First Name: ______________________ Last Name: ______________________ Student ID: __________

Chapter 4: Geometric and Hypergeometric Distributions

- **Geometric Distribution**

There are some similiarities but also critical differences between **binomial distribution** and **geometric distribution**.

- Similarities:

- Differences:

For example, in a board game, you cannot make your first move until a specific number is rolled of the die. This could take any number of trials before success is achieved.

Probability Function of a Geometric Distribution $\boldsymbol{X \sim Geo(p)}$

$$P(\boldsymbol{X} = \boldsymbol{x}) = P(\boldsymbol{x}) = q^x p$$

where p is the probability of success in each trial and q is the probability of failure.

The **geometric probability distribution** can be created by finding the probabilities for all values of x from 0 to ∞ (This is not practical for *us* to do, and the probability begins to converge to __________as the values of x become greater.)

Ex.1: What is the probability for getting out of jail in MONOPOLY

a) on your first roll? b) on your 4^{th} roll? c) within 3 rolls?

d) Graph the probability distribution for the number of rolls before rolling a double to get out of jail.

The **expected value (expectation)**, *E(X),* for a geometric distribution is the *predicted average of the amount of wait time* (# of failures) before success, and it is the sum of an infinite series. However, it is possible to show using Calculus that the expected value converges to a simple formula:

Expectation for a Geometric Distribution $\boldsymbol{X \sim Geo(p)}$

$$E(\boldsymbol{X}) = \sum_{0}^{\infty} x q^x p = \frac{q}{p}$$

e) Determine the expected number of rolls before rolling a double to get out of jail.

Ex.2: A basketball player has a success rate of 68% for scoring on free throws. What is the expected waiting time before the player misses the basket on a free throw?

Ex.3: Suppose you pass a traffic light on your way to school each day that alternates being green for 40 s and then amber or red for 60 s. What is the expected number of days you reach the intersection before you get a green light?

- **Hypergeometric Distribution**

Unlike binomial and geometric distributions we have encountered, **hypergeometric distributions** have trials whose probabilities ***do*** change from one trial to the next – making them **dependent** on one another.

For example, in one of class 5 lesson examples, we looked at the probability of winning one grand prize when buying 1/10/100 tickets of a lottery that promises 10 grand prizes of its 5 400 000 tickets.

In a hypergeometric distribution, each trial reduces the number of items in the population available for the next selection. The **random variable $\boldsymbol{X}$** for a **hypergeometric distribution** is number of successes (random draws for which the item drawn has a specified feature) in n dependent trials (draws), without replacement, from a *finite population* of size N that contains exactly K items with that feature.

Probability Function of a Hypergeometric Distribution $\boldsymbol{X \sim HyperGeo(N, K, n)}$

$$P(\boldsymbol{X} = \boldsymbol{x}) = P(\boldsymbol{x}) = \frac{C(K, x) \times C(N - K, n - x)}{C(N, n)}$$

where n is the number of dependent trials (draws),
N is the population size (the total number of items available for draws),
and K is the number of "successful" items in the population.

Ex.4: A committee of 8 is to be chosen from 6 men and 7 women.

a) Determine the probability distribution for the number of women on this committee.
b) What is the expected number of women on the committee?

The **expected value**, $E(\boldsymbol{X})$, for a hypergeometric distribution can be calculated by taking the sum of all $x_i * P(x_i)$, but the summation formula can also be simplified into the following:

Expectation for a Hypergeometric Distribution $\boldsymbol{X \sim HyperGeo(N, K, n)}$

$$E(\boldsymbol{X}) = \sum_{x=n-(N-K)}^{K} x \frac{C(K, x) \times C(N-K, n-x)}{C(N, n)} = \frac{nK}{N}$$

Ex.5: A box contains seven yellow, three green, five purple, and six red candies jumbled together. What is the expected number of red candies among five candies poured from the box?

Ex.6: In the spring, the Ministry of the Environment caught and tagged 450 raccoons in a wilderness area. The raccoons were released after being vaccinated against rabies. To estimate the raccoon population in the area, the ministry caught 50 raccoons during the summer. Of these 15 had tags.

a) Determine whether this situation can be modelled with a hypergeometric distribution.
b) Estimate the raccoon population in the wilderness area.

Ex.7: Recent work safety study at a plant has shown that approximately 35% of all industrial accidents are caused by failure of employees to follow instructions. The safety engineer decides to look at the accident reports (selected randomly and replaced in the system after reading) until she finds one that shows an accident caused by failure of employees to follow instructions.

a) On average, how many reports would the safety engineer expect to look at until she finds a report showing an accident caused by employee failure to follow instructions?
b) What is the probability that the safety engineer will have to examine at least three reports until she finds a report showing an accident caused by employee failure to follow instructions?

Ex.8: Seven cards are dealt from a standard deck.

a) What is the probability that three of the seven cards are hearts?
b) What is the expected number of hearts out of the seven cards dealt?

Ex.9: Sixty-five percent of people pass the provincial driver's exam on the first try. A group of 50 individuals who have taken the driver's exam is randomly selected.

a) What is the probability that at most 4 of the 50 selected pass their exams on the first try?
b) How many of the 50 can you expect to have passed the exam on the first try?