)$ has no expecected value.

(2) We know that:

$$\mathbb{P}(X = x \mid X < 5) = \frac{\mathbb{P}(X = x)}{\mathbb{P}(X < 5)}$$

If $x < 5$ and 0 otherwise.
So:

$$\mathbb{E}[X \mid X < 5] = \frac{1}{\mathbb{P}(X < 5)} \cdot \sum_{x=1}^{4} x \cdot \mathbb{P}(X = x)$$

We know:

$$\mathbb{P}(X < 5) = \frac{6}{\pi^2} \cdot \sum_{x=1}^{4} \frac{1}{x^2} = \frac{205}{24\pi^2}$$

And:

$$\sum_{x=1}^{4} x \cdot \mathbb{P}(X = x) = \frac{6}{\pi^2} \cdot \sum_{x=1}^{4} \frac{1}{x} = \frac{25}{2\pi^2}$$

So:

$$\boxed{\mathbb{E}[X \mid X < 5] = \frac{60}{41}}$$

Question 10.7:

A person randomly and uniformly chooses a number between 1 and 20 then rolls that many dice. What is the expected value and variance of the sum of these dice?

Let X be the number chosen, so $X \sim \text{Unif}[20]$, let Y be the sum of the dice, and let Y_i be the result of the ith roll. So:

$$Y = \sum_{i=1}^{X} X_i$$

Since X and the X_is are all independent of one another, and $X_i \sim \text{Unif}[6]$, we know that:

$$\mathbb{E}[Y] = \mathbb{E}[X] \cdot \mathbb{E}[X_i]$$

(As proven in a lecture and shown in question 1).
So:

$$\boxed{\mathbb{E}[Y] = 36\frac{3}{4}}$$

We know that:

$$\text{Var}(Y) = \text{Var}(\mathbb{E}[Y \mid X]) + \mathbb{E}[\text{Var}(Y \mid X)]$$

And:

$$\mathbb{E}[Y \mid X] = \sum_{i=1}^{X} \mathbb{E}[X_i] = X \cdot \mathbb{E}[X_i]$$

So:

$$\text{Var}(\mathbb{E}[Y \mid X]) = \text{Var}(X \cdot \mathbb{E}[X_i]) = \mathbb{E}[X_i]^2 \cdot \text{Var}(X)$$

And:

$$\text{Var}(Y \mid X) = \text{Var}\left(\sum_{i=1}^{X} X_i \,\middle|\, X\right)$$

Since X and the X_is are all independent, this is equal to:

$$= \sum_{i=1}^{X} \text{Var}(X_i \mid X) = \sum_{i=1}^{X} \text{Var}(X_i) = X \cdot \text{Var}(X_i)$$

So:

$$\mathbb{E}[\text{Var}(Y \mid X)] = \text{Var}(X_i) \cdot \mathbb{E}[X]$$

So:

$$\text{Var}(Y) = \mathbb{E}[X_i]^2 \cdot \text{Var}(X) + \mathbb{E}[X] \cdot \text{Var}(X_i) = 3.5^2 \cdot \frac{20^2 - 1}{12} + \frac{21}{2} \cdot \frac{35}{12}$$

That is:

$$\boxed{\text{Var}(Y) = 437\frac{15}{16}}$$

Question 10.8:

$\{X_k\}_{k\in\mathbb{N}}$ is a series of independent random variables which distribute geometrically over a parameter $p \neq 0.5$. Compute:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq 2\right)$$

We know by the law of large numbers that for every $\varepsilon > 0$:

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mathbb{E}\left[X\right]\right| \geq \varepsilon\right) = 0$$

Where $X \sim \mathrm{Geo}\left(p\right)$, so $\mathbb{E}\left[X\right] = \frac{1}{p}$, so:

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{p}\right| \geq \varepsilon\right) = 0$$

We know:

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{p}\right| \geq \varepsilon\right) = \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq \frac{1}{p} + \varepsilon\right) + \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \leq \frac{1}{p} - \varepsilon\right)$$

And since probabilities are nonnegative:

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq \frac{1}{p} + \varepsilon\right), \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \leq \frac{1}{p} - \varepsilon\right) \leq \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{p}\right| \geq \varepsilon\right)$$

So by the squeeze theorem:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq \frac{1}{p} + \varepsilon\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \leq \frac{1}{p} - \varepsilon\right) = 0$$

If $p < 0.5$ then $\frac{1}{p} > 2$, so let $\varepsilon = \frac{1}{p} - 2 > 0$, which means that $\frac{1}{p} - \varepsilon = 2$. So:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \leq \frac{1}{p} - \varepsilon\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \leq 2\right) = 0$$

Which means that:

$$\lim_{n\to\infty} 1 - \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \leq 2\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k > 2\right) = 1$$

And since:

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq 2\right) \geq \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k > 2\right)$$

This means that if $p < 0.5$:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq 2\right) = 1$$

And if $p > 0.5$, then $\frac{1}{p} < 2$, so let $\varepsilon = 2 - \frac{1}{p}$, which means that $\frac{1}{p} + \varepsilon = 2$. So:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq \frac{1}{p} + \varepsilon\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq 2\right) = 0$$

So if $p < 0.5$:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geq 2\right) = 0$$

That is, if $p < 0.5$, then the limit is 1, and if $p > 0.5$, the limit is 0.

Question 10.9:

A weighted coin has a probability of flipping heads with probability $\frac{1}{3}$. The coin is flipped n times independently. Let X be the length of the longest sequence of consecutive flips of heads. Let $\varepsilon > 0$, prove:

$$\mathbb{P}\left(X \geq (1+\varepsilon) \cdot \log_3 n\right) \longrightarrow 0$$

For $1 \leq i \leq n - \log_3 n$, let I_i be an indicator if the next (inclusive) $\log_3 n$ flips resulted in heads. Since each flip is independent, this has a probability of:

$$\frac{1}{3}^{\log_3 n} = \frac{1}{n}$$

So:

$$I_i \sim \text{Ber}\left(\frac{1}{n}\right)$$

Notice that if there's a sequence of at least $(1+\varepsilon)\log_3 n$ consecutive heads, suppose starting at the index i, then $I_j = 1$ for $i \leq j \leq i + \varepsilon \log_3 n$, and therefore:

$$\sum_{i=1}^{n-\log_3 n} I_i \geq \varepsilon \log_3 n$$

So this means that if $X \geq (1+\varepsilon)\log_3 n$ then $\sum I_i \geq \varepsilon \log_3 n$, so:

$$\mathbb{P}\left(X \geq (1+\varepsilon)\log_3 n\right) \leq \mathbb{P}\left(\sum_{i=1}^{n-\log_3 n} I_i \geq \varepsilon \log_3 n\right)$$

By Markov's Inequality:

$$\mathbb{P}\left(\sum_{i=1}^{n-\log_3 n} I_i \geq \varepsilon \log_3 n\right) \leq \frac{\mathbb{E}\left[\sum_{i=1}^{n-\log_3 n} I_i\right]}{\varepsilon \log_3 n} = \frac{\frac{n-\log_3 n}{n}}{\varepsilon \log_3 n}$$

The limit of this as n approaches infinity is 0 since the numerator approaches 1 ($\frac{\log_3 n}{n} \longrightarrow 0$), and the denominator approaches infinity.
So:

$$\mathbb{P}\left(\sum_{i=1}^{n-\log_3 n} I_i \geq \varepsilon \log_3 n\right) \longrightarrow 0$$

By the squeeze theorem. And similarly by the squeeze theorem:

$$\mathbb{P}\left(X \geq (1+\varepsilon)\log_3 n\right) \longrightarrow 0$$

As required.
(This is also true for $\mathbb{P}\left(X > (1+\varepsilon)\log_3 n\right)$ since it is less than or equal to $\mathbb{P}\left(X \geq (1+\varepsilon)\log_3 n\right)$.)