Chapter 5
Probability Distributions

Slide 1

Overview

Slide 2

This chapter will deal with the construction of

probability distributions

by combining the methods of *descriptive statistics* presented in Chapter 2 and those of *probability* presented in Chapter 3.

Probability Distributions will describe what will *probably* happen instead of what actually *did* happen.

Combining Descriptive Methods and Probabilities

Slide 3

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

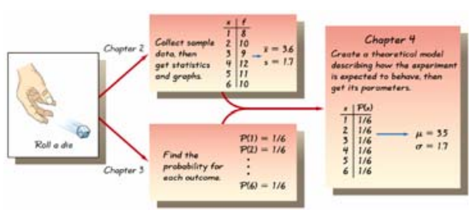


Figure 4-1

Definitions

Slide 4

- **A random variable is a variable (typically represented by *x*) that has a single numerical value, determined by chance, for each outcome of a procedure.**

- **A probability distribution is a graph, table, or formula that gives the probability for each value of the random variable.**

Definitions

Slide 5

- **A discrete random variable has 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 they result from a counting process.**

- **A continuous random variable has infinitely many values, and those values can be associated with measurements on a continuous scale in such a way that there are no gaps or interruptions.**

Graphs

Slide 6

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

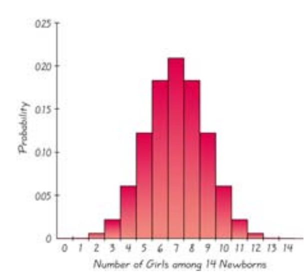


Figure 4-3

Requirements for Probability Distribution

Slide 7

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

where x assumes all possible values

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

for every individual value of x

Mean, Variance and Standard Deviation of a Probability Distribution

Slide 8

Mean $\mu = \sum (x \bullet P(x))$

Variance $\sigma^2 = \sum x^2 \bullet P(x) - \mu^2$

Standard Deviation $\sigma = \sqrt{\sigma^2}$

Using TI Probability Distribution

Slide 9

1) Enter *x*-values into **L1**,

2) Enter corresponding *P(x)* into **L2**,

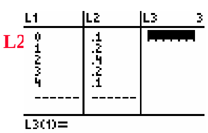
3) stat, calc, 1-var stats, **L1**, **,** , **L2**, enter

Note: n=1 and S_x=blank

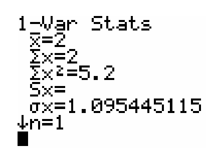
Using TI: Probability Histogram

Slide 10

1) Enter these numbers into **L1** and **L2**



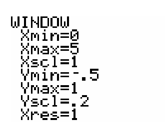
2) Then perform: stat, calc, 1-var stats, **L1**, **,** , **L2**, enter.



Using TI: Probability Histogram

Slide 11

3) Select Window, and enter the following setting.

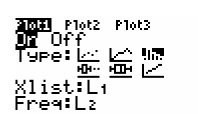


4) Then perform:
2nd , STATPLOT, ENTER to select PLOT1
Make sure other plots are turned off.

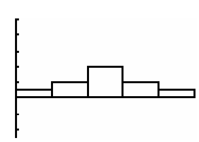
Using TI: Probability Histogram

Slide 12

5) Now enter the following setting for PLOT1



6) Now select GRAPH



Roundoff Rule for μ, σ, and σ^2

Slide 13

Round results by carrying one more decimal place than the number of decimal places used for the random variable x. If the values of x are integers, round μ, σ, and σ^2 and to one decimal place.

Identifying Unusual Results Range Rule of Thumb

Slide 14

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 = $\mu - 2\sigma$

Identifying Unusual Results Probabilities

Slide 15

Rare Event Rule

If, under a given assumption (such as the assumption that boys and girls are equally likely), the probability of a particular observed event (such as 13 girls in 14 births) is extremely small, we conclude that the assumption is probably not correct.

- Unusually high: x successes among n trials is an *unusually high* number of successes if $P(x$ or more) is very small (such as 0.05 or less).
- Unusually low: x successes among n trials is an *unusually low* number of successes if $P(x$ or fewer) is very small (such as 0.05 or less).

Definition

Slide 16

The expected value of a discrete random variable is denoted by E, and it represents the average value of the outcomes. It is obtained by finding the value of $\Sigma [x \cdot P(x)]$.

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

Definitions

Slide 17

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*.
4. The probabilities must remain *constant* for each trial.

Notation for Binomial Probability Distributions

Slide 18

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

$P(S) = p$ (p = probability of success)

$P(F) = 1 - p = q$ (q = probability of failure)

Notation (cont)

n denotes the number of fixed 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.

Important Hints

- Be sure that *x* and *p* both refer to the same category being called a success.
- When sampling without replacement, the events can be treated as if they were independent if the sample size is no more than 5% of the population size. (That is *n* is less than or equal to 0.05*N*.)

Methods for Finding Probabilities

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

Method 1: Using the Binomial Probability Formula

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

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 ($q = 1 - p$)

Binomial Probability Formula

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

Number of outcomes with exactly *x* successes among *n* trials

Probability of *x* successes among *n* trials for any one particular order

Binomial Probability Formula

$$P(x) = {}_nC_x \bullet p^x \bullet q^{n-x}$$

Number of outcomes with exactly *x* successes among *n* trials

Probability of *x* successes among *n* trials for any one particular order

Using TI: Binomial Distribution Evaluation of the formula

Slide 25

Example: **Find P(x=2) when n=6, and p=0.7.**

```
6 nCr 2*
```

1) Enter 6, MATH, PRB, nCr, 2, X

2) 0.7 ^2 X 0.3 ^ 4 , then enter

```
6 nCr 2*0.7^2*0.
3^4
```

Final Answer:

```
6 nCr 2*0.7^2*0.
3^4
             .059535
```


Method 2: Using Table A-1 in Appendix A

Slide 26

Part of Table A-1 is shown below. With *n* = 4 and *p* = 0.2 in the binomial distribution, the probabilities of 0, 1, 2, 3, and 4 successes are 0.410, 0.410, 0.154, 0.026, and 0.002 respectively.

From Table A-1:

n	*x*	*p* 0.20
4	0	0.410
	1	0.410
	2	0.154
	3	0.026
	4	0.002

→

Binomial probability distribution for $n = 4$ and $p = 0.2$

x	*P*(*x*)
0	0.410
1	0.410
2	0.154
3	0.026
4	0.002

Method 3: Using Technology

Slide 27

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

Using TI Binomial Distribution

Slide 28

$$P(x = a)$$

1) 2nd VARS(DISTR)

2) Arrow down to binompdf(

3) enter

4) n, , , p , a) enter

Using TI: Binomial Distribution

Slide 29

Example: **Find P(x=2) when n=6, and p=0.7.**

1) Select 2nd, VARS, arrow down to get to **0:binompdf(** enter to select

```
DISTR DRAW
4↑tpdf(
5:tcdf(
6:X²pdf(
7:X²cdf(
8:Fpdf(
9:Fcdf(
0:binompdf(
```

2) Enter **6,0.7, 2)**

```
binompdf(6,0.7,2
)
```


Using TI: Binomial Distribution

Slide 30

Example: **Find P(x=2) when n=6, and p=0.7.**

3) Enter to execute this operation and get the final answer.

```
binompdf(6,0.7,2
)
             .059535
```

This result was obtained earlier by directly using the binomial distribution formula.

Using TI
Binomial Distribution

Slide 31

$P(x \le a)$

1) 2nd VARS(DISTR)

2) Arrow down to binomcdf(

3) enter

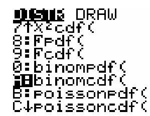
4) n, , , p , a) enter

Using TI: Binomial Distribution

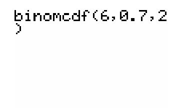
Slide 32

Example: **Find P(x<=2) when n=6, and p=0.7.**

1) Select 2nd, VARS, arrow down to get to **0:binomcdf(** enter to select



2) Enter **6,0.7, 2)**

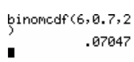


Using TI: Binomial Distribution

Slide 33

Example: **Find P(x<=2) when n=6, and p=0.7.**

3) Enter to execute this operation and get the final answer.



P(x<=2)=P(x=0)+(P(x=1)+P(x=2)

Using TI
Binomial Distribution

Slide 34

$P(x \ge a)$

1) Enter 1 – 2nd VARS(DISTR)

2) Arrow down to binomcdf(

3) enter

4) n, , , p , a–1) enter

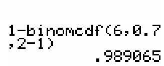
Using TI: Binomial Distribution

Slide 35

Example: **Find P(x>=2) when n=6, and p=0.7.**

1) Enter 1– Select 2nd, VARS, arrow down to get to **0:binomcdf(** enter to select, then type **6,0.7, 2–1)**

2) Enter to get the final answer



P(x>=2)=P(x=2)+P(x=3)+●●●+P(x=6)

Binomial Distribution: Formulas

Slide 36

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**

Example

Find the mean and standard deviation for the number of girls in groups of 14 births.

This scenario is a binomial distribution where:

$n = 14$

$p = 0.5$

$q = 0.5$

Using the binomial distribution formulas:

Example (cont)

This scenario is a binomial distribution where

$n = 14$

$p = 0.5$

$q = 0.5$

Using the binomial distribution formulas:

$\mu = (14)(0.5) = 7$ girls

$\sigma = \sqrt{(14)(0.5)(0.5)} = 1.9$ girls (rounded)

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:

$$\text{Maximum usual values} = \mu + 2\,\sigma$$

$$\text{Minimum usual values} = \mu - 2\,\sigma$$

Example

Determine whether 68 girls among 100 babies could easily occur by chance.

For this binomial distribution,

$\mu = 50$ girls

$\sigma = 5$ girls

$\mu + 2\,\sigma = 50 + 2(5) = 60$

$\mu - 2\,\sigma = 50 - 2(5) = 40$

The usual number girls among 100 births would be from 40 to 60. So 68 girls in 100 births is an unusual result.

Definition

The Poisson distribution is a discrete probability distribution that applies to occurrences of some event *over a specified interval*. The random variable x is the number of occurrences of the event in an interval. The interval can be time, distance, area, volume, or some similar unit.

Formula

$$P(x) = \frac{\mu^{x} \cdot e^{-\mu}}{x!} \quad \text{where } e \approx 2.71828$$

Poisson Distribution Requirements

- The random variable x is the number of occurrences of an event *over some interval.*
- The occurrences must be *random.*
- The occurrences must be *independent* of each other.
- The occurrences must be *uniformly distributed* over the interval being used.

Parameters

- The mean is μ.
- The standard deviation is $\sigma = \sqrt{\mu}$.

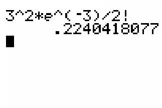
Using TI: Poisson Distribution Evaluation of the formula

Slide 43

Example: **Find P(x=2) when μ=3.**

1) Enter 3 ^ 2 X 2nd LN (e^(– 3) ÷ 2 MATH PRB !

2) then enter for **Final Answer:**



Using TI Poisson Distribution

Slide 44

$$P(x = a)$$

1) 2nd VARS(DISTR)

2) Arrow down to poissonpdf(

3) enter

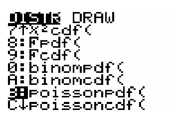
4) μ, , , a) enter

Using TI: Poisson Distribution

Slide 45

Example: **Find P(x=2) when μ=3.**

1) 2nd VARS arrow down to **poissonpdf(**



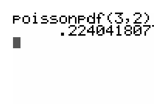
2) enter to select, then type **3 , 2)**

Using TI: Poisson Distribution

Slide 46

Example: **Find P(x=2) when μ=3.**

3) enter to get the **Final Answer:**



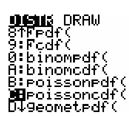
This result was obtained earlier by directly using the Poisson distribution formula.

Using TI: Poisson Distribution

Slide 47

Example: **Find P(x<=3) when μ=3.**

1) 2nd VARS arrow down to **poissoncdf(**



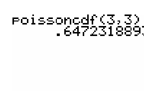
2) enter to select, then type **3 , 3)**

Using TI: Poisson Distribution

Slide 48

Example: **Find P(x<=3) when μ=3.**

3) enter to get the **Final Answer:**



P(x<=3)=P(x=0)+P(x=1)+P(x=2)+P(x=3)

Difference from a Binomial Distribution

Slide 49

The Poisson distribution differs from the binomial distribution in these fundamental ways:

- The binomial distribution is affected by the sample size n and the probability p, whereas the Poisson distribution is affected only by the mean μ.
- In a binomial distribution the possible values of the random variable are x are 0, 1, . . . n, but a Poisson distribution has possible x values of 0, 1, . . . , with no upper limit.

Example

Slide 50

World War II Bombs In analyzing hits by V-1 buzz bombs in World War II, South London was subdivided into 576 regions, each with an area of 0.25 km^2. A total of 535 bombs hit the combined area of 576 regions

If a region is randomly selected, find the probability that it was hit exactly twice.

The Poisson distribution applies because we are dealing with occurrences of an event (bomb hits) over some interval (a region with area of 0.25 km^2).

Example

Slide 51

The mean number of hits per region is

$$\mu = \frac{\text{number of bomb hits}}{\text{number of regions}} = \frac{535}{576} = 0.929$$

$$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \frac{0.929^2 \cdot 2.71828^{-0.929}}{2!} = \frac{0.863 \cdot 0.395}{2} = 0.170$$

The probability of a particular region being hit exactly twice is $P(2) = 0.170$.

Poisson as Approximation to Binomial

Slide 52

The Poisson distribution is sometimes used to approximate the binomial distribution when n is large and p is small.

Rule of Thumb

- $n \geq 100$
- $np \leq 10$

Poisson as Approximation to Binomial - μ

Slide 53

- $n \geq 100$
- $np \leq 10$

Value for μ

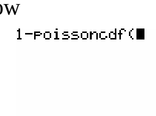
$$\mu = n \cdot p$$

Slide 54

Monterey Park Police department issues 5 J-walking citations in average during one week of school to ELAC students. Find the probability that at least 4 J-walking tickets will be issued this week by MPPD?

Solution: P(x>=4)=1–P(x<4)

1) Enter 1–2nd VARS arrow down to **poissoncdf(**



2) ENTER to select

Slide 55

Monterey Park Police department issues 5 J-walking citations in average during one week of school to ELAC students. Find the probability that at least 4 J-walking tickets will be issued this week by MPPD?

3) Type **7 , 4 – 1)**

4) Enter to execute this command and get the **final answer**.