Recitation 11: Expectation, Entropy

Problem 1 Coin Expectation

For a sequence of 10 tosses of a coin biased (heads-80%, tail-20%) .

1. What is the size of the sample space?
2. What is the probability that we obtain 7 heads?
3. What is the expected value for the number of heads?
4. What is the variance for the number of heads ?

Problem 2 Die game for grade

You can play the following game 100 times:You roll a dice. If the result is x and x is even, you get x extra points to your grade in CS 1802. If the result is odd you lose 3.5 points.
Should you play this game?

Problem 3 At crossroads ★ A lost tourist arrives at a point with 3 roads. The first road brings him back to the same point after 1 hours of walk. The second road brings him back to the same point after 6 hours of walk. The last road leads to the city after 2 hours of walk. There are no signs on the roads. Assuming that the tourist chooses a road equally likely at all times, what is the average time until the tourist arrives to the city ?

Problem 4 Permutation Expectation

We choose a random permutation $a[]$ of the numbers of indices $1 : n$.

(A) What is the expected value of $X=$ the number of elements such that $a_i = i$ (number of fixed points of the permutation)?

(B)(optional, no credit) ★★ Calculate $\mathbf{var}[X]$

(C) What is the expected number of $Y =$ inversions (pairs $i < j$ with $a_i > a_j$)?

(D) (optional, no credit) ★★★ Calculate $\mathbf{var}[Y]$

Problem 5 Breaking Eggs

Please watch this important video before starting this question:

https://www.youtube.com/watch?v=naQeNNaZYoA

Jimmy and my boy Neil Patrick Harris started out with a carton of 12 eggs. 8 were boiled, 4 were raw. Note that the initial probability of getting a raw egg is 1/3, but that changes once they start cracking them on their faces.

Suppose a single person (could be Jimmy or NPH) cracks 3 eggs. What is the expected number of raw eggs that get smashed on that person's face?

(A) Calculate using E[] definition

(B) Calculate using E[] linearity over sums and indicator RV

Problem 6 Entropy and Codes

You are given the following 2 codes for a, b, c, d, e, f, g, h:

Letter	Code 1
a	000
b	001
c	010
d	011
e	100
f	101
g	110
h	111

Letter	Code 2
a	11
b	10
c	011
d	010
e	0011
f	0010
g	0001
h	0000

1. Encode the following strings into bits using the 2 different codes (sequences of 0's and 1's).

   (a) abcdefgh

   (b) abbbadgh

   Which code is better?

2. Which code is better if you know that the frequency in which the letters appear is given by:

Letter	a	b	c	d	e	f	g	h
Frequency	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

3. Which code is better if you know that the frequency in which the letters appear is given by:

Letter	a	b	c	d	e	f	g	h
Frequency	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{7}{48}$

4. Which code is better if you know that the frequency in which the letters appear is given by:

Letter	a	b	c	d	e	f	g	h
Frequency	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$

5. Compute the entropy function $H(p)$ for all the above distributions. If the probability is not a power of 2 you don't need to evaluate the $\log_2$.

6. Decode the following sequence of bits into letters using code 1 and using code 2 (It is not always the case that a sequence of bits can be decoded by 2 different codes).

111000111010

Problem 7 Encode emojis (optional, no credit)

You work for the cell phone company BestPhone. Your company is offering an amazing new service to its customers, the ability to send text messages with emojis. Your company allows its costumers to send 256 different emojis.

1. BestPhone must pay 1 cent for each bit it sends over the network. It uses a fixed-length code to encode the different emojis. What is the minimum price that BestPhone must charge for sending an emoji in order not to lose money?

2. You did some research and you discovered that $\frac{1}{2}$ of the emojis that are being sent are the smiley emoji, $\frac{1}{256}$ of the emojis are the sad face, and the rest are equidistributed with probability $\frac{1}{512}$.
   You realize that BestPhone can do better.

   - Suggest a more efficient code.
   - What is the minimum price that BestPhone must charge for sending an emoji in order not to lose money? Explain how can this be done.