

Quiz Section 5

Review

- 1) **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 $[a, b]$ being equally likely. For example, a single roll of a fair die is $\text{Uniform}(1, 6)$.

- 2) **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 $\mathbb{P}(\text{head}) = p$.

- 3) **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 $X = X_1 + \dots + X_n \sim \text{Bin}(N_1 + \dots + N_n, p)$.

- 4) **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 $\mathbb{P}(\text{head}) = p$.

- 5) **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 $X = X_1 + \dots + X_n \sim \text{Poi}(\lambda_1 + \dots + \lambda_n)$.

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

$$p_X(k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}, \quad \text{where } n \leq N, \quad k \leq \min(K, n) \text{ and } k \geq \max(0, n - (N - K)).$$

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 $\text{Bin}(n, \frac{K}{N})$.

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

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

$\mathbb{E}[X] = \frac{r}{p}$ 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 $X = X_1 + \dots + X_n \sim \text{NegBin}(r_1 + \dots + r_n, p)$.

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)):

- how many of the next 10 fish I catch are blue, if I catch and release
- how many fish I had to catch until my first green fish, if I catch and release
- how many red fish I catch in the next five minutes, if I catch on average r red fish per minute
- whether or not my next fish is blue
- 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
- 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.

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

- For any random variable X , we have $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$.
- Let X, Y be random variables. Then, X and Y are independent if and only if $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.
- Let $X \sim \text{Binomial}(n, p)$ and $Y \sim \text{Binomial}(m, p)$ be independent. Then, $X + Y \sim \text{Binomial}(n + m, p)$.
- Let X_1, \dots, X_{n+1} be independent Bernoulli(p) random variables. Then, $\mathbb{E}[\sum_{i=1}^n X_i X_{i+1}] = np^2$.
- Let X_1, \dots, X_{n+1} be independent Bernoulli(p) random variables. Then, $Y = \sum_{i=1}^n X_i X_{i+1} \sim \text{Binomial}(n, p^2)$.
- If $X \sim \text{Bernoulli}(p)$, then $nX \sim \text{Binomial}(n, p)$.
- If $X \sim \text{Binomial}(n, p)$, then $\frac{X}{n} \sim \text{Bernoulli}(p)$.
- For any two independent random variables X, Y , we have $\text{Var}(X - Y) = \text{Var}(X) - \text{Var}(Y)$.

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 $X \sim \text{Geo}(p)$. Show that X is memoryless.

Task 5 – Continuous r.v. example

Suppose that X is a random variable with pdf

$$f_X(x) = \begin{cases} 2C(2x - x^2) & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

where C is an appropriately chosen constant.

- a) What must the constant C be for this to be a valid pdf?
- b) Using this C , what is $\mathbb{P}(X > 1)$?