Homework_04

February 18, 2025

Probability for Data Science

UC Berkeley, Spring 2025

Michael Xiao and Ani Adhikari

[1]:
```
from prob140 import *
from datascience import *
import numpy as np
from scipy import special
from scipy import stats

import matplotlib.pyplot as plt
%matplotlib inline
import matplotlib
matplotlib.style.use('fivethirtyeight')
```

1 Homework 4 (Due Tuesday, February 18th at 12 PM noon)

1.0.1 Instructions

Your homeworks will generally have two components: a written portion and a portion that also involves code. Written work should be completed on paper, and coding questions should be done in the notebook. Start the work for the written portions of each section on a new page. You are welcome to LaTeX your answers to the written portions, but staff will not be able to assist you with LaTeX related issues.

It is your responsibility to ensure that both components of the lab are submitted completely and properly to Gradescope. **Make sure to assign each page of your pdf to the correct question. Refer to the bottom of the notebook for submission instructions.**

1.0.2 How to Do Your Homework

The point of homework is for you to try your hand at using what you've learned in class. The steps to follow:

- Go to lecture and sections, and also go over the relevant text sections before starting on the homework. This will remind you what was covered in class, and the text will typically contain examples not covered in lecture. The weekly Study Guide will list what you should read.
- Work on some of the practice problems before starting on the homework.
- Attempt the homework problems by yourself with the text, section work, and practice materials all at hand. Sometimes the week's lab will help as well. The two steps above will help this step go faster and be more fruitful.
- At this point, seek help if you need it. Don't ask how to do the problem — ask how to get started, or for a nudge to get you past where you are stuck. Always say what you have already tried. That helps us help you more effectively.
- For a good measure of your understanding, keep track of the fraction of the homework you can do by yourself or with minimal help. It's a better measure than your homework score, and only you can measure it.

1.0.3 Rules for Homework

- Every answer should contain a calculation or reasoning. For example, a calculation such as $(1/3)(0.8)+(2/3)(0.7)$ or `sum([(1/3)*0.8, (2/3)*0.7])`is fine without further explanation or simplification. If we want you to simplify, we'll ask you to. But just $\binom{5}{2}$ by itself is not fine; write "we want any 2 out of the 5 frogs and they can appear in any order" or whatever reasoning you used. Reasoning can be brief and abbreviated, e.g. "product rule" or "not mutually exclusive."
- You may consult others (see "How to Do Your Homework" above) but you must write up your own answers using your own words, notation, and sequence of steps.
- We'll be using Gradescope. You must submit the homework according to the instructions at the end of homework set.

1.1 We will not grade assignments which do not have pages selected for each question.

1.2 1. Aces and Face Cards

A standard deck consists of 52 cards of which 4 are aces, 4 are kings, and 12 (including the four kings) are "face cards" (Jacks, Queens, and Kings).

Cards are dealt at random without replacement from a standard deck till all the cards have been dealt.

Find the expectation of the following. Each can be done with almost no calculation if you use symmetry.

a) The number of face cards among the first 5 cards

b) The number of face cards that *do not* appear among the first 13 cards

c) The number of aces among the first 5 cards minus the number of kings among the last 5 cards

d) The number of cards before the first face card

e) The number of cards strictly in between the first face card and the last face card

f) The number of aces before the first face card

1.3 2. Collecting Distinct Values

This exercise is a workout in finding expectations by using all the tools at your disposal. If an answer doesn't appear to fit into a formula that has already been proven, it's a very good idea to try to write the variable as a sum of simpler variables.

a) A fair die is rolled n times. Find the expected number of times the face with six spots appears.

b) A fair die is rolled n times. Find the expected number of faces that *do not* appear, and say what happens to this expectation as n increases.

c) Use your answer to Part **b** to find the expected number of distinct faces that *do* appear in n rolls of a die.

d) Find the expected number of times you have to roll a die till you have seen all of the faces. This is a version of what is known as the *coupon collector's problem*.

1.4 3. Unbiased Estimators

a) A population of known size N contains an unknown number G of good elements. Let X be the number of good elements in a simple random sample of size n drawn from this population. Use X to construct an unbiased estimator of G.

b) Would your answer to Part **a** have been different if X had been the number of good elements in a random sample drawn with replacement from the population? Why or why not?

c) A flattened die lands 1 and 6 with chance $p/2$ each, and the other faces 2, 3, 4, and 5 with chance $(1-p)/4$ each. Here $p \in (0,1)$ is an unknown number. Let $X_1, X_2, \dots, X_n$ be the results of n rolls of this die. First find $E(|X_1 - 3.5|)$, and use the answer to construct an unbiased estimator of p based on all of $X_1, X_2, \dots, X_n$.

1.5 4. Waiting for a Success Run

Required Reading: Before you start this section, carefully read **Answer 2** of Section 9.3.4. That is the method you will apply in this exercise.

In a sequence of i.i.d. Bernoulli (p) trials, let H represent "heads" or "success", that is, the event that occurs with probability p. Let $W_{H,n}$ be the number of trials till you get n heads in a row.

Remember that "till" means "up to and including" the n heads in a row. Thus if the sequence of tosses starts out as TTTHTHHH then $W_{H,1} = 4$, $W_{H,2} = 7$, and $W_{H,3} = 8$.

(a) [These results are in the book but it will help to have them here for reference. You don't have to prove them here.] Write $E(W_{H,1})$ and $E(W_{H,2})$ as math expressions involving **only** positive terms of the form $1/p^k$ for positive integer k.

(b) Three heads in a row. In Answer 2 of Section 9.3.4, we condition on $W_{H,1}$ to get $E(W_{H,2})$. Now condition on $W_{H,2}$ to find $E(W_{H,3})$. As above, your answer should involve only positive terms of the form $1/p^k$ for positive integer k. Find the numerical value of $E(W_{H,3})$ in the case $p = 1/2$.

(c) n heads in a row. Guess an expression for $E(W_{H,n})$ for $n \geq 1$. Prove it by induction, as follows: Assume that your guess is true for n, and show that then it must also be true for $n+1$.

(d) Define a function `ev_W_run` that takes p and n as arguments and returns $E(W_{H,n})$. Just replace the ellipsis by an expression. Do not add any other lines of code.

[2]:
```
import numpy as np

def ev_W_run(p, n):
    return sum((1 / p) ** k for k in np.arange(1, n+1))
```

Check that your function works by running the cell below and confirming that the answer is the same as what you got in Part **b**.

[3]:
```
ev_W_run(0.5, 3)
```

[3]: 14.0

Now use `ev_W_run` to find the expectation of each of the following random variables.

(1) The number of tosses of a fair coin till 10 consecutive heads appear

(2) The number of rolls of a fair die till six consecutive sixes appear

(3) The number of times a random number generator is run till 000 appears, if the generator draws at random with replacement from the 10 digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9

(4) The number of days till a robot types $ZZZZZ$ if the robot types at the rate of 10 letters per second (without breaks) and chooses letters at random with replacement from the 26 upper case letters of the English alphabet

[4]:
```
# Expected number of:

# (1) fair coin tosses till 10 consecutive heads
ans_1 = ev_W_run(1/2, 10)
```

```python
# (2) rolls of a die till 6 consecutive sixes
ans_2 = ev_W_run(1/6, 6)

# (3) runs of a random number generator till 000
ans_3 = ev_W_run(1/10, 3)

# (4) days till ZZZZZ by robot typist
ans_4 = ev_W_run(1/26, 5)

ans_1, ans_2, ans_3, ans_4
```

[4]: (2046.0, 55986.0, 1110.0, 12356630.0)

1.6 Submission Instructions

Many assignments throughout the course will have a written portion and a code portion. Please follow the directions below to properly submit both portions.

1.6.1 Written Portion

- Scan all the pages into a PDF. You can use any scanner or a phone using applications such as CamScanner. Please **DO NOT** simply take pictures using your phone.
- Please start a new page for each question. If you have already written multiple questions on the same page, you can crop the image in CamScanner or fold your page over (the old-fashioned way). This helps expedite grading.
- It is your responsibility to check that all the work on all the scanned pages is legible.
- If you used LaTeX to do the written portions, you do not need to do any scanning; you can just download the whole notebook as a PDF via LaTeX.

1.6.2 Code Portion

- Save your notebook using `File > Save Notebook`.
- Generate a PDF file using `File > Save and Export Notebook As > PDF`. This might take a few seconds and will automatically download a PDF version of this notebook.
  - If you have issues, please post a follow-up on the general Homework 4 Ed thread.

1.6.3 Submitting

- Combine the PDFs from the written and code portions into one PDF. Here is a useful tool for doing so.
- Submit the assignment to Homework 4 on Gradescope.
- **Make sure to assign each page of your pdf to the correct question.**
- **It is your responsibility to verify that all of your work shows up in your final PDF submission.**

If you are having difficulties scanning, uploading, or submitting your work, please read the Ed Thread on this topic and post a follow-up on the general Homework 4 Ed thread.

1.7 We will not grade assignments which do not have pages selected for each question.

```
[ ]:
```

1a. Symmetry → each card in deck is equally likely to get drawn

$X \sim$ number of face cards among first 5 cards

$\sim \text{Hypergeometric}(52, 12, 5)$

total deck

drawing face card

first 5 cards chosen

$$P(X=k) = \frac{\binom{12}{k}\binom{40}{5-k}}{\binom{52}{5}}$$

$E[X] = n \cdot \frac{G}{N}$ textbook 8.5.2

$= 5 \cdot \frac{12}{52}$

$= \frac{60}{52}$

$= 1.15$

∴ You can expect about 1.15 face cards among the first 5 cards

1b. $X \sim$ # face cards NOT appearing among first 13 cards

$$52 - 13 = 39$$

remaining cards after first 13 drawn, find expected # face cards that will appear

$$\sim \text{Hypergeometric}(52, 39, 12)$$

$$P(X=k) = \frac{\binom{39}{k}\binom{13}{12-k}}{\binom{52}{12}}$$

$$E[X] = n \cdot \frac{G}{N}$$

$$= 12 \cdot \frac{39}{52}$$

$$= \frac{468}{52}$$

$$= 9$$

$\therefore$ You can expected 9 face cards to NOT appear among first 13 cards

1c. $X \sim$ # aces among first 5 cards minus # kings among last 5 cards

$A \sim$ # aces among first 5 cards

$K \sim$ # kings among last 5 cards

$\sim$ Hypergeometric$(52, 4, 5)$

$$P(A=a) = \frac{\binom{4}{a}\binom{48}{5-a}}{\binom{52}{5}}$$

$$P(K=k) = \frac{\binom{4}{k}\binom{48}{5-k}}{\binom{52}{5}}$$

$$E[A] = n \times \frac{G}{N} = 5 \times \frac{4}{52} = \frac{20}{52}$$

$$E[K] = n \times \frac{G}{N} = 5 \times \frac{4}{52} = \frac{20}{52}$$

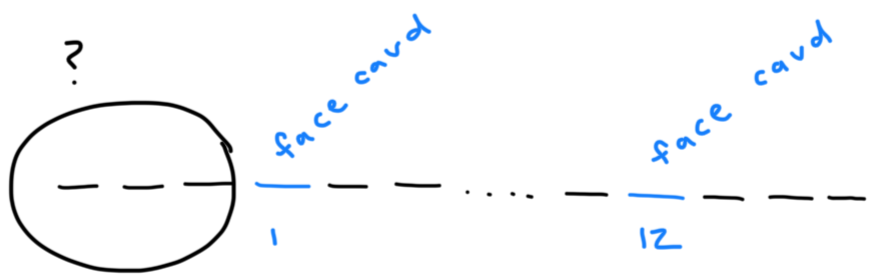
$$E[X] = E[A] - E[K] = \frac{20}{52} - \frac{20}{52} = 0$$

$\therefore$ The expected number of aces among first 5 cards minus the expected number of kings among last 5 cards is 0.

1d. $X \sim$ # cards before first face card

$\therefore$ method of gaps

- sample without replacement
- first gap = first face card
  - ↳ expected # non-face cards before first "gap"



$$E[X] = \frac{N-n}{n+1}$$ textbook 8.4.7

$$= \frac{52-12}{12+1}$$

$$= \frac{40}{13}$$

$$\approx 3.08$$

$\therefore$ The expected number of cards before first face card is 3.08

1e. $X \sim$ # cards strictly in between first face card and last face card

∴ because of symmetry, we expect first face card to appear $\approx \frac{1}{13}$ through the deck + last face card to appear the same as the first face card, but at the end

$X_0 \sim$ # cards before first face card

$$E[X_0] = \frac{40}{13}$$

$$\approx 3.08 \quad \text{(part d)}$$

$X_1 \sim$ position of first face card

$$E[X_1] = E[X_0] + \underline{1}$$

$$= \frac{40}{13} + 1$$

$$= \frac{53}{13}$$

$$\approx 4.08$$

adding 1 let's us take into account the position of first face card, as we are adding it with the expected # cards before it shows

X_{12} ~ position of last face card

$$E[X_{12}] = 52 - E[X_1]$$
$$= 52 - \frac{53}{13}$$
$$= \frac{623}{13}$$
$$\approx 47.92$$

$$\therefore E[X] = E[X_{12}] - E[X_1] \underline{- 1}$$
$$= \frac{623}{13} - \frac{53}{13} - 1$$
$$= \frac{570}{13} - 1$$
$$= \frac{557}{13}$$
$$\approx 42.85$$

subtracting by 1 ensures we find # cards STRICTLY between the first + last face card, excluding both of them

$\therefore$ The expected number of cards strictly in between the first and last face card is 42.85

1f. $X \sim$ # aces before first face card

$\therefore$ indicators

- 4 total aces
- how many will there be?

$$I_j = \begin{cases} 1 & j^{th} \text{ ace appears before first face card} \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = E[I_1] + E[I_2] + E[I_3] + E[I_4]$$

$$X \sim \text{Hypergeometric}(\underbrace{52}_{N}, \underbrace{\tfrac{53}{13}}_{G}, \underbrace{4}_{n})$$

total # cards in deck

4 trials to get all aces

good cards possible to show before first face card

$$E[X] = n \cdot \frac{G}{N} \quad \text{textbook 8.5.2}$$

$$= 4 \times \frac{\frac{53}{13}}{52}$$

$$= \frac{212}{676}$$

$$\approx 0.314$$

$\therefore$ The expected number of aces to appear before first face card is 0.314.

2a. $X \sim$ # faces with six spots appears

$\sim \text{Binomial}(\underline{n}, \underline{\tfrac{1}{6}})$

n trials ↩ ↪ fair dice, iid

$I_i \sim$ trial i is success

$$X = I_1 + I_2 + \dots + I_n$$

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

$$\therefore E[X] = E[I_1] + E[I_2] + \dots + E[I_n]$$

$$= np \longrightarrow \text{textbook } 8.5.1$$

$$= n \cdot \frac{1}{6}$$

$$= \frac{n}{6}$$

26. $X \sim$ # faces that don't appear

$\therefore$ die = six faces

$$P\left(\begin{array}{c}\text{face } i \text{ NOT appearing} \\ \text{after one roll}\end{array}\right) = 1 - P(\text{face } i \text{ appears})$$

$$= 1 - \frac{1}{6}$$

$$= \frac{5}{6}$$

$$P\left(\begin{array}{c}\text{face } i \text{ NOT appearing} \\ \text{after } n \text{ rolls}\end{array}\right) = \left(\frac{5}{6}\right)^n$$

$$I_i = \begin{cases} 1 & \text{face } i \text{ doesn't appear after } n \text{ rolls} \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = E[I_1] + E[I_2] + \dots + E[I_6]$$

$$= \left(\frac{5}{6}\right)^n + \left(\frac{5}{6}\right)^n + \dots + \left(\frac{5}{6}\right)^n$$

$$= 6 \times \left(\frac{5}{6}\right)^n$$

$\therefore$ as n increases, the expected value of faces that don't appear approaches 0, meaning that all the faces on the die will appear at least once

2c. $X \sim$ # distinct faces that do appear in n rolls of a die

$$p\left(\begin{array}{l}\text{distinct faces that} \\ \text{do appear in n rolls}\end{array}\right) = 1 - p\left(\begin{array}{c}\text{face i NOT appearing} \\ \text{after n rolls}\end{array}\right)$$

$$= 1 - \left(\frac{5}{6}\right)^n$$

$$I_1 = \begin{cases} 1 & \text{face i appears in n rolls} \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = E[I_1] + E[I_2] + \dots + E[I_6]$$

$$= \left(1 - \left(\frac{5}{6}\right)^n\right) + \left(1 - \left(\frac{5}{6}\right)^n\right) + \dots + \left(1 - \left(\frac{5}{6}\right)^n\right)$$

$$= 6\left(1 - \left(\frac{5}{6}\right)^n\right)$$

$$= 6 - 6\left(\frac{5}{6}\right)^n$$

$\therefore$ as n increases, we saw that in pt. b the expected value approaches 0, leaving behind $6 - 0 = 6$. The expected value of distinct faces that do appear in n rolls of a die is 6, meaning all faces of a die will appear.

2d. $X \sim$ # rolls till you've seen all faces

$X_1 \sim$ first roll shows 1st new face, ALWAYS

$X_2 \sim$ roll until 2nd new face

$\vdots$

$X_6 \sim$ roll until 6th / final new face

$$X = X_1 + X_2 + \ldots + X_6$$

$$X \sim \text{Geometric}(p)$$

↳ probability is always changing depending on new faces found

$$P(X_1) = P(1^{st} \text{ new face}) = \frac{6}{6}$$

$$P(X_2) = P(2^{nd} \text{ new face}) = 1 - P(1^{st} \text{ face rolled})$$

$$= 1 - \frac{1}{6}$$

$$= \frac{5}{6}$$

$$P(X_3) = P(3^{rd} \text{ new face}) = 1 - P(2 \text{ faces rolled})$$

$$= 1 - \frac{2}{6}$$

$$= \frac{4}{6}$$

$$P(X_4) = P(4^{th} \text{ new face}) = 1 - P(3 \text{ faces rolled})$$

$$= 1 - \frac{3}{6}$$

$$= \frac{3}{6}$$

$$P(X_5) = P(5^{th} \text{ new face}) = 1 - P(4 \text{ faces rolled})$$

$$= 1 - \frac{4}{6}$$

$$= \frac{2}{6}$$

$$P(X_6) = P(\text{final face}) = 1 - P(5 \text{ faces rolled})$$

$$= 1 - \frac{5}{6}$$

$$= \frac{1}{6}$$

$$\therefore E[X] = E[X_1] + E[X_2] + \dots + E[X_6]$$

$$= \sum_{k=1}^{6} \frac{1}{p} \quad \text{textbook 8.2.6}$$

$$= \frac{1}{6/6} + \frac{1}{5/6} + \frac{1}{4/6} + \frac{1}{3/6} + \frac{1}{2/6} + \frac{1}{1/6}$$

$$= \frac{6}{6} + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1}$$

$$= \frac{882}{60}$$

$$\approx 14.7$$

$\therefore$ You will be expected to do 14.7 rolls to see all faces on a die

3a. $X \sim \text{Hypergeometric}(N, G, n)$

$\therefore$ simple random sample (w/o replacement) + fixed n

$$E[X] = n \cdot \frac{G}{N}$$ textbook 8.5.2

$$N \times E[X] = nG$$

$$G = \frac{N \times E[X]}{n}$$

$E[X] = X$, as X is unbiased estimator for $E[X]$

$$\hat{G} = \frac{NX}{n}$$

3b. $X \sim \text{Binomial}\left(n, \frac{G}{N}\right)$

$\therefore$ iid + fixed n

$$E[X] = n \cdot p$$ textbook 8.5.1

$$E[X] = n \times \frac{G}{N}$$

$$N \times E[X] = nG$$

$$G = \frac{N \cdot E[X]}{n}$$

$$\hat{G} = \frac{NX}{n}$$

My answer to part a wouldn't have been different if X had been the number of good elements in a random sample with replacement from the population, as the expectation of E[X] is the same when sampling with replacement (binomial dist.) and without replacement (hypergeometric dist.). Also, sampling with replacement gives the same probability of $\frac{G}{N}$ of drawing a good element, which matches the population proportion of good elements $\frac{G}{N}$ when sampling w/o replacement.

3c. $P(X_1 = 1) = \frac{p}{2}$

$P(X_1 = 6) = \frac{p}{2}$

$P(X_1 = 2) = \frac{1-p}{4}$

$P(X_1 = 3) = \frac{1-p}{4}$

$$P(X_1 = 4) = \frac{1-p}{4}$$

$$P(X_1 = 5) = \frac{1-p}{4}$$

$$E[|X_1 - 3.5|] = |1-3.5|\left(\frac{p}{2}\right) + |6-3.5|\left(\frac{p}{2}\right)$$
$$+ |2-3.5|\left(\frac{1-p}{4}\right) + |3-3.5|\left(\frac{1-p}{4}\right)$$
$$+ |4-3.5|\left(\frac{1-p}{4}\right) + |5-3.5|\left(\frac{1-p}{4}\right)$$
$$= \frac{2.5}{2}p + \frac{2.5}{2}p + \frac{1.5}{4} - \frac{1.5}{4}p + \frac{0.5}{4}$$
$$- \frac{0.5}{4}p + \frac{0.5}{4} - \frac{0.5}{4}p + \frac{1.5}{4} - \frac{1.5}{4}p$$
$$= 2.5p + 1 - p$$
$$= 1.5p + 1$$

$$\therefore E[|X_1 - 3.5|] = 1.5p + 1$$

$$E[|X_1 - 3.5|] - 1 = 1.5p$$

$$p = \frac{E[|X_1 - 3.5|] - 1}{1.5}$$

$$E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

textbook 8.4.3 — sample mean

acts as estimator to solve for p, as $E[|x_i - 3.5|]$ is not known

$$E[|x_i - 3.5|] = \frac{1}{n}\sum_{i=1}^{n} |x_i - 3.5|$$

$$\therefore \hat{p} = \frac{\frac{1}{n}\sum_{i=1}^{n} |x_i - 3.5| - 1}{1.5}$$

4a. $x = p(1) + (1-p)(1+x)$

$$x = p + (1-p) + (1-p)x$$

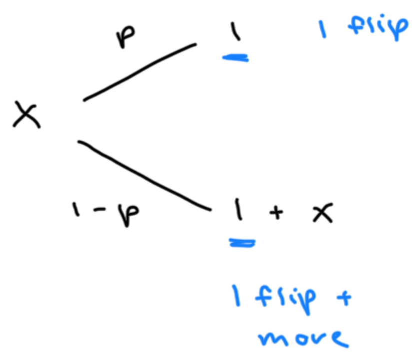
$$x = p + 1 - p + x - px$$

$$px = 1$$

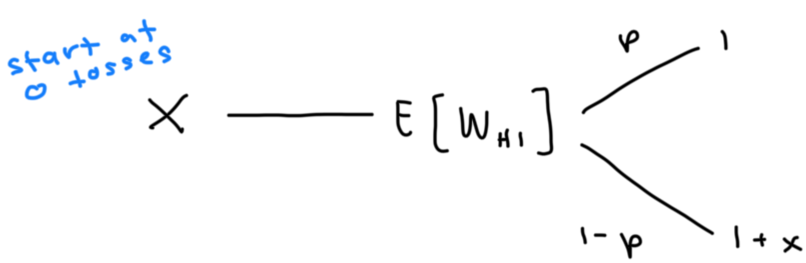
$$x = \frac{1}{p}$$

$$\therefore E[W_H] = \frac{1}{p}$$

$$X = E[W_H]$$



$$X_1 = E[W_{H2}]$$



$$x = \frac{1}{p} + p(1) + (1-p)(1+x)$$

$$x = \frac{1}{p} + p + (1-p) + (1-p)x$$

$$x = \frac{1}{p} + p + 1 - p + x - px$$

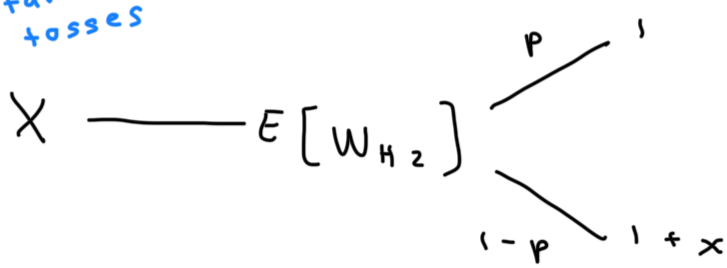
$$px = \frac{1}{p} + 1$$

$$x = \frac{1}{p^2} + \frac{1}{p}$$

$$\therefore E[W_{H2}] = \frac{1}{p^2} + \frac{1}{p}$$

4b. $x = E[W_{H3}]$

start at 0 tosses



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

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

$$x = \frac{1}{p^2} + \frac{1}{p} + p + 1 - p + x - px$$

$$px = \frac{1}{p^2} + \frac{1}{p} + 1$$

$$x = \frac{1}{p^3} + \frac{1}{p^2} + \frac{1}{p}$$

$$\therefore E[W_{H3}] = \frac{1}{p^3} + \frac{1}{p^2} + \frac{1}{p}$$

$$= \frac{1}{(1/2)^3} + \frac{1}{(1/2)^2} + \frac{1}{(1/2)}$$

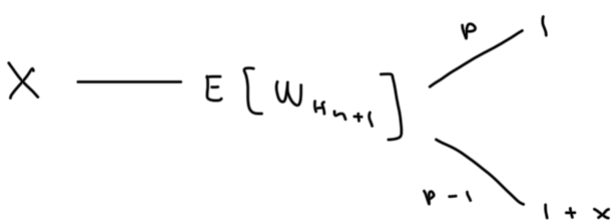
$$= 8 + 4 + 2$$

$$= 14$$

4c. $X = E[W_{Hn}]$

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

$$X = E[W_{H_{n+1}}]$$



$$x = \left(\sum_{i=1}^{n} \frac{1}{p^i} \right) + p + (1-p) + (1-p)x$$

$$x = \left(\sum_{i=1}^{n} \frac{1}{p^i}\right) + p + 1 - p + x - px$$

$$px = \left(\sum_{i=1}^{n} \frac{1}{p^i}\right) + 1$$

$$x = \left(\sum_{i=1}^{n} \frac{1}{p^{i+1}}\right) + \frac{1}{p}$$

$$x = \left(\sum_{i=2}^{n+1} \frac{1}{p^i}\right) + \frac{1}{p}$$

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

4d.

1. fair coin $(p) : \frac{1}{2}$
(heads, tails)

10 consecutive heads $(n) = 10$

2. fair die $(p) : \frac{1}{6}$
(spots: 1, 2, 3, 4, 5, 6)

6 consecutive sixes $(n) = 6$

3. random # generator
WITH replacement (p) $= \frac{1}{10}$
(0, 1, 2, 3, 4, 5, 6, 7, 8, 9)

consecutive 000 (n) $= 3$

4. choose random letter
WITH replacement (p) $= \frac{1}{26}$
(A, B, C, ..., X, Y, Z)

consecutive ZZZZZ (n) $= 5$