

Section 5: Solutions

Review of Main Concepts

- **Independence:** Random variable X and event E are independent iff

$$\forall x, \quad \mathbb{P}(X = x \cap E) = \mathbb{P}(X = x)\mathbb{P}(E)$$

Random variables X and Y are independent iff

$$\forall x \forall y, \quad \mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

In this case, we have $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ (the converse is not necessarily true).

- **i.i.d. (independent and identically distributed):** Random variables X_1, \dots, X_n are i.i.d. (or iid) iff they are mutually independent and have the same probability mass function.
- **Independence of functions of a r.v.:** If X and Y are independent and $g(\cdot), h(\cdot)$ are functions mapping real numbers to real numbers, then $g(X)$ and $h(Y)$ are independent. (See if you can prove this!)
- **Variance of Independent Variables:** If X is independent of Y , $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. 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 , $\text{Var}(aX + bY + c) = a^2\text{Var}(X) + b^2\text{Var}(Y)$.
- Review: Zoo of Discrete Random Variables

- (a) **Uniform:** $X \sim \text{Uniform}(a, b)$ ($\text{Unif}(a, b)$ for short), for integers $a \leq b$, iff X has the following probability mass function:

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

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

- (b) **Bernoulli (or indicator):** $X \sim \text{Bernoulli}(p)$ ($\text{Ber}(p)$ 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 $\mathbb{P}(\text{head}) = p$.

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

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

$\mathbb{E}[X] = np$ 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 $\text{Bin}(1, p) \equiv \text{Ber}(p)$. As $n \rightarrow \infty$ and $p \rightarrow 0$, with $np = \lambda$, then $\text{Bin}(n, p) \rightarrow \text{Poi}(\lambda)$. If X_1, \dots, X_n are independent Binomial r.v.'s, where $X_i \sim \text{Bin}(N_i, p)$, then $X = X_1 + \dots + X_n \sim \text{Bin}(N_1 + \dots + N_n, p)$.

- (d) **Geometric:** $X \sim \text{Geometric}(p)$ ($\text{Geo}(p)$ for short) iff X has the following probability mass function:

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

$\mathbb{E}[X] = \frac{1}{p}$ 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 $\mathbb{P}(\text{head}) = p$.

(e) **Poisson:** $X \sim \text{Poisson}(\lambda)$ ($\text{Poi}(\lambda)$ for short) iff X has the following probability mass function:

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

$\mathbb{E}[X] = \lambda$ 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, \dots, X_n are independent Poisson r.v.'s, where $X_i \sim \text{Poi}(\lambda_i)$, then $X = X_1 + \dots + X_n \sim \text{Poi}(\lambda_1 + \dots + \lambda_n)$.

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

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

$\mathbb{E}[X] = \frac{r}{p}$ and $\text{Var}(X) = \frac{r(1-p)}{p^2}$. 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, \dots, X_n are independent Negative Binomial r.v.'s, where $X_i \sim \text{NegBin}(r_i, p)$, then $X = X_1 + \dots + X_n \sim \text{NegBin}(r_1 + \dots + r_n, p)$.

(g) **Hypergeometric:** $X \sim \text{HyperGeometric}(N, K, n)$ ($\text{HypGeo}(N, K, n)$ for short) iff X has the following probability mass function:

$$p_X(k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}, \quad k = \max\{0, n+K-N\}, \dots, \min\{K, n\}$$

$\mathbb{E}[X] = n \frac{K}{N}$. 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 $\text{Bin}(n, \frac{K}{N})$.

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 **Solution:**

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 **Solution:**

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 **Solution:**

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 **Solution:**

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) (optional) 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 **Solution:**

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) (optional) how many fish I have to catch until I catch three red fish, if I catch and release **Solution:**

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 $\frac{R}{N}$.

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

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? **Solution:**

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 $\mathbb{E}[X_i] = \frac{1}{0.2} = 5$. $\mathbb{E}[\sum_{i=1}^{10} X_i] = \sum_{i=1}^{10} \mathbb{E}[X_i] = \sum_{i=1}^{10} 5 = 10 \cdot 5 = 50$.

- (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? **Solution:**

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 $P(10 \leq Y \leq 12)$. 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? **Solution:**

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

3. Variance of a Product

Let X, Y, Z be independent random variables with means μ_X, μ_Y, μ_Z and variances $\sigma_X^2, \sigma_Y^2, \sigma_Z^2$, respectively. Find $\text{Var}(XY - Z)$. **Solution:**

First notice that $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \implies \mathbb{E}[X^2] = \text{Var}(X) + \mathbb{E}[X]^2 = \sigma_X^2 + \mu_X^2$, and same for Y .

$$\begin{aligned} \text{Var}(XY) &= \mathbb{E}[X^2 Y^2] - \mathbb{E}[XY]^2 \text{ (by theorem in class)} \\ &= \mathbb{E}[X^2] \mathbb{E}[Y^2] - (\mathbb{E}[X] \mathbb{E}[Y])^2 \text{ (by independence)} \\ &= \mathbb{E}[X^2] \mathbb{E}[Y^2] - \mathbb{E}[X]^2 \mathbb{E}[Y]^2 \\ &= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2 \mu_Y^2 \end{aligned}$$

By independence,

$$\begin{aligned} \text{Var}(XY - Z) &= \text{Var}(XY) + \text{Var}(-Z) = \text{Var}(XY) + \text{Var}(Z) \\ &= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2 \mu_Y^2 + \sigma_Z^2 \end{aligned}$$

4. 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 $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$. **Solution:**

True, since $0 \leq \text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$, since the squaring necessitates the result is non-negative.

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

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?), $\mathbb{E}[XY] = \mathbb{E}[X X^2] = \mathbb{E}[X^3] = \mathbb{E}[X] = 0$, 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, $X + Y \sim \text{Binomial}(n + m, p)$. **Solution:**

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

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

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 $nX \sim \text{Binomial}(n, p)$. **Solution:**

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

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

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

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

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

5. Fun with Poissons

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

- (a) Show that $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$ [This was done in class.] **Solution:**

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 $P(X = k | X + Y = n) = P(W = k)$ where $W \sim \text{Bin}(n, \frac{\lambda_1}{\lambda_1 + \lambda_2})$ **Solution:**

$$\begin{aligned}
 P(X = k | X + Y = n) &= \frac{P(X = k \cap X + Y = n)}{P(X + Y = n)} \\
 &= \frac{P(X = k \cap Y = n - k)}{P(X + Y = n)} \\
 &= \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)} && \text{[X and Y are independent]} \\
 &= \frac{e^{-\lambda_1} \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!}}{e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^n}{n!}} \\
 &= \frac{\lambda_1^k}{k!} \cdot \frac{\lambda_2^{n-k}}{(n-k)!} \\
 &= \frac{(\lambda_1 + \lambda_2)^n}{n!} \\
 &= \frac{n!}{k!(n-k)!} \cdot \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n} \\
 &= \binom{n}{k} \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^k (\lambda_1 + \lambda_2)^{n-k}} \\
 &= \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-k} \\
 &= \binom{n}{k} p^k (1-p)^{n-k}, \text{ where } p = \frac{\lambda_1}{\lambda_1 + \lambda_2}
 \end{aligned}$$

6. 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.

Solution:

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

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