Chapter 4 Outline: Discrete Random Variables

Key Concepts

- A discrete random variable $X$ can take on either a finite or countably infinite number of values.

- The probability mass function specifies the probability that $X$ is equal to $x_i$, for each $x_i$ in the sample space:
  \[ p_i = P(X = x_i) \]
  The probability mass function is the discrete analog of the density function in the continuous case.

- The expectation of a discrete random variable is given by
  \[ E(X) = \mu_X = \sum_{i=1}^{n} p_i x_i \]

- The variance of a discrete random variable is given by
  \[ Var(X) = \sigma^2_X = \sum_{i=1}^{n} (x_i - \mu_X)^2 \]

- The standard deviation of $X$ is the square root of the variance.

Useful Distributions:

- The Bernoulli distribution with parameter $p$. Possible outcomes are $x_0 = 0$ and $x_1 = 1$. The probability mass function is
  \[ p_0 = P(X = 0) = (1 - p) \]
  \[ p_1 = P(X = 1) = p \]
  The mean is $p$ and the variance is $p(1 - p)$. Bernoulli random variables are often used to model situations where the outcome of a trial is either a success or a failure, the former with probability $p$ and the latter with probability $1 - p$.

- The binomial distribution with parameters $n$ and $p$. Possible outcomes are any integer between 0 and $n$. The probability mass function is
  \[ p_i = P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i} \]
  The mean is $np$ and the variance is $np(1 - p)$. Binomial random variables are used to model situations in which $n$ Bernoulli trials are performed in succession (where $n$ is fixed and chosen in advance.) The value of $X$ represents the total number of successes.

- The Poisson distribution with parameter $\lambda$. Possible outcomes are any integer greater than 0. The probability mass function is
  \[ p_i = P(X = i) = \frac{e^{-\lambda} \lambda^i}{i} \]
The mean and variance are both equal to $\lambda$. Poisson random variables are often used to model the number of events that occur in a specified time window. The parameter $\lambda$ is adjusted to reflect how many events typically occur ($\lambda$ is small if there are few events, large if there are many events.)

Miscellaneous

- If $n$ is big, calculating the factorials in the expression for the binomial random variable is not viable. In this case, the Poisson random variable can sometimes be used to approximate the Binomial. The approximation works best in the case when $n$ is large and $p$ is small (i.e $np$ is moderate.) In this case, the best approximation uses parameter $\lambda = np$.

- This approximation is useful for computational, not conceptual reasons.