Section 5 – Solutions

Review

1) Variance. Var $(X) = \mathbb{E}[(X - \mathbb{E}[\mathbb{X}])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ Var $(aX + b) =$ ___Var (X).

Notice that since this is an expectation of a non-negative random variable ($(X-\mu)^2$), variance is always non-negative.

Var $(aX + b) = a^2$Var (X)

2) Independence. Two random variables X and Y are **independent** if ____________________.

When two random variables are independent, we have $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ (the converse is not necessarily true).

$\forall x \in \Omega_X, \forall y \in \Omega_Y$, the following holds true: $\mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$

3) Variance and Independence. For any two independent random variables X and Y, Var $(X + Y) =$ ____________________

This depends on independence, whereas linearity of expectation always holds. Note that this combined with the above shows that $\forall a, b, c \in \mathbb{R}$ and if X is independent of Y, $\text{Var}(aX + bY + c) = a^2\text{Var}(X) + b^2\text{Var}(Y)$.

Var $(X + Y) =$ Var $(X) + Var(Y)$

4) i.i.d. (independent and identically distributed): Random variables $X_1, \ldots, X_n$ are i.i.d. (or iid) iff they are independent and have the same probability mass function.

5) 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$(1, 6)$.

6) 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 $\mathbb{P}(\text{head}) = p$.

7) 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$(N_1 + \ldots + N_n, p)$.

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

We may or may not cover the next three in class.

9) 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 $\operatorname{Var}(X) = \lambda$. 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$(\lambda_1 + \ldots + \lambda_n)$.

10) 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 $\mathbb{E}[X] = n\frac{K}{N}$. ($\operatorname{Var}(X) = n \cdot \frac{K(N-K)(N-n)}{N^2(2N-1)}$ 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 $\left(n, \frac{K}{N}\right)$.

11) 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 $\operatorname{Var}(X) = \frac{r(1-p)}{p^2}$. 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

Since this is the same as saying how many of my next 10 trials (fish) are a success (are blue), this is a binomial distribution. Specifically, since we are doing catch and release, the probability of a given fish being blue is $\frac{B}{N}$ and each trial is independent. Thus:

$$\text{Bin}\left(10, \frac{B}{N}\right)$$

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

Once again, each catch is independent, so this is asking how many trials until we see a success, hence it is a geometric distribution:

$$\text{Geo}\left(\frac{G}{N}\right)$$

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

This is asking for the number of occurrences of event given an average rate, which is the definition of the Poisson distribution. Since we're looking for events in the next 5 minutes, that is our time unit, so we have to adjust the average rate to match (r per minute becomes $5r$ per 5 minutes).

$$\text{Poi}(5r)$$

d) whether or not my next fish is blue

This is the same as the binomial case, but it's only one trial, so it is necessarily Bernoulli.

$$\text{Ber}\left(\frac{B}{N}\right)$$

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

We have not covered the Hypergeometric RV in class, but its definition is the number of successes in n draws (without replacement) from N items that contain K successes in total. In this case, we have 10 draws (without replacement because we do not catch and release), and out of the N fish, B are blue (a success).

$$\text{HypGeo}(N, B, 10)$$

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

Negative binomial is another RV we didn't cover in class. It models the number of trials with probability of success p, until you get r successes. In this case, as before, our trials are caught fish (with replacement this time) and our success is if the fish are red, which happens with probability $\frac{R}{N}$.

$$\text{NegBin}\left(3, \frac{R}{N}\right)$$

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?

The number of matches you have to fight until you win 10 times can be modeled by $\sum_{i=1}^{10} X_i$ where $X_i \sim \text{Geometric}(0.2)$ is the number of matches you have to fight to go from $i-1$ wins to i wins, including the match that gets you your i^{th} win, where every match has a 0.2 probability of success. Recall $\mathbb{E}[X_i] = \frac{1}{0.2} = 5$. $\mathbb{E}[\sum_{i=1}^{10} X_i] = \sum_{i=1}^{10} \mathbb{E}[X_i] = \sum_{i}^{10} \frac{1}{0.2} = 10 \cdot 5 = 50$.

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?

You can go to the championship if you win more than or equal to 10 times this year. Let Y be the number of matches you win out of the 12 matches. Note that $Y \sim \text{Binomial}(12, 0.2)$. Since the max number you can win is 12 (there are 12 matches), we are looking for $P(10 \leqslant Y \leqslant 12)$. Thus, since Y is discrete, we are interested in

$$\mathbb{P}(Y = 10) + \mathbb{P}(Y = 11) + \mathbb{P}(Y = 12) = \sum_{i=10}^{12} \binom{12}{i} 0.2^i (1 - 0.2)^{12-i}$$

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?

The number of times you go to the championship can be modeled by $Y \sim \text{Binomial}(20, p)$. So, $E[Y] = 20 \cdot p$.

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 $\mathbb{E}[X^2] \geqslant \mathbb{E}[X]^2$.

True. $\text{Var}(X)$ is the expectation of a square so $\text{Var}(X) \geqslant 0$. Then we have $\mathbb{E}[X^2] - \mathbb{E}[X]^2 = \text{Var}(X) \geqslant 0$ which is equivalent to what we need to prove.

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

False. The forward implication is true, but the reverse is not. For example, if $X \sim \text{Uniform}(-1, 1)$ (equally likely to be in $\{-1, 0, 1\}$), and $Y = X^2$, we have $\mathbb{E}[X] = 0$, so $\mathbb{E}[X]\,\mathbb{E}[Y] = 0$. However, since $X = X^3$ (why? X takes on only 3 values $-1, 0, 1$ which are the 3 solutions of the equation $x^3 - x = 0$), $\mathbb{E}[XY] = \mathbb{E}[XX^2] = \mathbb{E}[X^3] = \mathbb{E}[X] = 0$, we have that $\mathbb{E}[X]\,\mathbb{E}[Y] = 0 = \mathbb{E}[XY]$. However, X and Y are not independent; indeed, $\mathbb{P}(Y = 0|X = 0) = 1 \neq \frac{1}{3} = \mathbb{P}(Y = 0)$.

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

True. X is the sum of n independent Bernoulli trials, and Y is the sum of m. So $X + Y$ is the sum of $n + m$ independent Bernoulli trials, so $X + Y \sim \text{Binomial}(n + m, p)$.

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

True. Notice that $X_i X_{i+1}$ is also Bernoulli (only takes on 0 and 1), but is 1 iff both are 1, so $X_i X_{i+1} \sim \text{Bernoulli}(p^2)$. The statement holds by linearity, since $\mathbb{E}[X_i X_{i+1}] = p^2$.

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

False. They are all Bernoulli p^2 as determined in the previous part, but they are not independent. Indeed, $\mathbb{P}(X_1 X_2 = 1|X_2 X_3 = 1) = \mathbb{P}(X_1 = 1) = p \neq p^2 = \mathbb{P}(X_1 X_2 = 1)$.

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

False. The range of X is $\{0, 1\}$, so the range of nX is $\{0, n\}$. nX cannot be Bin(n, p), otherwise its range would be $\{0, 1, ..., n\}$.

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

False. Again, the range of X is $\{0, 1, ..., n\}$, so the range of $\frac{X}{n}$ is $\{0, \frac{1}{n}, \frac{2}{n}, ..., 1\}$. Hence it cannot be Ber$(p)$, otherwise its range would be $\{0, 1\}$.

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

False. $\text{Var}(X - Y) = \text{Var}(X + (-Y)) = \text{Var}(X) + (-1)^2\text{Var}(Y) = \text{Var}(X) + \text{Var}(Y)$.

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.

Let's note that if $X \sim Geo(p)$, then $\mathbb{P}(X > k) = \mathbb{P}(\text{no successes in the first } k \text{ trials}) = (1-p)^k$.

$$\begin{aligned}
\mathbb{P}(X > k + i \mid X > k) &= \frac{\mathbb{P}(X > k \mid X > k + i)\,\mathbb{P}(X > k + i)}{\mathbb{P}(X > k)} && \text{[Bayes Theorem]} \\
&= \frac{\mathbb{P}(X > k + i)}{\mathbb{P}(X > k)} && [\mathbb{P}(X > k \mid X > k + i) = 1] \\
&= \frac{(1-p)^{k+i}}{(1-p)^k} && [\mathbb{P}(X > k) = (1-p)^k] \\
&= (1-p)^i \\
&= \mathbb{P}(X > i)
\end{aligned}$$

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 $P(X + Y = n) = e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}$

$$\begin{aligned}
P(X + Y = n) &= \sum_{k=0}^{n} P(X = k \cap Y = n - k) \\
&= \sum_{k=0}^{n} P(X = k)P(Y = n - k) && \text{[X and Y are independent]} \\
&= \sum_{k=0}^{n} e^{-\lambda_1}\frac{\lambda_1^k}{k!}\, e^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{\lambda_1^k}{k!}\,\frac{\lambda_2^{n-k}}{(n-k)!} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{1}{k!(n-k)!}\,\lambda_1^k\lambda_2^{n-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\,\lambda_1^k\lambda_2^{n-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^{n} \binom{n}{k}\,\lambda_1^k\lambda_2^{n-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!}\,(\lambda_1 + \lambda_2)^n && \text{[Binomial Theorem]}
\end{aligned}$$

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

$$\begin{aligned}
P(X = k \mid X + Y = n) &= \frac{P(X = k \cap X + Y = n)}{P(X + Y = n)} \\
&= \frac{P(X = k \cap Y = n - k)}{P(X + Y = n)} \\
&= \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)} && \text{[X and Y are independent]} \\
&= \frac{e^{-\lambda_1} \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!}}{e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^n}{n!}} \\
&= \frac{\frac{\lambda_1^k}{k!} \cdot \frac{\lambda_2^{n-k}}{(n-k)!}}{\frac{(\lambda_1+\lambda_2)^n}{n!}} \\
&= \frac{n!}{k!(n-k)!} \cdot \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n} \\
&= \binom{n}{k} \frac{\lambda_1^k \, \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^k \, (\lambda_1 + \lambda_2)^{n-k}} \\
&= \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-k} \\
&= P(W = k)
\end{aligned}$$

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 $\mathrm{Var}(X)$.

Let X_i be the indicator that bin i has exactly one ball, for each $i = 1, ..., m$. Since $X = \sum_i X_i$, we can use the computational formula for variance:

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \mathbb{E}\left[\left(\sum_{i=1}^m X_i\right)^2\right] - \left(\mathbb{E}\left[\sum_{i=1}^m X_i\right]\right)^2 \\
&= \mathbb{E}\left[\sum_{i \neq j} X_i X_j + \sum_{i=1}^m X_i^2\right] - \left(\sum_{i=1}^m \mathbb{E}[X_i]\right)^2 && \text{[Expand square of sum]} \\
&= \sum_{i \neq j} \mathbb{E}[X_i X_j] + \sum_{i=1}^m \mathbb{E}[X_i] - \left(\sum_{i=1}^m \mathbb{E}[X_i]\right)^2,
\end{aligned}$$

where the last line followed from linearity of expectation and recognizing that $X_i^2 = X_i$, since it can only take on the values 0 or 1.

One has

$$\begin{aligned}
\mathbb{E}[X_i] &= 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) && \text{[Definition of Expectation]} \\
&= \mathbb{P}(X_i = 1) \\
&= \binom{n}{1} \cdot \Big(\frac{1}{m}\Big)^1 \Big(\frac{m-1}{m}\Big)^{n-1} \\
&= \frac{n}{m}\Big(\frac{m-1}{m}\Big)^{n-1}
\end{aligned}$$

which is putting only one ball out of n balls into ith bin.
For $j \in 1, ..., n, j \neq i$,

$$\mathbb{E}[X_i X_j] = \binom{n}{1}\binom{n-1}{1}\Big(\frac{1}{m}\Big)^1\Big(\frac{1}{m}\Big)^1\Big(\frac{m-2}{m}\Big)^{n-2} = \frac{n(n-1)}{m^2}\Big(\frac{m-2}{m}\Big)^{n-2}$$

which is putting only one ball out of n balls into ith bin and only one ball out of $n-1$ balls into jth bin.
Noting that $\sum_{i \neq j}$ has $m(m-1)$ terms, and the rest of the sums have m terms, we find

$$\mathrm{Var}(X) = m(m-1) \cdot \frac{n(n-1)}{m^2}\Big(\frac{m-2}{m}\Big)^{n-2} + m \cdot \frac{n}{m}\Big(\frac{m-1}{m}\Big)^{n-1} - m^2\Big[\frac{n}{m}\Big(\frac{m-1}{m}\Big)^{n-1}\Big]^2$$