Probability and Statistics Homework #8

Ari Feiglin

Question 8.1:

Given a row of 16 balls, 10 of which are red and 6 are blue, compute the expected number of adjacent color swaps.

Answer:

Let's generalize to n red balls and m blue balls.
Let:

$$X_i := \begin{cases} 1 & \text{color } i \neq \text{color } i+1 \\ 0 & \text{color } i \text{ equals color } i+1 \end{cases}$$

For $i \in [n+m-1]$.
Let X be the number of adjacent color swaps. This means that:

$$X = \sum_{i=1}^{n+m-1} X_i$$

As X_i indicates whether or not there is a color swap between the ith and $i+1$th color.
Thus:

$$\mathbb{E}[X] = \sum_{i=1}^{n+m-1} \mathbb{E}[X_i]$$

By linearity of expected value. Furthermore, since $X_i \in \{0,1\}$, this means:

$$\mathbb{E}[X_i] = \mathbb{P}(X_i = 1)$$

(As $\mathbb{E}[X_i] = 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1)$.)
So what is the probability that X_i is 1? This is the probability that the ith and $i+1$th ball have different colors.
Let C_i denote the color of the ith ball. We know:

$$X_i = 1 \iff C_i \neq C_{i+1}$$

So:

$$\mathbb{P}(X_i = 1) = \mathbb{P}(C_i = \text{red}, C_{i+1} = \text{blue}) + \mathbb{P}(C_i = \text{blue}, C_{i+1} = \text{red})$$

By symmetry, these two probabilities are equal (we can first assign i a color, or $i+1$ and thus the events are equivalent). And we know:

$$\mathbb{P}(C_i = \text{red}, C_{i+1} = \text{blue}) = \frac{n}{n+m} \cdot \frac{m}{n+m-1} = \frac{nm}{(n+m-1)(n+m)}$$

As this is the probability we choose a red ball from the $n+m$ balls (probability $\frac{n}{n+m}$) and then independently a blue ball from the $n+m-1$ remaining balls (probability $\frac{m}{n+m-1}$).
Thus:

$$\mathbb{E}[X_i] = 2 \cdot \frac{n \cdot m}{(n+m-1)(n+m)}$$

Therefore:

$$\mathbb{E}[X] = \sum_{i=1}^{n+m-1} 2\frac{n \cdot m}{(n+m-1)(n+m)} = \frac{2 \cdot n \cdot m}{n+m}$$

In our case, $n = 10$ and $m = 6$, so:

$$\boxed{\mathbb{E}[X] = \frac{15}{2}}$$

■

Question 8.2:

Prove the following: $X \perp\!\!\!\perp Y$ if and only if for all bound functions $g, h\colon \mathbb{R} \longrightarrow \mathbb{R}$:

$$\mathbb{E}\left[g(X) \cdot h(Y)\right] = \mathbb{E}\left[g(X)\right] \cdot \mathbb{E}\left[h(Y)\right]$$

Answer:

Let us first show that if $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ is bound then $\mathbb{E}\left[f(X)\right]$ exists. Suppose $|f(x)| \leq \omega$. All we need to do is to show the following sum converges:

$$\sum_{x \in \mathbb{R}} |f(x)| \cdot \mathbb{P}\left(f(X) = x\right)$$

We know that this is less than or equal to:

$$\leq \sum_{x \in \mathbb{R}} \omega \cdot \mathbb{P}\left(f(X) = x\right) = \omega \cdot \sum_{x \in \mathbb{R}} \mathbb{P}\left(f(X) = x\right)$$

And as always, $\sum\limits_{x \in \mathbb{R}} \mathbb{P}\left(f(X) = x\right) = 1$. So this is equal to ω.

Thus, the sum is positive and bound, which means it converges. Therefore $f(X)$ has an expected value.

$(\implies)$ Suppose $X \perp\!\!\!\perp Y$ and let $g, h\colon \mathbb{R} \longrightarrow \mathbb{R}$ be bound. I will prove that $g(X) \perp\!\!\!\perp h(Y)$.

$$\mathbb{P}\left(g(X) = x, h(Y) = y\right) = \mathbb{P}\left(X \in g^{-1}(x), Y \in h^{-1})(y)\right)$$

Since X and Y are independent, this is equal to:

$$= \mathbb{P}\left(X \in g^{-1}(x)\right) \cdot \mathbb{P}\left(Y \in h^{-1}(y)\right) = \mathbb{P}\left(g(X) = x\right) \cdot \mathbb{P}\left(h(Y) = y\right)$$

Therefore $g(X) \perp\!\!\!\perp h(Y)$. And we know that the expected value of the product of two independent random variables is the product of their expected values, that is:

$$\mathbb{E}\left[g(X) \cdot h(Y)\right] = \mathbb{E}\left[g(X)\right] \cdot \mathbb{E}\left[h(Y)\right]$$

$(\impliedby)$ Let S_1 and S_2 be two subsets of $\mathbb{R}$. I will prove that:

$$\mathbb{P}\left(X \in S_1, Y \in S_2\right) = \mathbb{P}\left(X \in S_1\right) \cdot \mathbb{P}\left(Y \in S_2\right)$$

Let:

$$g \coloneqq \mathbb{1}_{S_1} \quad h \coloneqq \mathbb{1}_{S_2}$$

Thus:

$$\mathbb{P}\left(g(X) = 1\right) = \mathbb{P}\left(X \in S_1\right) \implies g(X) \sim \mathrm{Ber}\left(\mathbb{P}\left(X \in S_1\right)\right)$$

Since $g(X) \in \{1, 0\}$. Similarly for h:

$$h(Y) \sim \mathrm{Ber}\left(\mathbb{P}\left(Y \in S_2\right)\right)$$

This means:

$$\mathbb{E}\left[g(X)\right] = \mathbb{P}\left(X \in S_1\right) \quad \mathbb{E}\left[h(Y)\right] = \mathbb{P}\left(Y \in S_2\right)$$

Furthremore, g and h are both bound (by 1), so:

$$\mathbb{E}\left[g(X) \cdot h(Y)\right] = \mathbb{E}\left[g(X)\right] \cdot \mathbb{E}\left[h(Y)\right]$$

But $\mathbb{E}\left[g(X) \cdot h(Y)\right] = \mathbb{P}\left(g(X) = 1, h(Y) = 1\right)$ as $g(X) \cdot h(Y) \in \{0, 1\}$. And this is simply the event $X \in S_1, Y \in S_2$, since g and h are indicators. So:

$$\mathbb{P}\left(X \in S_1, Y \in S_2\right) = \mathbb{P}\left(X \in S_1\right) \cdot \mathbb{P}\left(Y \in S_2\right)$$

For every $S_1, S_2 \subseteq \mathbb{R}$. Therefore $X \perp\!\!\!\perp Y$, as required.

$\blacksquare$

Question 8.3:

We choose two numbers (with repetition) uniformly and independently from $[n]$. Let X and Y be the firsta and second number respectively. Compute the expected value of $M \coloneqq \max\{X, Y\}$.

Answer:

We know:

$$\mathbb{P}(M = m) = \mathbb{P}(X = m, Y < m) + \mathbb{P}(X < m, Y = m) + \mathbb{P}(X = m, Y = m)$$

As these disjoint events make up the event $M = m$. We know:

$$\mathbb{P}(X = m, Y < m) = \frac{1}{n} \cdot \frac{m-1}{n}$$

As X and Y are independent, and $\mathbb{P}(X = m) = \frac{1}{n}$ and $\mathbb{P}(Y < m) = \frac{m-1}{n}$ as their probabilities are uniform. Similarly:

$$\mathbb{P}(X < m, Y = m) = \frac{1}{n} \cdot \frac{m-1}{n}$$

And:

$$\mathbb{P}(X = m, Y = m) = \mathbb{P}(X = m) \cdot \mathbb{P}(Y = m) = \frac{1}{n^2}$$

So:

$$\mathbb{P}(M = m) = \frac{2m-1}{n^2}$$

We also know $M \in [n]$ since $X, Y \in [n]$. So:

$$\mathbb{E}[M] = \sum_{m=1}^{n} m \cdot \frac{2m-1}{n^2} = \frac{1}{n^2} \cdot \sum_{m=1}^{n} 2m^2 - m$$

We know that:

$$\sum_{m=1}^{n} m^2 = \frac{n(n+1)(2n+1)}{6}$$

And

$$\sum_{m=1}^{n} m = \frac{n}{2} \cdot (n+1)$$

So:

$$\sum_{m=1}^{n} 2m^2 - m = \frac{n(n+1)(2n+1)}{3} - \frac{n(n+1)}{2} = \frac{n(n+1)(4n-1)}{6}$$

Which means that:

$$\boxed{\mathbb{E}[M] = \frac{(n+1)(4n-1)}{6n}}$$

■

Question 8.4:

Suppose X is a random variable such that $X \sim \text{Unif}\left([n]\right)$. And $(Y_1, \ldots, Y_n)$ are random variables that are pairwise independent and independent of X. Furthermore, $Y_i \sim \text{Ber}\left(p\right)$ where $p = 0.5$. Compute the following expected values:

$$a_1 \coloneqq \mathbb{E}\left[\sum_{i \leq X} Y_i \;\middle|\; Y_1 = \cdots = Y_X = 1\right] \quad a_2 \coloneqq \mathbb{E}\left[\sum_{i \leq X} Y_i\right]$$

Explain why for large enough ns, a_1 is less than a_2.

Answer:

Lemma 8.4.1:

If $\{A_i\}_{i \in I}$ is countable and partitions Ω, then:

$$\mathbb{E}\left[X\right] = \sum_{i \in I} \mathbb{P}\left(A_i\right) \cdot \mathbb{E}\left[X \mid A_i\right]$$

Proof:

We know:

$$\mathbb{E}\left[X\right] = \sum_{x \in \mathbb{R}} x \cdot \mathbb{P}\left(X = x\right)$$

From the law of total probability, we know this is equal to:

$$= \sum_{x \in \mathbb{R}} \left(x \cdot \sum_{i \in I} \mathbb{P}\left(X = x \mid A_i\right) \cdot \mathbb{P}\left(A_i\right)\right)$$

Since this must be absolutely convergant (as per the definition of expected values), we can reorder the summation like so:

$$= \sum_{i \in I} \left(\mathbb{P}\left(A_i\right) \cdot \sum_{x \in X} x \cdot \mathbb{P}\left(X = x \mid A_i\right)\right) = \sum_{i \in I} \mathbb{P}\left(A_i\right) \cdot \mathbb{E}\left[X = x \mid A_i\right]$$

As required.

$\blacksquare$

Firstly, note that if $Y_1 = \cdots = Y_X = 1$ then:

$$\sum_{i \leq X} Y_i = \sum_{i \leq X} 1 = X$$

So:

$$a_1 = \mathbb{E}\left[X \mid Y_1 = \cdots = Y_X = 1\right]$$

By the definition of expected values:

$$a_1 = \sum_{k=1}^{n} k \cdot \mathbb{P}\left(X = k \mid Y_1 = \cdots = Y_X = 1\right)$$

And by Baye's Law:

$$\mathbb{P}\left(X = k \mid Y_1 = \cdots = Y_X = 1\right) = \mathbb{P}\left(Y_1 = \cdots = Y_X \mid X = k\right) \cdot \frac{\mathbb{P}\left(X = k\right)}{\mathbb{P}\left(Y_1 = \cdots = Y_X = 1\right)}$$

Using the law of total probability:

$$\mathbb{P}\left(Y_1 = \cdots = Y_X = 1\right) = \sum_{k=1}^{n} \mathbb{P}\left(Y_1 = \cdots = Y_X \mid X = k\right) \cdot \mathbb{P}\left(X = k\right) = \frac{1}{n} \cdot \sum_{k=1}^{n} \mathbb{P}\left(Y_1 = \cdots = Y_k = 1\right)$$

And since Y_i are all independent, this is equal to:

$$= \frac{1}{n} \cdot \sum_{k=1}^{n} p^k = \frac{1}{n} \cdot \frac{p(p^n - 1)}{p - 1}$$

So:

$$\mathbb{P}\left(X = k \mid Y_1 = \cdots = Y_X = 1\right) = p^k \cdot \frac{p-1}{p(p^n - 1)} = p^{k-1} \cdot \frac{p-1}{p^n - 1}$$

And therefore:

$$a_1 = \frac{p-1}{p^n - 1} \cdot \sum_{k=1}^{n} k p^{k-1}$$

Notice that the sum:

$$\sum_{k=1}^{n} k p^{k-1}$$

Is the derivative (with respect to p) of the sum:

$$\sum_{k=1}^{n} p^k = \frac{p(p^n - 1)}{p - 1}$$

And is therefore equal to:

$$\frac{((n+1)p^n - 1) \cdot (p-1) - p(p^n - 1)}{(p-1)^2} = \frac{n \cdot p^{n+1} - (n+1) \cdot p^n + 1}{(p-1)^2}$$

So:

$$a_1 = \frac{n \cdot p^{n+1} - (n+1) \cdot p^n + 1}{(p-1)(p^n - 1)} = \frac{np^n \cdot (p-1) - p^n + 1}{(p-1)(p^n - 1)} = \frac{np^n}{p^n - 1} - \frac{1}{p-1}$$

We can simplify this to:

$$\frac{1}{1-p} - \frac{n}{p^{-n} - 1}$$

Since $p = 0.5$ in our case, this equals:

$$\boxed{a_1 = 2 - \frac{n}{2^n - 1}}$$

By lemma 8.4.1, we know:

$$a_2 = \mathbb{E}\left[\sum_{i \le X} Y_i\right] = \sum_{k=1}^{n} \mathbb{P}\left(X = k\right) \cdot \mathbb{E}\left[\sum_{i \le X} Y_i \;\middle|\; X = k\right]$$

The random variable $\left(\sum_{i \le X} Y_i \;\middle|\; X = k\right)$ is equivalent to the random variable $\left(\sum_{i=1}^{k} Y_i \;\middle|\; X = k\right)$, and since Y_i and X are independent, this is equivalent to the random variable $\sum_{i=1}^{k} Y_i$. Since the expected value is linear, its expect value is:

$$\sum_{i=1}^{k} \mathbb{E}\left[Y_i\right]$$

But since $\{Y_i\}$ all distribute the same way (they are equal distributively), they have the same expected value, so this is equal to:

$$k \cdot \mathbb{E}\left[Y_1\right]$$

This means:

$$a_2 = \sum_{k=1}^{n} \mathbb{P}\left(X = k\right) \cdot k\mathbb{E}\left[Y_1\right] = \mathbb{E}\left[Y_1\right] \cdot \sum_{k=1}^{n} k \cdot \mathbb{P}\left(X = k\right) = \mathbb{E}\left[Y_1\right] \cdot \mathbb{E}\left[X\right]$$

Since $Y_1 \sim \mathrm{Ber}\left(\frac{1}{2}\right)$, this is equal to (we calculated the expected value of X above):

$$\frac{1}{2} \cdot \frac{n+1}{2} = \frac{n+1}{4}$$

All in all:

$$\boxed{a_2 = \frac{n+1}{4}}$$

So how come $a_2 > a_1$ for ns sufficiently large? Well, if we know that $Y_1 = \cdots = Y_X$ are all equal to 1, then X is probably not so large, since the larger X is, the less probable this is. So the sum of $Y_1 + \cdots + Y_X$ in this case (a_1) is generally less than the sum if we don't have any information about $Y_1, \ldots, Y_X$.

■

Question 8.5:

n letters are destined for n different recipients. The letters are placed into n preprepared envelopes. Unfortunately, there was a mixup and the letters were put into the envelopes randomly and uniformly.

(1) What is the expected number of letters which arrive at their intended recipient if each envelope can only hold a single letter?

(2) What is the expected number of letters which arrive at their intended recipient if each envelope can hold an arbitrary number of letters?

Answer:

Let X_i indicate whether or not the ith recipient received the correct letter. Let X be the number of recipients which received the correct letter, which means that X is the sum of the X_is:

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

Thus:

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Since X_i is an indicator (as in it has a bernoulli distribution), its expected value is the probability is equals 1 (since $X_i \in \{1, 0\}$ and multiplying by 0 doesn't affect the sum). So:

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{P}(X_i = 1)$$

(1) In this case, the probability that X_i is 1 is equal to $\frac{1}{n}$ since the distribution is uniform.
Another way of thinking about this is to think of the distribution as a permutation of $[n]$. So $X_i = 1$ if and only if $\sigma(i) = i$. There are $(n-1)!$ permutations which satisfy this (as we know what $\sigma(i)$ is and the rest of the $\sigma(j)$s must form a permutation over $[n] \setminus \{i\}$ which has a cardinality of $(n-1)!$), and $n!$ total permutations. Since the probability is uniform, the probability that $X_i = 1$ is thus:

$$\mathbb{P}(X_i = 1) = \frac{(n-1)!}{n!} = \frac{1}{n}$$

(2) In this case, the probability that X_i is 1 is also equal to $\frac{1}{n}$.
The reasoning for this is because you can think of the distribution of the letters as a function $[n] \longrightarrow [n]$. The number of distributions such that $X_i = 1$ is the number of functions such that $f(i) = i$. The number of functions like this is n^{n-1}, as we know what $f(i)$ and the rest of the $f(j)$s form a function $[n] \setminus \{i\} \longrightarrow [n]$, which has cardinality n^{n-1}. Since the probability is uniform, it is equal to:

$$\mathbb{P}(X_i = 1) = \frac{n^{n-1}}{n^n} = \frac{1}{n}$$

So in both cases, $\mathbb{P}(X_i = 1) = \frac{1}{n}$, which means in both cases the expected value is:

$$\sum_{i=1}^{n} \frac{1}{n} = \frac{n}{n} = 1$$

So in both cases:

$$\boxed{\mathbb{E}[X] = 1}$$

■

Question 8.6:

In a small town there are n residents. The probability a resident is friends with another resident is p. Calculate the expected number of friends a resident in the town has.

Answer:

Let X_i indicate if this resident is friends with the ith *other* resident (there are $n-1$ of these since there are $n-1$ other residents).
We know $X_i \sim \text{Ber}\left(p\right)$, so $\mathbb{E}\left[X_i\right] = p$.
And let X be the number of friends the resident has. This is just the sum of X_is (since the sum is the number of X_is which equal 1, the number of friends):

$$X = \sum_{i=1}^{n-1} X_i$$

Thus:

$$\mathbb{E}\left[X\right] = \sum_{i=1}^{n-1} \mathbb{E}\left[X_i\right] = \sum_{i=1}^{n-1} p = p(n-1)$$

So:

$$\boxed{\mathbb{E}\left[X\right] = p(n-1)}$$

■

Question 8.7:

We flip a fair coin n times. Let X be the number of times two consecutive flips resulted in different outcomes. Compute the expected value of X.

Answer:

Suppose $X = k$, how many ways are there to do this? First, we choose between which two flips the result will change. There are $\binom{n-1}{k}$ ways, as this are $n-1$ spots between flips. Then we must choose which side the coin will land on for the first flip, there are 2 options. The rest of the series of flips is uniquely determined by these two variables. We know that the flips after the first until the first swap must have the same result, then from the first to the second swap must have the same result, and so on. So that means the number of ways to flip the coin such that $X = k$ is $\binom{n-1}{k} \cdot 2$.
The probability of each of these flips is $\frac{1}{2^n}$, since there are 2^n total ways to flip the coin (think of it as a function from $[n]$ to $\{\text{heads}, \text{tails}\}$), and the probability is uniform. So that means:

$$\mathbb{P}(X = k) = \binom{n-1}{k} \cdot \frac{1}{2^{n-1}}$$

We also know that at most there can be $n-1$ swaps, since there are only n flips. Therefore:

$$\mathbb{E}[X] = \sum_{k=0}^{n-1} k \cdot \frac{1}{2^{n-1}} \cdot \binom{n-1}{k} = \frac{1}{2^{n-1}} \cdot \sum_{k=1}^{n-1} k \cdot \binom{n-1}{k}$$

We know that:

$$\sum_{k=1}^{n} k \cdot \binom{n}{k} = n \cdot 2^{n-1}$$

Since the sum represents the "length" of $\mathcal{P}([n])$: if we took every set in $\mathcal{P}([n])$ and added their lengths we'd get the sum.
But on the other hand, for each element $k \in [n]$, there are 2^{n-1} sets in $\mathcal{P}([n])$ which contain it (take any set in $\mathcal{P}([n] \setminus \{k\})$, and add k to it. This is a bijecive function, and $|\mathcal{P}([n] \setminus \{k\})| = 2^{n-1}$). Therefore every element in $[n]$ contributes 2^{n-1} to the sum, and as there are n elements, the sum is equal to $n \cdot 2^{n-1}$.
Therefore:

$$\mathbb{E}[X] = \frac{1}{2^{n-1}} \cdot (n-1) \cdot 2^{n-2} = \frac{n-1}{2}$$

So:

$$\boxed{\mathbb{E}[X] = \frac{n-1}{2}}$$

■

Question 8.8:

A fair coin is flipped until it flips two consecutive heads. Compute the expected value of the number of flips until it stops.

Answer:

Let X be the number of flips. Let R_i indicate whether the ith flip landed on heads. By lemma 8.4.1, we can see:

$$\begin{aligned}\mathbb{E}[X] = &\mathbb{P}(R_1 = 0) \cdot \mathbb{E}[X \mid R_1 = 0] \\ &+ \mathbb{P}(R_1 = 1, R_2 = 1) \cdot \mathbb{E}[X \mid R_1 = 1, R_2 = 1] \\ &+ \mathbb{P}(R_1 = 1, R_2 = 0) \cdot \mathbb{E}[X \mid R_1 = 1, R_2 = 0]\end{aligned}$$

We know that $\mathbb{P}(R_1 = 0) = \frac{1}{2}$ since the coin is fair. And $\mathbb{P}(R_1 = 1, R_2 = 1) = \mathbb{P}(R_1 = 1, R_2 = 0) = \frac{1}{4}$ since the flips are fair, so we just multiply $\mathbb{P}(R_1 = x)$ and $\mathbb{P}(R_2 = y)$ (which are both $\frac{1}{2}$).
If $R_1 = 0$, then we are at the same state that we were at the beginning, we have no heads flipped. So the expected number of how many more flips will need to do until we get two heads is $\mathbb{E}[X]$, but we've already flipped once, so:

$$\mathbb{E}[X \mid R_1 = 0] = \mathbb{E}[X] + 1$$

If $R_1 = 1, R_2 = 1$, then we've already flipped two consecutive heads, and thus $\mathbb{E}[X \mid R_1 = 1, R_2 = 1] = 2$.
If $R_1 = 1, R_2 = 0$, then once again we're at the same state as the beginning, but we've "wasted" two flips, so:

$$\mathbb{E}[X \mid R_1 = 1, R_2 = 0] = \mathbb{E}[X] + 2$$

Therefore:

$$\mathbb{E}[X] = \frac{1}{2} \cdot (\mathbb{E}[X] + 1) + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot (\mathbb{E}[X] + 2)$$

Rearranging, we get:

$$\frac{1}{4} \cdot \mathbb{E}[X] = \frac{3}{2} \implies \boxed{\mathbb{E}[X] = 6}$$

■