Section 5

Review

1) Uniform: $X \sim$ Uniform(a, b) (Unif(a, b) for short), for integers $a \leqslant b$, iff X has the following probability mass function:

$$p_X(k) = \frac{1}{b-a+1}, \quad k = a, a+1, \ldots, b$$

$\mathbb{E}[X] = \frac{a+b}{2}$ and $\text{Var}(X) = \frac{(b-a)(b-a+2)}{12}$. This represents each integer from $[a, b]$ being equally likely. For example, a single roll of a fair die is Uniform$(1, 6)$.

2) Bernoulli (or indicator): $X \sim$ Bernoulli(p) (Ber(p) for short) iff X has the following probability mass function:

$$p_X(k) = \begin{cases} p, & k = 1 \\ 1-p, & k = 0 \end{cases}$$

$\mathbb{E}[X] = p$ and $\text{Var}(X) = p(1-p)$. An example of a Bernoulli r.v. is one flip of a coin with $\mathbb{P}(\text{head}) = p$.

3) Binomial: $X \sim$ Binomial(n, p) (Bin(n, p) for short) iff X is the sum of n iid Bernoulli(p) random variables. X has probability mass function

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

$\mathbb{E}[X] = np$ and $\text{Var}(X) = np(1-p)$. An example of a Binomial r.v. is the number of heads in n independent flips of a coin with $\mathbb{P}(\text{head}) = p$. Note that Bin$(1, p) \equiv$ Ber(p). As $n \to \infty$ and $p \to 0,$ with $np = \lambda$, then Bin$(n, p) \to$ Poi(λ). If $X_1, \ldots, X_n$ are independent Binomial r.v.'s, where $X_i \sim$ Bin(N_i, p), then $X = X_1 + \ldots + X_n \sim$ Bin$(N_1 + \ldots + N_n, p)$.

4) Geometric: $X \sim$ Geometric(p) (Geo(p) for short) iff X has the following probability mass function:

$$p_X(k) = (1-p)^{k-1} p, \quad k = 1, 2, \ldots$$

$\mathbb{E}[X] = \frac{1}{p}$ and $\text{Var}(X) = \frac{1-p}{p^2}$. An example of a Geometric r.v. is the number of independent coin flips up to and including the first head, where $\mathbb{P}(\text{head}) = p$.

5) Poisson: $X \sim$ Poisson(λ) (Poi(λ) for short) iff X has the following probability mass function:

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \ldots$$

$\mathbb{E}[X] = \lambda$ and $\text{Var}(X) = \lambda$. An example of a Poisson r.v. is the number of people born during a particular minute, where λ is the average birth rate per minute. If $X_1, \ldots, X_n$ are independent Poisson r.v.'s, where $X_i \sim$ Poi(λ_i), then $X = X_1 + \ldots + X_n \sim$ Poi$(\lambda_1 + \ldots + \lambda_n)$.

6) Hypergeometric: $X \sim$ HyperGeometric(N, K, n) (HypGeo(N, K, n) for short) iff X has the following probability mass function:

$$p_X(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \qquad \text{where } n \leqslant N,\ k \leqslant \min(K, n) \text{ and } k \geqslant \max(0, n-(N-K)).$$

We have $\mathbb{E}[X] = n\frac{K}{N}$. $(\text{Var}(X) = n \cdot \frac{K(N-K)(N-n)}{N^2(2N-1)}$ which is not very memorable.) This represents the number of successes drawn, when n items are drawn from a bag with N items (K of which are successes, and $N-K$ failures) without replacement. If we did this with replacement, then this scenario would be represented as Bin$\left(n, \frac{K}{N}\right)$.

7) Negative Binomial: $X \sim$ NegativeBinomial(r, p) (NegBin(r, p) for short) iff X is the sum of r iid Geometric(p) random variables. X has probability mass function

$$p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \ldots$$

$\mathbb{E}[X] = \frac{r}{p}$ and $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$. An example of a Negative Binomial r.v. is the number of independent coin flips up to and including the r^{th} head, where $\mathbb{P}(\text{head}) = p$. If $X_1, \ldots, X_n$ are independent Negative Binomial r.v.'s, where $X_i \sim$ NegBin(r_i, p), then $X = X_1 + \ldots + X_n \sim$ NegBin$(r_1 + \ldots + r_n, p)$.

Task 1 – Pond fishing

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

a) how many of the next 10 fish I catch are blue, if I catch and release

b) how many fish I had to catch until my first green fish, if I catch and release

c) how many red fish I catch in the next five minutes, if I catch on average r red fish per minute

d) whether or not my next fish is blue

e) how many of the next 10 fish I catch are blue, if I do not release the fish back to the pond after each catch

f) how many fish I have to catch until I catch three red fish, if I catch and release

Task 2 – Best Coach Ever!!

You are a hardworking boxer. Your coach tells you that the probability of your winning a boxing match is 0.2 independently of every other match.

a) How many matches do you expect to fight until you win 10 times and what kind of random variable is this?

b) You only get to play 12 matches every year. To win a spot in the Annual Boxing Championship, a boxer needs to win at least 10 matches in a year. What is the probability that you will go to the Championship this year and what kind of random variable is the number of matches you win out of the 12?

c) Let p be your answer to part (b). How many times can you expect to go to the Championship in your 20 year career?

Task 3 – True or False?

Identify the following statements as true or false (true means always true). Justify your answer.

a) For any random variable X, we have $\mathbb{E}[X^2] \geqslant \mathbb{E}[X]^2$.

b) Let X, Y be random variables. Then, X and Y are independent if and only if $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$.

c) Let $X \sim$ Binomial(n, p) and $Y \sim$ Binomial(m, p) be independent. Then, $X + Y \sim$ Binomial$(n + m, p)$.

d) Let $X_1, ..., X_{n+1}$ be independent Bernoulli(p) random variables. Then, $\mathbb{E}[\sum_{i=1}^{n} X_i X_{i+1}] = np^2$.

e) Let $X_1, ..., X_{n+1}$ be independent Bernoulli(p) random variables. Then, $Y = \sum_{i=1}^{n} X_i X_{i+1} \sim$ Binomial(n, p^2).

f) If $X \sim$ Bernoulli(p), then $nX \sim$ Binomial(n, p).

g) If $X \sim$ Binomial(n, p), then $\frac{X}{n} \sim$ Bernoulli(p).

h) For any two independent random variables X, Y, we have $\mathrm{Var}(X - Y) = \mathrm{Var}(X) - \mathrm{Var}(Y)$.

Task 4 – Memorylessness

We say that a random variable X is memoryless if $\mathbb{P}(X > k + i \mid X > k) = \mathbb{P}(X > i)$ for all non-negative integers k and i. The idea is that X does not *remember* its history. Let $X \sim Geo(p)$. Show that X is memoryless.

Task 5 – Fun with Poissons

Let $X \sim Poisson(\lambda_1)$ and $Y \sim Poisson(\lambda_2)$, where X and Y are independent.

a) Show that $X + Y \sim Poisson(\lambda_1 + \lambda_2)$. To show that a random variable is distributed according to a particular distribution, we must show that they have the same PMF. Thus, we are trying to show that $P(X + Y = n) = e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}$

b) Show that $P(X = k \mid X + Y = n) = P(W = k)$ where $W \sim Bin(n, \frac{\lambda_1}{\lambda_1+\lambda_2})$

Task 6 – Balls and Bins

Throw n balls into m bins, where m and n are positive integers. Let X be the number of bins with exactly one ball. Compute $\mathrm{Var}(X)$.