Chapter 4:
Probabilistic features of certain data Distributions
Pages 93- 111

Key words

Probability distribution , random variable , Bernolli distribution, Binomail distribution, Poisson distribution

The Random Variable (X):

- When the values of a variable (height, weight, or age) can't be predicted in advance, the variable is called a random variable.
- An example is the adult height.
- When a child is born, we can't predict exactly his or her height at maturity.

4.2 Probability Distributions for Discrete Random Variables

- **Definition**:
- The probability distribution of a discrete random variable is a table, graph, formula, or other device used to specify all possible values of a discrete random variable along with their respective probabilities.

The Cumulative Probability Distribution of X, F(x):

- It shows the probability that the variable X is less than or equal to a certain value, $P(X \leq x)$.

Example 4.2.1 page 94:

Number of Programs	frequency	P(X=x)	F(x)= P(X≤ x)
1	62	0.2088	0.2088
2	47	0.1582	0.3670
3	39	0.1313	0.4983
4	39	0.1313	0.6296
5	58	0.1953	0.8249
6	37	0.1246	0.9495
7	4	0.0135	0.9630
8	11	0.0370	1.0000
Total	297	1.0000	

- See figure 4.2.1 page 96
- See figure 4.2.2 page 97

- **Properties of probability distribution of discrete random variable.**

1. $0 \le P(X = x) \le 1$
2. $\sum P(X = x) = 1$
3. $P(a \le X \le b) = P(X \le b) - P(X \le a-1)$
4. $P(X < b) = P(X \le b-1)$

- **Example 4.2.2 page 96: (use table in example 4.2.1)**

What is the probability that a randomly selected family will be one who used three assistance programs?

- **Example 4.2.3 page 96: (use table in example 4.2.1)**

What is the probability that a randomly selected family used either one or two programs?

- **Example 4.2.4 page 98: (use table in example 4.2.1)**

  What is the probability that a family picked at random will be one who used two or fewer assistance programs?

- **Example 4.2.5 page 98: (use table in example 4.2.1)**

  What is the probability that a randomly selected family will be one who used fewer than four programs?

- **Example 4.2.6 page 98: (use table in example 4.2.1)**

  What is the probability that a randomly selected family used five or more programs?

Example 4.2.7 page 98: (use table in example 4.2.1)

What is the probability that a randomly selected family is one who used between three and five programs, inclusive?

4.3 The Binomial Distribution:

- The binomial distribution is one of the most widely encountered probability distributions in applied statistics. It is derived from a process known as a Bernoulli trial.
- Bernoulli trial is :

When a random process or experiment called a trial can result in only one of two mutually exclusive outcomes, such as dead or alive, sick or well, the trial is called a Bernoulli trial.

The Bernoulli Process

- A sequence of Bernoulli trials forms a Bernoulli process under the following conditions

1- Each trial results in one of two possible, mutually exclusive, outcomes. One of the possible outcomes is denoted (arbitrarily) as a success, and the other is denoted a failure.

2- The probability of a success, denoted by p, remains constant from trial to trial. The probability of a failure, 1-p, is denoted by q.

3- The trials are independent, that is the outcome of any particular trial is not affected by the outcome of any other trial

- The probability distribution of the binomial random variable **X**, the number of successes in **n** independent trials is:

$$f(x) = P(X = x) = \binom{n}{x} p^X q^{n-X} \quad , \quad x = 0,1,2,....,n$$

- Where $\binom{n}{x}$ is the number of combinations of **n** distinct objects taken **x** of them at a time.

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

$$x! = x(x-1)(x-2)....(1)$$

* Note: 0! =1

Properties of the binomial distribution

1. $f(x) \geq 0$
2. $\sum f(x) = 1$
3. The parameters of the binomial distribution are *n* and *p*
4. $\mu = E(X) = np$
5. $\sigma^2 = \text{var}(X) = np(1-p)$

Example 4.3.1 page 100

- If we examine all birth records from the North Carolina State Center for Health statistics for year 2001, we find that 85.8 percent of the pregnancies had delivery in week 37 or later (full- term birth).

If we randomly selected five birth records from this population what is the probability that exactly three of the records will be for full-term births?

Exercise: example 4.3.2 page 104

Example 4.3.3 page 104

- Suppose it is known that in a certain population 10 percent of the population is color blind. If a random sample of 25 people is drawn from this population, find the probability that

a) Five or fewer will be color blind.

b) Six or more will be color blind

c) Between six and nine inclusive will be color blind.

d) Two, three, or four will be color blind.

- **Exercise: example 4.3.4 page 106**