Statistics 1
Chapter 7

Discrete Probability Distributions

Chapter Seven

Discrete Probability Distributions

GOALS

When you have completed this chapter, you will be able to:

ONE

Define the terms probability distribution and random variable.

TWO

Distinguish between a discrete and continuous probability distribution.

THREE

Calculate the mean, variance, and standard deviation of a discrete probability distribution.

FOUR

Describe the characteristics and compute probabilities using the binomial probability distribution.

Chapter Six *continued*

A Survey of Probability Concepts

GOALS

When you have completed this chapter, you will be able to:

FIVE

Describe the characteristics and compute probabilities using the hypergeometric distribution.

SIX

Describe the characteristics and compute the probabilities using the Poisson distribution.

Random Variables

- A random variable is a numerical value determined by the outcome of an experiment.
- EXAMPLE 1: Consider a random experiment in which a coin is tossed three times. Let X be the number of heads. Let H represent the outcome of a head and T the outcome of a tail.

EXAMPLE 1 *continued*

- The *sample space* for such an experiment will be: TTT, TTH, THT, THH, HTT, HTH, HHT, HHH.
- Thus the possible values of X (number of heads) are x = 0,1,2,3.

EXAMPLE 1 *continued*

- The outcome of zero heads occurred once.
- The outcome of one head occurred three times.
- The outcome of two heads occurred three times.
- The outcome of three heads occurred once.
- From the definition of a random variable, X as defined in this experiment, is a *random variable*.

Probability Distributions

- A probability distribution is a listing of all the outcomes of an experiment and their associated probabilities. For EXAMPLE 1

Number of Heads	Probability of Outcomes
0	1/8 = .125
1	3/8 = .375
2	3/8 = .375
3	1/8 = .125
Total	8/8 = 1

Characteristics of a Probability Distribution

- The probability of an outcome must always be between 0 and 1.
- The sum of all mutually exclusive outcomes is always 1.

Discrete Random Variable

- A discrete random variable is a variable that can assume only certain clearly separated values resulting from a count of some item of interest.
- EXAMPLE 2: Let X be the number of heads when a coin is tossed 3 times. Here the values for X are x = 0,1,2,3.

Continuous Random Variable

- A continuous random variable is a variable that can assume one of an infinitely large number of values.
- Examples: Height of a basketball player, the length of a nap.

The Mean of a Discrete Probability Distribution

- The mean:
  - reports the central location of the data.
  - is the long-run average value of the random variable.
  - is also referred to as its expected value, E(x), in a probability distribution.
  - is a weighted average.

The Mean of a Discrete Probability Distribution

- The mean is computed by the formula:

$$\mu = E(x) = \Sigma[x * P(x)]$$

- where μ represents the mean and P(x) is the probability of the various outcomes x.

The Variance of a Discrete Probability Distribution

- The variance measures the amount of spread (variation) of a distribution.
- The variance of a discrete distribution is denoted by the Greek letter σ^2 (sigma squared).
- The standard deviation is obtained by taking the square root of σ^2.

The Variance of a Discrete Probability Distribution

- The variance of a discrete probability distribution is computed from the formula

$$\sigma^2 = \Sigma[(x-\mu)^2 * P(x)]$$

EXAMPLE 2

- Dan Desch, owner of College Painters, has studied his records for the past 20 weeks and reports the following number of houses painted per week:

# of Houses Painted	Weeks
10	5
11	6
12	7
13	2

EXAMPLE 2 *continued*

- Probability Distribution:

Number of houses painted, X	Probability, P(X)
10	.25
11	.30
12	.35
13	.10
Total	1

EXAMPLE 2 *continued*

- Compute the mean number of houses painted per week:

$$\mu = E(x) = \Sigma[xP(x)]$$

$$= (10)(.25) + (11)(.30) + (12)(.35) + (13)(.10)$$

$$= 11.3$$

EXAMPLE 2 *continued*

- Compute the variance of the number of houses painted per week:

$$\sigma^2 = \Sigma[(x-\mu)^2 P(x)]$$
$$= .4225 + .0270 + .1715 + .2890$$
$$= .91$$

Binomial Probability Distribution

- The binomial distribution has the following characteristics:
  - An outcome of an experiment is classified into one of two mutually exclusive categories - success or failure.
  - The data collected are the results of counts.
  - The probability of success stays the same for each trial.
  - The trials are independent.

Binomial Probability Distribution

- To construct a binomial distribution, let
  - n be the number of trials
  - x be the number of observed successes
  - π be the probability of success on each trial

Binomial Probability Distribution

- the formula for the binomial probability distribution is:

$$P(x) = \frac{n!}{x!(n-x)!}\pi^{x}(1-\pi)^{n-x}$$

EXAMPLE 3

- The Department of Labor for the state of Alabama reports that 20% of the workforce in Mobile is unemployed. From a sample of 14 workers, calculate the following probabilities using the formula for the binomial probability distribution (n=14, π pp=.2,):
  - three are unemployed: P(x=3)=.250

EXAMPLE 3 *continued*

- *Note:* These are also examples of cumulative probability distributions:
  - three or more are unemployed: $P(x>3)=.250+.172+.086+.032+.009+.002=.551$
  - at least one of the workers is unemployed: $P(x>1) = 1-P(x=0) = 1-.044 = .956$
  - at most two of the workers are unemployed: $P(x<2)=.044+.154+.250=.448$

Mean & Variance of the Binomial Distribution

- The mean is given by:

$$\mu = n\pi$$

- The variance is given by:

$$\sigma^2 = n\pi(1-\pi)$$

EXAMPLE 4

- From EXAMPLE 3, recall that π =.2 and n=14.
- Hence, the mean = $n\pi$ = 14(.2) = 2.8.
- The variance = $n\pi(1-\pi)$ = (14)(.2)(.8) =2.24.

Finite Population

- A finite population is a population consisting of a fixed number of known individuals, objects, or measurements.
- Examples include: The number of students in this class, the number of cars in the parking lot.

Hypergeometric Distribution

- Formula:

$$P(x) = \frac{({}_S C_x)({}_{N-S} C_{n-x})}{{}_N C_n}$$

- where N is the size of the population, S is the number of successes in the population, x is the number of successes of interest, n is the sample size, and C is a combination.

Hypergeometric Distribution

- Use the hypergeometric distribution to find the probability of a specified number of successes or failures if:
  - the sample is selected from a finite population without replacement (recall that a criteria for the binomial distribution is that the probability of success remains the same from trial to trial).
  - the size of the sample n is greater than 5% of the size of the population N .

EXAMPLE 5

- The National Air Safety Board has a list of 10 reported safety violations by ValueJet. Suppose only 4 of the reported violations are actual violations and the Safety Board will only be able to investigate five of the violations. What is the probability that three of five violations randomly selected to be investigated are actually violations?

EXAMPLE 5 *continued*

$$P(3) = \frac{{}_4C_3 * {}_6C_2}{{}_{10}C_5} = \frac{4*15}{252} = .238$$

N	= 10
S	= 4
x	= 3
n	= 5

Poisson Probability Distribution

- The binomial distribution of probabilities will become more and more skewed to the right as the probability of success become smaller.
- The limiting form of the binomial distribution where the probability of success π is very small and n is large is called the Poisson probability distribution.

Poisson Probability Distribution

- The Poisson distribution can be described mathematically using the formula:

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

- where μ is the arithmetic mean number of occurrences in a particular interval of time, e is the constant 2.71828, and x is the number of occurrences.

Poisson Probability Distribution

- The mean number of successes μ can be determined in binomial situations by $n\pi$, where n is the number of trials and π the probability of success.
- The variance of the Poisson distribution is also equal to $n\pi$.

EXAMPLE 6

- The Sylvania Urgent Care facility specializes in caring for minor injuries, colds, and flu. For the evening hours of 6-10 PM the mean number of arrivals is 4.0 per hour.
- What is the probability of 4 arrivals in an hour? P(4) = (4^4)(e^-4)/4!=.1954.