Practice Problem Set 5: Distributions and Expectation

Reading: LLM Chapter 19 (in the 2017 revision), Schaums 5.1-5.4, 5.7-5.9, 5.11. 6.1-6.2.

Exercise 1. (Problem adapted from an earlier exam.) We have a hash table storing 3500-bit items in 40 bins. To populate this hash table, we have a hash function h that produces random, independent outputs such that

$$h : \{0,1\}^{3500} \rightarrow \{1,2,3,...,40\}$$

To insert item y in the hash table, we simply store it in $h(y)$-th bin.
(Thus, to insert item x with $h(x) = 22$, we store x in the 22^{nd} bin.)
Starting with an empty table, we insert a sequence of items into the table, one after the other.

1. What is the probability that the first item lands in the first bin?
2. Suppose we inserted $t = 80$ items in the hash table. Let X be a random variable that counts the number of items that hashed to the first bin. What kind of random variable is X? (Uniform, Binomial, Bernoulli, or Geometric)?
3. What is the *expected number* of items that are stored in the first bin?
4. Let T be the number of items that we need to insert until *exactly one* item lands in the first bin. What kind of random variable is $T + 1$? (Uniform, Binomial, Bernoulli, or Geometric)?
5. Determine $\text{Ex}[T]$.
6. What is the probability that $T > 200$?
7. Suppose each bin can store a maximum of 4 items. Suppose we want to insert $t = 200$ items into the hash table. What is the probability that *at least one bin* in the hash table will overflow?

Exercise 2. (Problem adapted from an earlier exam) A binary erasure channel is a common communications channel model.

Transmitter – Receiver

In this model, whenever a Transmitter wishes to send a bit b to a Receiver, the bit is *"erased"* before it is arrives at the Receiver with probability p, and received correctly with probability $1 - p$. Each bit is erased independently, and the Receiver can always tell if a bit was erased.

1. Suppose you send an m bit message over the binary erasure channel. What is the probability that none of the bits in the message were erased?
2. Suppose that the Receiver can correctly decode the m-bit message if no more than 2 bits are erased. What is the probability that the Receiver can correctly decode the message?
3. Let X be a random variable that counts number of erasures. What kind of random variable is X? (Uniform, Binomial, Bernoulli, or Geometric)?

4. Suppose $p = \frac{1}{50}$. How small should m be to ensure that the **expected number** of erasures is less than or equal to 2?

Exercise 3. We say probability distribution function (PDF) for random variable X is *symmetric* about a constant a if the following holds for all x:

$$\Pr[X = a - x] = \Pr[X = a + x]$$

1. Prove that if the PDF of X is symmetric about a, then $\text{Ex}[X] = a$.
2. Is the PDF of a uniform distribution U on the interval $[1, 10]$ is symmetric about $\text{Ex}[U]$? (Try plotting the PDF of U)
3. Show that the PDF of the binomial distribution $B(n, p)$ is symmetric about $\text{Ex}[B(n, p)]$ if $p = \frac{1}{2}$.
4. Assume that $p = \frac{1}{3}$. Show that the PDF of the binomial distribution is not symmetric about $\text{Ex}[B(n, p)]$.
5. We roll a fair die with 4-sides (numbered $1, 2, 3, 4$) twice, obtaining values X and Y. Let $Z = \max(X, Y)$. Is the PDF of Z symmetric about $\text{Ex}[Z]$? (Try plotting the PDF of Z.)

Exercise 4. (Problem adapted from an earlier exam) Suppose we have a random variable Y such that

$$Y = \begin{cases} 0 & \text{with probability } \frac{1}{3} \\ 1 & \text{with probability } p \\ 2 & \text{otherwise} \end{cases} \tag{1}$$

Let G be a geometric random variable with parameter $\frac{1}{2}$. (Recall that G counts the number of times you flip a fair coin until you see the first 'heads' (including the final coin flip).) Now let X be random variable such that

$$X = \begin{cases} G & \text{if } Y = 0 \\ 1 & \text{if } Y = 1 \\ G + 1 & \text{if } Y = 2 \end{cases} \tag{2}$$

1. What is $\Pr[X = 1]$?
2. What is the probability that $\Pr[Y = 1 | X = 1]$?

Exercise 5. Here is a useful theorem:

Theorem 1. *For any **independent** random variables X and Y, it follows that* $\text{Ex}[XY] = \text{Ex}[X]\text{Ex}[Y]$.

1. First show that $\text{Ex}[XY] = \text{Ex}[X]\text{Ex}[Y]$ does NOT necessarily hold if X and Y are not independent random variables.
2. Now, prove the theorem.

Exercise 6. True or False? If you think the statement is true, **provide a proof**. If you think the statement is false, provide a **counterexample**.
(**NOTE:** We have not covered variance this week, so skip any parts of this practice problem that mentions variance (*i.e.,* $\text{Var}[\cdot]$) for now, but do go back to it when you are preparing for the final exam!)

1. The expected value of any random variable is non-negative. That is, for any random variable X, $\text{Ex}[X] \geq 0$.

2. The variance of any random variable is non-negative. That is, for any random variable X, $\text{Var}[X] = \text{Ex}[(X - \text{Ex}[X])^2] \geq 0$.

3. For any random variable X, then $\text{Ex}[X^3] = (\text{Ex}[X])^3$.

4. For any random variable X, then $\text{Ex}[3X] = 3\text{Ex}[X]$.

5. For any random variable X with $\text{Ex}[X] = 0$, it follows that $\text{Var}[X] = 0$.

6. For any constant a and b and random variable X, it follows that $\text{Var}[aX + b] = a^2\text{Var}[X]$.

7. For all *independent* random variables X, Y, then

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \tag{3}$$

8. Equation (3) holds even if X and Y are *not* independent.