Chapter 5
Some Discrete Probability Distributions

5.3 The Binomial Distribution

Objectives:

1. **To define the Bernoulli process and the binomial distribution.**
2. **To find the probabilities of a binomial r.v.**
3. **To calculate μ and σ^2 of a binomial r.v.**

Definition: (Bernoulli Process (Experiment)) It is any experiment that satisfies the following;

1. The experiment consists of **n** repeated trials.
2. Each trial has **only** two outcomes, namely a **S**uccess or a **F**ailure.
3. The probability of the Success, denoted by ***p***, is **fixed**, sampling with replacement, during all the trials.
4. The repeated trials are **independent**.

Definition: (Binomial Distribution) If X: the **number of successes in a Bernoulli process** then X is called a **binomial r.v.** with two parameters ***n***: the number of trials and ***p***: the probability of success, $q = 1 - p$. The probability of getting x successes is given by;

$$f(x) = P(X = x) = b(x; n, p) = \binom{n}{x} p^x q^{n-x} \ ; \ x = 0,1,\ldots,n,$$

which is the pdf of X.

Ex.1 (5.4/119): If $n = 4$ and $p = 3/4$, then find $P(X = 2)$.

Definition: (Binomial Cumulative Distribution Function) If X is a binomial r.v. then the cdf of X is given by; $F(x) = P(X \le x) = \sum_{t=0}^{x} b(t; n, p) = B(x; n, p)$, for which certain numerical values are given in table A.1 on pp. 661-666 for some values of n & p.

Ex.2 (5.5/120): If $n = 15$ and $p = 0.4$, then find;

a. P(At least 10). b. P(Between 3 & 8 inclusive).
c. P(Exactly 5). d. P(At most 7). e. P(Less than 10).

Ex.3 (5.6/120): Let X: # of defective items and let Y: # of shipments with at least one defective items among 20 tested.

a. If $n = 20$ and $p = 0.03$, then find $P(X \geq 1)$.
b. If $n = 10$ and $p = P(X \geq 1)$, then find $P(Y = 3)$.

Theorem: If X is a binomial r.v. with a pdf $b(x;n,p)$ then;

1. $\mu = E(X) = np$ 2. $\sigma^2 = V(X) = npq$.

Ex.4: Refer to Ex.2 and find μ and σ^2.

Ex.5: Problem (12/125). Let W: # of cars out of state, then n=9 & p=1/4.

Ex.6: Problem (15/125). Let Z: # of mice contracting the disease, then n=5 & p = 0.6.

5.4 The Hypergeometric (HG) Distribution

Objectives:

1. **To define the HG distribution and calculate its probabilities.**

2. To recognize the relationship with the binomial r.v.
3. To calculate μ and σ^2 of a HG r.v.

Definition: (Hypergeometric Experiment) It is any experiment done **without replacement** and satisfies the following;

1. A random sample of size n is selected from N items.
2. A group of k items of the N items may be classified as successes and $N - k$ failures.

Definition: (Hypergeometric Distribution) If X: the **number of successes in a sample of size *n* selected from a group of *N* items**, then X is called a **HG r.v.** with three parameters ***N***, ***n***, and ***k***. The probability of getting x successes is given by;

$$P(X = x) = h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \; ; \; \max\{0, n-(N-k)\} \le x \le \min\{k, n\}$$

which is the pdf of X.

Ex.1 (5.12/128): $N = 40$, $k = 3$, and $n = 5$. Find $P(X = 1)$.

Theorem: The **mean** and **variance** of a HG distribution is $\mu = E(X) = \frac{nk}{N}$ and $\sigma^2 = V(X) = \frac{N-n}{N-1} n \frac{k}{N}\left(1 - \frac{k}{N}\right)$ respectively.

Ex.2 (5.13/129): Referring to Ex.1, find μ and σ^2 for the r.v. X.

Relationship To The Binomial Distribution

1. HG probabilities can be approximated by binomial probabilities when $\frac{n}{N} \le 0.05$ for a good approximation.
2. Note that $\mu = E(X) = np = \frac{nk}{N}$ and $\sigma^2 = V(X) = n\frac{k}{N}\left(1 - \frac{k}{N}\right)$ for the binomial distribution with the difference of the **correction factor** $\frac{N-n}{N-1} \to 1$ as n gets sufficiently small relative to N.

Ex.3 (5.14/129): N=5000, k=1000, and n=10. Find $P(X$=3). Note that the exact probability is h(3; 5000, 10, 1000) = 0.2015. What is the expected number of blemished tires in the sample of 10 tires purchased from the shipment?

Ex.4: Problem (1/131). **)**: N=52, k=12, and n=7.

Ex.5: Problem (18/132). **)**: N=9, k=?, n=5, G=3, B=2, and R=4. Find P(Both blue balls & at least 1 red ball).

5.5 The Geometric Distribution

Objectives:

1. To define the geometric distribution.
2. To find the probabilities of a geometric r.v.
3. To calculate μ and σ^2 of a geometric r.v.

Definition: If repeated trials can result in a **S**uccess with probability p and a **F**ailure with probability $q=1-p$, then the probability distribution of the r.v. X: **the number of trials <u>on which the</u> (needed to) first Success occur**, is $g(x;p)=pq^{x-1}, x=1,2,...$

Ex.1 (5.17/134): If $p = 1/100 = 0.01$, then find $P(X = 5)$.
Ex.2 (5.18/135): If $p = 1/20 = 0.05$, then find $P(X = 5)$.

Theorem: The **mean** and **variance** of a geometric distribution is $\mu_X = E(X) = \dfrac{1}{p}$ and $\sigma_X^2 = V(X) = \dfrac{1-p}{p^2} = \dfrac{q}{p^2}$ respectively.

Ex.3: Refer to Ex.1 and find the mean, *the average number of tests needed to get the first defective item*, and the variance.

5.6 Poisson Distribution and Poisson Process

Objectives:

1. **To define the Poisson r.v. and the Poisson process.**
2. **To find the probabilities of a geometric r.v.**
3. **To calculate μ and σ² of a geometric r.v.**
4. **To use the Poisson distribution to approximate the binomial probabilities.**

Definition: If the r.v. X: **the number of outcomes occurring during a Poisson process**, usually denoted by **t**, then X has a Poisson distribution with a pdf; $p(x;\lambda t) = \dfrac{e^{-\lambda t}(\lambda t)^x}{x!}$, $x = 0, 1, 2, ...$
where λ: Average number of outcomes per unit time or region and t: Length of the time interval or the size of the region.

Properties of the Poisson process

1. The number of outcomes in any interval is independent of the other disjoint intervals. Poisson process has no memory.
2. The probability that a single outcome will occur in a very short time interval is proportional to the length of the interval.
3. The probability that more than one outcome will occur in such a very short time interval is negligible.

Note: Certain values of the Poisson probability sum (**Poisson cdf**), i.e. $F(x) = P(X \le x) = \sum_{t=0}^{x} p(t;\lambda t) = P(x;\lambda t)$, are given in table A.2 on pp. 667-669 for some values of $\lambda t = \mu$.

Ex.1 (5.19/137): If Ave. = 4/1millisecond, then find $P(X = 6)$. $\lambda t = 4$.
Ex.2 (5.20/137): If $\lambda t = 10$, then find $P(X > 15)$.

Theorem: Each of the **mean** and **variance** of a Poisson distribution is λt.

Ex.3: Problem (3/139). $\lambda = 5$, $t = 1$.
Ex.4: Problem (10/139). $\lambda = 6$, $t = 1$.

Theorem: Let X:B(n,p), then as $n \to \infty$, $p \to 0$ and $\lambda t = \mu = np$ remains constant, b(x; n, p) → p(x; μ).

Ex.5 (5.21/138): If $p = 0.005$, $n = 400$, then find $P(X = 1)$.