Chapter 5
Probability Distributions

5-1 Review and Preview

5-2 Probability Distributions

5-3 Binomial Probability Distributions

5-4 Parameters for Binomial Distributions

Review and Preview

This chapter combines the methods of descriptive statistics presented in Chapter 2 and 3 and those of probability presented in Chapter 4 to describe and analyze probability distributions.

Probability distributions describe what will probably happen instead of what actually did happen, and they are often given in the format of a graph, table, or formula.

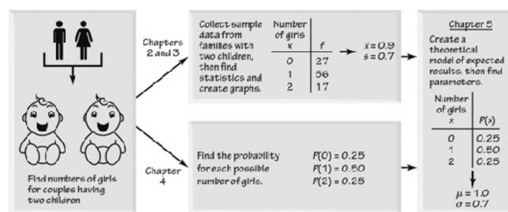
Preview

In order to fully understand probability distributions, we must first understand the concept of a random variable, and be able to distinguish between discrete and continuous random variables. In this chapter we focus on discrete probability distributions. In particular, we discuss binomial and Poisson probability distributions.

11

Combining Descriptive Methods and Probabilities

In this chapter we will construct probability distributions by presenting possible outcomes along with the relative frequencies we expect.



Chapter 5
Probability Distributions

5-1 Review and Preview

5-2 Probability Distributions

5-3 Binomial Probability Distributions

5-4 Parameters for Binomial Distributions

Key Concept

This section introduces the important concept of a probability distribution, which gives the probability for each value of a variable that is determined by chance.

Give consideration to distinguishing between outcomes that are likely to occur by chance and outcomes that are “unusual” in the sense they are not likely to occur by chance.

Random Variable Probability Distribution

- Random Variable
  a variable (typically represented by *x*) that has a single numerical value, determined by chance, for each outcome of a procedure
- Probability Distribution
  a description that gives the probability for each value of the random variable, often expressed in the format of a graph, table, or formula

Discrete and Continuous Random Variables

- Discrete Random Variable
  either a finite number of values or countable number of values, where “countable” refers to the fact that there might be infinitely many values, but that they result from a counting process
- Continuous Random Variable
  has infinitely many values, and those values can be associated with measurements on a continuous scale without gaps or interruptions.

Probability Distribution: Requirements

1. There is a numerical random variable *x* and its values are associated with corresponding probabilities.
2. The sum of all probabilities must be 1.

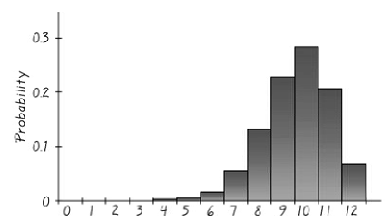
$$\sum P(x) = 1$$

3. Each probability value must be between 0 and 1 inclusive.

$$0 \le P(x) \le 1$$

Graphs

The probability histogram is very similar to a relative frequency histogram, but the vertical scale shows probabilities.



Probability Histogram for Number of Mexican-American Jurors Among 12

Mean, Variance and Standard Deviation of a Probability Distribution

$\mu = \Sigma[x \cdot P(x)]$	Mean
$\sigma^2 = \Sigma[(x-\mu)^2 \cdot P(x)]$	Variance
$\sigma^2 = \Sigma[(x^2 \cdot P(x)] - \mu^2$	Variance (shortcut)
$\sigma = \sqrt{\Sigma[(x^2 \cdot P(x)] - \mu^2}$	Standard Deviation

Expected Value

The expected value of a discrete random variable is denoted by *E*, and it represents the mean value of the outcomes. It is obtained by finding the value of

$$\Sigma[x \cdot P(x)].$$

$$E = \Sigma[x \cdot P(x)]$$

Example

The following table describes the probability distribution for the number of girls in two births.

Find the mean, variance, and standard deviation.

x	$P(x)$		
0	0.25		
1	0.50		
2	0.25		
Total			

Example

The following table describes the probability distribution for the number of girls in two births.

Find the mean, variance, and standard deviation.

x	$P(x)$	$x \cdot P(x)$	$(x-\mu)^2 \cdot P(x)$
0	0.25	0.00	0.25
1	0.50	0.50	0.00
2	0.25	0.50	0.25
Total		1.00	0.50

Example

The following table describes the probability distribution for the number of girls in two births.

Find the mean, variance, and standard deviation.

$$\text{Mean} = \mu = \sum\left[x \cdot P(x)\right] = 1.0$$

$$\text{Variance} = \sigma^2 = \sum\left[(x-\mu)^2 \cdot P(x)\right] = 0.5$$

$$\text{Standard Deviation} = \sigma = \sqrt{0.5} = 0.707$$

Identifying *Unusual* Results Range Rule of Thumb

According to the range rule of thumb, most values should lie within 2 standard deviations of the mean.

We can therefore identify "unusual" values by determining if they lie outside these limits:

Maximum usual value = $\mu + 2\sigma$

Minimum usual value = $\mu - 2\sigma$

Example – continued

We found for families with two children, the mean number of girls is 1.0 and the standard deviation is 0.7 girls.

Use those values to find the maximum and minimum usual values for the number of girls.

$$\text{maximum usual value} = \mu + 2\sigma = 1.0 + 2(0.7) = 2.4$$

$$\text{minimum usual value} = \mu - 2\sigma = 1.0 - 2(0.7) = -0.4$$

Identifying Unusual Results Probabilities

Rare Event Rule for Inferential Statistics

If, under a given assumption (such as the assumption that a coin is fair), the probability of a particular observed event (such as 992 heads in 1000 tosses of a coin) is extremely small, we conclude that the assumption is probably not correct.

Identifying Unusual Results Probabilities

Using Probabilities to Determine When Results Are Unusual

- ❖ Unusually high: *x* successes among *n* trials is an unusually high number of successes if $P(x \text{ or more}) \leq 0.05$.
- ❖ Unusually low: *x* successes among *n* trials is an unusually low number of successes if $P(x \text{ or fewer}) \leq 0.05$.

Chapter 5 Probability Distributions

Key Concept

This section presents a basic definition of a binomial distribution along with notation and methods for finding probability values.

Binomial probability distributions allow us to deal with circumstances in which the outcomes belong to two relevant categories such as acceptable/defective or survived/died.

Binomial Probability Distribution

A binomial probability distribution results from a procedure that meets all the following requirements:

1. The procedure has a fixed number of trials.
2. The trials must be independent. (The outcome of any individual trial doesn't affect the probabilities in the other trials.)
3. Each trial must have all outcomes classified into two categories (commonly referred to as success and failure).
4. The probability of a success remains the same in all trials.

Notation for Binomial Probability Distributions

S and F (success and failure) denote the two possible categories of all outcomes; *p* and *q* will denote the probabilities of S and F, respectively, so

$P(\text{S}) = p$ (*p* = probability of success)

$P(\text{F}) = 1 - p = q$ (*q* = probability of failure)

Notation (continued)

n	denotes the fixed number of trials.
x	denotes a specific number of successes in *n* trials, so *x* can be any whole number between 0 and *n*, inclusive.
p	denotes the probability of success in one of the *n* trials.
q	denotes the probability of failure in one of the *n* trials.
P(*x*)	denotes the probability of getting exactly *x* successes among the *n* trials.

Caution

- Be sure that x and p both refer to the same category being called a success.
- When sampling without replacement, consider events to be independent if $n < 0.05N$.

Example

- When an adult is randomly selected, there is a 0.85 probability that this person knows what Twitter is.
- Suppose we want to find the probability that exactly three of five randomly selected adults know of Twitter.
- Does this procedure result in a binomial distribution?

  Yes. There are five trials which are independent. Each trial has two outcomes and there is a constant probability of 0.85 that an adult knows of Twitter.

Methods for Finding Probabilities

We will now discuss three methods for finding the probabilities corresponding to the random variable x in a binomial distribution.

Method 1: Using the Binomial Probability Formula

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

$$\text{for } x = 0, 1, 2, \ldots, n$$

where

n = number of trials

x = number of successes among n trials

p = probability of success in any one trial

q = probability of failure in any one trial ($q = 1 - p$)

Method 2: Using Technology

STATDISK, Minitab, Excel, SPSS, SAS and the TI-83/84 Plus calculator can be used to find binomial probabilities.

STATDISK

MINITAB

x	P(x)
0	0.000977
1	0.014648
2	0.087891
3	0.263672
4	0.395508
5	0.237305

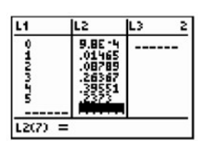
Method 2: Using Technology

STATDISK, Minitab, Excel and the TI-83 Plus calculator can all be used to find binomial probabilities.

EXCEL

	A	B
1	0	0.000977
2	1	0.014648
3	2	0.087891
4	3	0.263672
5	4	0.395508
6	5	0.237305

TI-83 PLUS Calculator



Method 3: Using Table A-1 in Appendix A

Part of Table A-1 is shown below. With n = 12 and p = 0.80 in the binomial distribution, the probabilities of 4, 5, 6, and 7 successes are 0.001, 0.003, 0.016, and 0.053 respectively.

n	x	p 0.80
	4	0.001
	5	0.003
	6	0.016
	7	0.053

→

x	p
4	0.001
5	0.003
6	0.016
7	0.053

Strategy for Finding Binomial Probabilities

- Use computer software or a TI-83/84 Plus calculator, if available.
- If neither software nor the TI-83/84 Plus calculator is available, use Table A-1, if possible.
- If neither software nor the TI-83/84 Plus calculator is available and the probabilities can't be found using Table A-1, use the binomial probability formula.

Example

Given there is a 0.85 probability that any given adult knows of Twitter, use the binomial probability formula to find the probability of getting exactly three adults who know of Twitter when five adults are randomly selected.

We have:

$$n = 5, \quad x = 3, \quad p = 0.85, \quad q = 0.15$$

We want:

$$P(3)$$

Example

We have:

$$n = 5, \quad x = 3, \quad p = 0.85, \quad q = 0.15$$

$$\begin{aligned} P(3) &= \frac{5!}{(5-3)!3!} \cdot 0.85^{3} \cdot 0.15^{5-3} \\ &= \frac{5!}{2!3!} \cdot 0.614125 \cdot 0.0225 \\ &= (10)(0.614125)(0.0225) \\ &= 0.138 \end{aligned}$$

Rationale for the Binomial Probability Formula

$$P(x) = \underbrace{\frac{n!}{(n-x)!x!}}_{\text{The number of outcomes with exactly } x \text{ successes among } n \text{ trials}} \cdot p^{x} \cdot q^{n-x}$$

Rationale for the Binomial Probability Formula

$$P(x) = \underbrace{\frac{n!}{(n-x)!x!}}_{\text{Number of outcomes with exactly } x \text{ successes among } n \text{ trials}} \cdot \underbrace{p^{x} \cdot q^{n-x}}_{\text{The probability of } x \text{ successes among } n \text{ trials for any one particular order}}$$

Chapter 5
Probability Distributions

Key Concept

In this section we consider important characteristics of a binomial distribution including center, variation and distribution. That is, given a particular binomial probability distribution we can find its mean, variance and standard deviation.

A strong emphasis is placed on interpreting and understanding those values.

Binomial Distribution: Formulas

Mean $\mu = n \cdot p$

Variance $\sigma^2 = n \cdot p \cdot q$

Std. Dev. $\sigma = \sqrt{n \cdot p \cdot q}$

Where

n = number of fixed trials

p = probability of success in one of the n trials

q = probability of failure in one of the n trials

Interpretation of Results

It is especially important to interpret results. The range rule of thumb suggests that values are unusual if they lie outside of these limits:

maximum usual value = $\mu + 2\sigma$

minimum usual value = $\mu - 2\sigma$

Example

McDonald's has a 95% recognition rate. A special focus group consists of 12 randomly selected adults.

For such a group, find the mean and standard deviation.

$$\mu = np = 12(0.95) = 11.4$$

$$\sigma = \sqrt{npq} = \sqrt{12(0.95)(0.05)} = 0.754983 = 0.8 \text{ (rounded)}$$

Example - continued

Use the range rule of thumb to find the minimum and maximum usual number of people who would recognize McDonald's.

$$\mu + 2\sigma = 11.4 + 2(0.8) = 13 \text{ people}$$

$$\mu - 2\sigma = 11.4 - 2(0.8) = 9.8 \text{ people}$$

If a particular group of 12 people had all 12 recognize the brand name of McDonald's, that would ***not*** be unusual.

