9 Chapter 5 - Several Useful Discrete Distributions

He told his friends,
"You know the law of averages says,
Anything will happen that can. (That's what it sez.)
But the last time the Cubs won a National League pennant
Was the year that we dropped the bomb on Japan. [10]
The Cubs, they made me a criminal,
Sent me down a wayward path,
Stole my youth from me. (... and that's the truth)
I'd forsake my teachers,
To go sit in the bleachers,
In flagrant truancy".

The Dying Cubs Fan's Last Request
Steve Goodman
Big Ear Music, Inc./Red Pajamas Music, Inc.(ASCAP)

1. Previously, we introduce the topic of probability distributions. A number of variables can be represented by established, probability distributions. This chapter introduces three probability distribution functions - binomial, poisson and hypergeometric. We will examine the binomial and poisson. The next chapter introduces the normal and standard normal distributions

2. The probability distributions discussed in this chapter are part of the set of **discrete** probability distributions. These distributions are discrete because they have the following properties.

   (a) The events described by these distributions are discrete (countable).

   (b) $P(x = k) > 0$ for any k in the set of discrete events.

   (c) Shape of the function can best be described as a jump function and best represented graphically using a barchart.

   (d) $P(x \leq k) = \sum_{i=1}^{k} P(x = i)$

   (e) $P(x \geq k) = \sum_{i=k}^{N} P(x = i)$

   (f) $P(a \leq x \leq b) = P(x \leq b) - P(x < a)$

   (g) $\sum_{i=1}^{N} P(x = i) = 1.0$

[10] 1945. The last time the Cubs won the World Series was 1908.

3. One major motivation for understanding these probability distributions is that we can determine the likelihood of real events. If it is understood that a set of events can be represented by a probability distribution, we know have an important tool for determining the likelihood of events. The event of the simple description described above is one case we will examine in this section.

4. **Binomial Distribution**

   (a) Suppose the states of the world consist of two states. Let's refer to one value as a success, the other value as a failure. Events are repeated for n trials. Trials are independent, that is, what happened in the past does not affect the outcome of the present or future events. For each trial, there is a probability of success (p) and failure(q). The questions that a researcher may have is, what is the probability that, after n trials, a person has k successes. If the events can be represented as following a binomial distribution, we can find this probability.

      i. In economics, certain choices can be described this way. A unemployed person in the midst of a job search faces one of two outcome in each time period - success (finds a job) or failure (does not find a job). Suppose a time period consist of a week. A binomial distribution can help us estimate the probability that, after 10 time periods, a person is still unemployed.

   (b) Discrete probability density function (pdf) of the binomial distributions

$$P(x = k) = C_k^n p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

   (c) Discrete cumulative probability density function (cdf) of a binomial distribution

$$P(x \leq j) = \sum_{i=0}^{j} P(x = i)$$

   (d) Mean, Variance and Standard Deviation

$$\mu = np$$

$$\sigma^2 = nqp$$

$$\sigma = \sqrt{nqp}$$

   (e) Rule of thumb - If the sample size (n) is large relative to the population size (N)(if $\frac{n}{N} \geq 0.05$) then the probability assigned to the set of events is not binomial.

5. **Poisson Distribution**

(a) Suppose one knows that, in a given time period, it is expected that a given event will occur μ times (Another way of saying this is that it will occur on the average μ times.) What is the probability that the number of events is k, not μ?

   i. With the previous distribution, the researcher needed to know the probability of success and failure in each time period. The number of trials was also important. Here one only needs to know the average number of times an event is expected to occur. The number of trials is not important.

(b) Discrete probability density function (pdf) for a Poisson distribution.

$$P(x = k) = \frac{\mu^k e^{-u}}{k!}$$

where e is equal to the exponential value.

(c) Discrete cumulative probability density function (cdf) of a binomial distribution

$$P(x \leq j) = \sum_{i=0}^{j} P(x = i)$$

(d) Mean, Variance and Standard Deviation

$$\mu = \mu$$

$$\sigma^2 = \mu$$

$$\sigma = \sqrt{\mu}$$

(e) When the sample size is large and $\mu < 7$, The Poisson probability distribution outcome approximates that Binomial distribution outcome.