7 Distributions

7.1 Binomial Distribution

When to use: If there is n trials with probability of success p and x successes. The x successes can occur at any trial (This is why we have to multiply by ${}_n\mathrm{C}_x$).

Requirements:

- n independent trials: This means the outcome of the previous trial does not affect the next trial
- Each trial has to be a *Bernoulli Trial*: Each trial must only have two outcomes (1 or 0, yes or no, success or failure)
- Probability of success in each trial is constant: p does not change on any trial

Let X be the number of successes over running the experiment n times

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \tag{1}$$

We should note that $x \in \mathbb{Z}^+$ and $x \in \{0, 1, 2, .., n\}$

- The mean of a binomial random variable X is $\mu = np$
- The variance of a binomial random variable X is $\sigma^2 = np(1-p)$
- The standard deviation of a binomial random variable X is $\sigma = \sqrt{np(1-p)}$

7.2 Geometric Distribution

When to use: If we perform X trials with probability of success p. We continue to perform trials until we reach a success p.

Requirements:

- X independent trials
- Bernoulli trials
- Probability of success in each trial is constant

Let X be the number of trials until the first success

$$P(X = x) = (1-p)^{x-1} p \tag{2}$$

We should note that $x \in \mathbb{Z}^+$

- The mean of a geometric random variable is $\mu = 1/p$
- The variance of a geometric random variable is $\sigma^2 = (1-p)/p^2$

7.3 Negative Binomial Distribution

When to use: When you have x trials with probability of success p and r successes. The difference between binomial and negative binomial is that negative binomial distributions stop at the r^{th} success. The wording of the question may use the phrase "Until the rth success", so we know our last trial must be a success.

Requirements:

- x independent trials
- Bernoulli trials
- r successes required to occur

$$P(X = x) = \binom{n-1}{r-1} p^r (1-p)^{n-r} \tag{3}$$

We should note that $x \in \mathbb{Z}^+$ and $x \in \{r, r+1, r+2, ...\}$.

- The mean of a negative geometric random variable is $\mu = r/p$
- The variance of a negative geometric random variable is $\sigma^2 = r(1-p)/p^2$

7.4 Hyper geometric Distribution

When to use: When you have a finite population of N objects with two classes: K objects conform and $K-1$ non-conforming. We want to sample n objects **without replacement**.

Requirements:

- A set of N objects split into two classes of K conforming and $K-1$ non-conforming
- Bernoulli trials
- Sample without replacement

Let X be the number of objects we sample from the working class

$$P(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \tag{4}$$

- The mean of a hyper geometric random variable is $\mu = np = n\dfrac{K}{N}$
- The variance of a hyper geometric random variable is $\sigma^2 = np(1-p)\left(\dfrac{N-n}{N-1}\right)$ where $p = \dfrac{K}{N}$
- $\left(\dfrac{N-n}{N-1}\right)$ is known as the correction factor for sampling without replacement

7.5 Poisson Distribution

When to use: When we know an average per unit and want to create a distribution of probability based on the average.

Requirements:

- λ : average per unit
- T the amount of units we are going to base our distribution

Let X be the number of events in a Poisson process.

$$P(X = x) = \frac{e^{-\lambda T}(\lambda T)^x}{x!} \tag{5}$$