Chapter 5

Discrete and Continuous Probability Distributions

Chapter Goals

After completing this chapter, you should be able to:

- Apply the binomial distribution to applied problems
- Compute probabilities for the Poisson and hypergeometric distributions
- Find probabilities using a normal distribution table and apply the normal distribution to business problems
- Recognize when to apply the uniform and exponential distributions

Probability Distributions

Discrete Probability Distributions
- Binomial
- Poisson
- Hypergeometric

Continuous Probability Distributions
- Normal
- Uniform
- Exponential
Discrete Probability Distributions

- A discrete random variable is a variable that can assume only a countable number of values
  - Many possible outcomes:
    - number of complaints per day
    - number of TV's in a household
    - number of rings before the phone is answered
  - Only two possible outcomes:
    - gender: male or female
    - defective: yes or no
    - spreads peanut butter first vs. spreads jelly first

Continuous Probability Distributions

- A continuous random variable is a variable that can assume any value on a continuum (can assume an uncountable number of values)
  - thickness of an item
  - time required to complete a task
  - temperature of a solution
  - height, in inches
- These can potentially take on any value, depending only on the ability to measure accurately.

The Binomial Distribution
The Binomial Distribution

Characteristics of the Binomial Distribution:

- A trial has only two possible outcomes – “success” or “failure”
- There is a fixed number, n, of identical trials
- The trials of the experiment are independent of each other
- The probability of a success, p, remains constant from trial to trial
- If p represents the probability of a success, then \((1-p) = q\) is the probability of a failure

Binomial Distribution Settings

- A manufacturing plant labels items as either defective or acceptable
- A firm bidding for a contract will either get the contract or not
- A marketing research firm receives survey responses of “yes I will buy” or “no I will not”
- New job applicants either accept the offer or reject it

Counting Rule for Combinations

A combination is an outcome of an experiment where x objects are selected from a group of n objects

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

where:

- \(n! = n(n - 1)(n - 2) \ldots (2)(1)\)
- \(x! = x(x - 1)(x - 2) \ldots (2)(1)\)
- \(0! = 1\) (by definition)
Binomial Distribution Formula

\[ P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x} \]

- \( P(x) \) is the probability of \( x \) successes in \( n \) trials, with probability of success \( p \) on each trial.
- \( x \) = number of 'successes' in sample, \( x = 0, 1, 2, \ldots, n \)
- \( p \) = probability of "success" per trial
- \( q \) = probability of "failure" = \( 1 - p \)
- \( n \) = number of trials (sample size)

Example: Flip a coin four times, let \( x = \) # heads:
- \( n = 4 \)
- \( p = 0.5 \)
- \( q = (1 - .5) = .5 \)
- \( x = 0, 1, 2, 3, 4 \)

Binomial Distribution

The shape of the binomial distribution depends on the values of \( p \) and \( n \).

Mean

- Here, \( n = 5 \) and \( p = .1 \)
- Here, \( n = 5 \) and \( p = .5 \)

Variance and Standard Deviation

\[ \mu = E(x) = np \]
\[ \sigma^2 = npq \]
\[ \sigma = \sqrt{npq} \]

Where:
- \( n \) = sample size
- \( p \) = probability of success
- \( q = (1 - p) = \) probability of failure
### Binomial Characteristics

**Examples**

- \( n = 5, p = 0.1 \)
  
  \[ \mu = np = 5 \times 0.1 = 0.5 \]
  
  \[ \sigma = \sqrt{npq} = \sqrt{5 \times 0.1 \times 0.9} = 0.6708 \]

- \( n = 5, p = 0.5 \)
  
  \[ \mu = np = 5 \times 0.5 = 2.5 \]
  
  \[ \sigma = \sqrt{npq} = \sqrt{5 \times 0.5 \times 0.5} = 1.118 \]

### Using Binomial Tables

<table>
<thead>
<tr>
<th>x</th>
<th>p=.50</th>
<th>p=.45</th>
<th>p=.40</th>
<th>p=.35</th>
<th>p=.30</th>
<th>p=.25</th>
<th>p=.20</th>
<th>p=.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0010</td>
<td>0.0117</td>
<td>0.0439</td>
<td>0.0763</td>
<td>0.1172</td>
<td>0.1580</td>
<td>0.1988</td>
<td>0.2396</td>
</tr>
<tr>
<td>1</td>
<td>0.0098</td>
<td>0.0638</td>
<td>0.2051</td>
<td>0.2461</td>
<td>0.2870</td>
<td>0.3280</td>
<td>0.3690</td>
<td>0.4100</td>
</tr>
<tr>
<td>2</td>
<td>0.0439</td>
<td>0.1462</td>
<td>0.2870</td>
<td>0.3280</td>
<td>0.3690</td>
<td>0.4100</td>
<td>0.4510</td>
<td>0.4920</td>
</tr>
<tr>
<td>3</td>
<td>0.0980</td>
<td>0.2531</td>
<td>0.3690</td>
<td>0.4510</td>
<td>0.5320</td>
<td>0.6130</td>
<td>0.6940</td>
<td>0.7750</td>
</tr>
<tr>
<td>4</td>
<td>0.1520</td>
<td>0.3081</td>
<td>0.4510</td>
<td>0.6130</td>
<td>0.7750</td>
<td>0.9370</td>
<td>1.0990</td>
<td>1.2610</td>
</tr>
<tr>
<td>5</td>
<td>0.2061</td>
<td>0.3632</td>
<td>0.6130</td>
<td>0.9370</td>
<td>1.2610</td>
<td>1.5950</td>
<td>1.9290</td>
<td>2.2630</td>
</tr>
<tr>
<td>6</td>
<td>0.2602</td>
<td>0.4183</td>
<td>0.9370</td>
<td>1.5950</td>
<td>2.2630</td>
<td>2.9310</td>
<td>3.6090</td>
<td>4.2870</td>
</tr>
<tr>
<td>7</td>
<td>0.3143</td>
<td>0.4734</td>
<td>1.5950</td>
<td>2.9310</td>
<td>4.2870</td>
<td>5.6690</td>
<td>7.0570</td>
<td>8.4450</td>
</tr>
<tr>
<td>8</td>
<td>0.3683</td>
<td>0.5285</td>
<td>2.9310</td>
<td>5.6690</td>
<td>8.4450</td>
<td>10.3370</td>
<td>12.2250</td>
<td>14.1130</td>
</tr>
<tr>
<td>9</td>
<td>0.4223</td>
<td>0.5836</td>
<td>5.6690</td>
<td>10.3370</td>
<td>14.1130</td>
<td>17.0290</td>
<td>19.9450</td>
<td>22.8610</td>
</tr>
<tr>
<td>10</td>
<td>0.4763</td>
<td>0.6387</td>
<td>8.4450</td>
<td>17.0290</td>
<td>22.8610</td>
<td>25.6550</td>
<td>28.4490</td>
<td>31.2430</td>
</tr>
</tbody>
</table>

**Examples:**

- \( n = 10, p = 0.35, x = 3 \): \( P(x = 3|n=10, p = 0.35) = 0.2522 \)
- \( n = 10, p = 0.75, x = 2 \): \( P(x = 2|n=10, p = 0.75) = 0.0004 \)

### Using PHStat

Select PHStat / Probability & Prob. Distributions / Binomial...
Using PHStat

- Enter desired values in dialog box

Here: n = 10
p = .35

Output for x = 0 to x = 10 will be generated by PHStat
Optional check boxes for additional output

PHStat Output

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0531</td>
</tr>
<tr>
<td>1</td>
<td>0.1603</td>
</tr>
<tr>
<td>2</td>
<td>0.2134</td>
</tr>
<tr>
<td>3</td>
<td>0.2522</td>
</tr>
<tr>
<td>4</td>
<td>0.2522</td>
</tr>
<tr>
<td>5</td>
<td>0.1966</td>
</tr>
</tbody>
</table>

P(x = 3 | n = 10, p = .35) = .2522

P(x > 5 | n = 10, p = .35) = .0949

The Poisson Distribution
The Poisson Distribution

Characteristics of the Poisson Distribution:

- The outcomes of interest are rare relative to the possible outcomes.
- The average number of outcomes of interest per time or space interval is \( \lambda \).
- The number of outcomes of interest are random, and the occurrence of one outcome does not influence the chances of another outcome of interest.
- The probability of that an outcome of interest occurs in a given segment is the same for all segments.

Poisson Distribution Formula

\[
P(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}
\]

where:
- \( t \) = size of the segment of interest
- \( x \) = number of successes in segment of interest
- \( \lambda \) = expected number of successes in a segment of unit size
- \( e \) = base of the natural logarithm system (2.71828...)

Poisson Distribution Characteristics

- Mean
  \[
  \mu = \lambda t
  \]
- Variance and Standard Deviation
  \[
  \sigma^2 = \lambda t \\
  \sigma = \sqrt{\lambda t}
  \]

where:
- \( \lambda \) = number of successes in a segment of unit size
- \( t \) = the size of the segment of interest
Using Poisson Tables

<table>
<thead>
<tr>
<th>x</th>
<th>0.00</th>
<th>0.05</th>
<th>0.10</th>
<th>0.15</th>
<th>0.20</th>
<th>0.25</th>
<th>0.30</th>
<th>0.35</th>
<th>0.40</th>
<th>0.45</th>
<th>0.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>6</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>7</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Example: Find \( P(x = 2) \) if \( \lambda = 0.05 \) and \( t = 100 \)

\[
P(x = 2) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} = \frac{(0.05)^2 e^{-0.50}}{2!} = 0.0758
\]
The Hypergeometric Distribution

Probability Distributions

Discrete Probability Distributions

Binomial
Poisson
Hypergeometric

"n" trials in a sample taken from a finite population of size N
Sample taken without replacement
Trials are dependent
Concerned with finding the probability of "x" successes in the sample where there are "X" successes in the population

Hypergeometric Distribution Formula

(Two possible outcomes per trial)

\[ P(x) = \frac{C_{N-X}^{n-x} \cdot C_x^x}{C_N^n} \]

Where

- \( N \) = Population size
- \( X \) = number of successes in the population
- \( n \) = sample size
- \( x \) = number of successes in the sample
- \( n - x \) = number of failures in the sample
Hypergeometric Distribution

Formula

Example: 3 Light bulbs were selected from 10. Of the 10 there were 4 defective. What is the probability that 2 of the 3 selected are defective?

\[ N = 10 \quad n = 3 \]
\[ X = 4 \quad x = 2 \]

\[ P(x = 2) = \frac{C_{N-X}^n \cdot C_{X}^x}{C_{N}^n} = \frac{C_{6}^3 \cdot C_{4}^2}{C_{10}^3} = \frac{6 \cdot 6}{120} = 0.3 \]

Example: 3 Light bulbs were selected from 10. Of the 10 there were 4 defective. What is the probability that 2 of the 3 selected are defective?

\[ N = 10 \quad n = 3 \]
\[ X = 4 \quad x = 2 \]

\[ P(x = 2) = \frac{C_{N-X}^n \cdot C_{X}^x}{C_{N}^n} = \frac{C_{6}^3 \cdot C_{4}^2}{C_{10}^3} = \frac{6 \cdot 6}{120} = 0.3 \]

Hypergeometric Distribution in PHStat

Select:

PHStat / Probability & Prob. Distributions / Hypergeometric ...

Hypergeometric Distribution in PHStat

(continued)

\[ P(x = 2) = 0.3 \]
The Normal Distribution

- "Bell Shaped"
- Symmetrical
- Mean, Median and Mode are Equal
  - Location is determined by the mean, \( \mu \)
  - Spread is determined by the standard deviation, \( \sigma \)
  - The random variable has an infinite theoretical range: \( -\infty \) to \( +\infty \)

The Normal Distribution

- By varying the parameters \( \mu \) and \( \sigma \), we obtain different normal distributions
The Normal Distribution Shape

Changing $\mu$ shifts the distribution left or right.

Changing $\sigma$ increases or decreases the spread.

---

Finding Normal Probabilities

Probability is measured by the area under the curve.

$P(a \leq x \leq b)$

---

Probability as Area Under the Curve

The total area under the curve is 1.0, and the curve is symmetric, so half is above the mean, half is below.

$P(-\infty < x < \mu) = 0.5$

$P(\mu < x < \infty) = 0.5$

$P(-\infty < x < \infty) = 1.0$
Empirical Rules

What can we say about the distribution of values around the mean? There are some general rules:

\[ \mu \pm 1\sigma \text{ encloses about } 68\% \text{ of } x's \]

\[ \mu \pm 2\sigma \text{ covers about } 95\% \text{ of } x's \]

\[ \mu \pm 3\sigma \text{ covers about } 99.7\% \text{ of } x's \]

The Empirical Rule (continued)

- \( \mu \pm 2\sigma \) covers about 95% of x's
- \( \mu \pm 3\sigma \) covers about 99.7% of x's

Importance of the Rule

- If a value is about 2 or more standard deviations away from the mean in a normal distribution, then it is far from the mean.
- The chance that a value that far or farther away from the mean is highly unlikely, given that particular mean and standard deviation.
The Standard Normal Distribution

- Also known as the “z” distribution
- Mean is defined to be 0
- Standard Deviation is 1

Values above the mean have positive z-values, values below the mean have negative z-values.

The Standard Normal

- Any normal distribution (with any mean and standard deviation combination) can be transformed into the standard normal distribution (z)
- Need to transform x units into z units

Translation to the Standard Normal Distribution

- Translate from x to the standard normal (the “z” distribution) by subtracting the mean of x and dividing by its standard deviation:

$$z = \frac{x - \mu}{\sigma}$$
Example

- If $x$ is distributed normally with mean of 100 and standard deviation of 50, the $z$ value for $x = 250$ is

$$z = \frac{x - \mu}{\sigma} = \frac{250 - 100}{50} = 3.0$$

- This says that $x = 250$ is three standard deviations (3 increments of 50 units) above the mean of 100.

Comparing $x$ and $z$ units

Note that the distribution is the same, only the scale has changed. We can express the problem in original units ($x$) or in standardized units ($z$).

The Standard Normal Table

- The Standard Normal table in the textbook (Appendix D) gives the probability from the mean (zero) up to a desired value for $z$.

Example:

$$P(0 < z < 2.00) = .4772$$
The Standard Normal Table

The value within the table gives the probability from \( z = 0 \) up to the desired \( z \) value.

<table>
<thead>
<tr>
<th>( z )</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>...</td>
</tr>
<tr>
<td>2.0</td>
<td>0.4772</td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The row shows the value of \( z \) to the first decimal point.
The column gives the value of \( z \) to the second decimal point.

\[ P(0 < z < 2.00) = 0.4772 \]

General Procedure for Finding Probabilities

To find \( P(a < x < b) \) when \( x \) is distributed normally:

- Draw the normal curve for the problem in terms of \( x \)
- Translate \( x \)-values to \( z \)-values
- Use the Standard Normal Table

Z Table example

Suppose \( x \) is normal with mean 8.0 and standard deviation 5.0. Find \( P(8 < x < 8.6) \)

Calculate \( z \)-values:

\[ z = \frac{x - \mu}{\sigma} = \frac{8 - 8}{5} = 0 \]

\[ z = \frac{x - \mu}{\sigma} = \frac{8.6 - 8}{5} = 0.12 \]

\[ P(8 < x < 8.6) = P(0 < z < 0.12) \]
Z Table example

Suppose \( x \) is normal with mean 8.0 and standard deviation 5.0. Find \( P(8 < x < 8.6) \)

\[ \mu = 8 \]
\[ \sigma = 5 \]

\[ P(8 < x < 8.6) \]

\[ P(0 < z < 0.12) \]

\[ z = 0.12 \]

Solution: Finding \( P(0 < z < 0.12) \)

Standard Normal Probability Table (Portion)

<table>
<thead>
<tr>
<th>( z )</th>
<th>.00</th>
<th>.01</th>
<th>.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0000</td>
<td>0.0040</td>
<td>0.0080</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0398</td>
<td>0.0438</td>
<td>0.0478</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0793</td>
<td>0.0832</td>
<td>0.0871</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1179</td>
<td>0.1217</td>
<td>0.1255</td>
</tr>
</tbody>
</table>

\[ P(8 < x < 8.6) = P(0 < z < 0.12) = 0.0478 \]

Finding Normal Probabilities

- Suppose \( x \) is normal with mean 8.0 and standard deviation 5.0.
- Now find \( P(x < 8.6) \)
Finding Normal Probabilities

Suppose \( x \) is normal with mean 8.0 and standard deviation 5.0.

Now Find \( P(x < 8.6) \)

\[
P(x < 8.6) = P(z < 0.12) = P(z < 0) + P(0 < z < 0.12)
\]

\[
= .5 + .0478 = .5478
\]

Upper Tail Probabilities

Suppose \( x \) is normal with mean 8.0 and standard deviation 5.0.

Now Find \( P(x > 8.6) \)

\[
P(x > 8.6) = P(z > 0.12) = P(z > 0) - P(0 < z < 0.12)
\]

\[
= .5 - .0478 = .4522
\]
Suppose \( x \) is normal with mean 8.0 and standard deviation 5.0.

Now find \( P(7.4 < x < 8) \)

\[
\begin{align*}
Z & = \frac{7.4 - 8.0}{5} \\
& = -0.12
\end{align*}
\]

The Normal distribution is symmetric, so we use the same table even if \( z \)-values are negative:

\[
P(7.4 < x < 8) = P(-0.12 < z < 0) = 0.0478
\]

(continued)

We can use Excel and PHStat to quickly generate probabilities for any normal distribution.

We will find \( P(8 < x < 8.6) \) when \( x \) is normally distributed with mean 8 and standard deviation 5.
The Uniform Distribution

Continuous Probability Distributions

- Normal
- Uniform
- Exponential
The Uniform Distribution

- The uniform distribution is a probability distribution that has equal probabilities for all possible outcomes of the random variable.

The Continuous Uniform Distribution:

\[ f(x) = \begin{cases} 
  \frac{1}{b-a} & \text{if } a \leq x \leq b \\
  0 & \text{otherwise}
\end{cases} \]

where
- \( f(x) \) = value of the density function at any \( x \) value
- \( a \) = lower limit of the interval
- \( b \) = upper limit of the interval

Example: Uniform Probability Distribution

Over the range \( 2 \leq x \leq 6 \):

\[ f(x) = \frac{1}{6 - 2} = \frac{1}{4} = 0.25 \text{ for } 2 \leq x \leq 6 \]
The Exponential Distribution

- Used to measure the time that elapses between two occurrences of an event (the time between arrivals)

- Examples:
  - Time between trucks arriving at an unloading dock
  - Time between transactions at an ATM Machine
  - Time between phone calls to the main operator

The probability that an arrival time is equal to or less than some specified time \( a \) is

\[
P(0 \leq x \leq a) = 1 - e^{-\lambda a}
\]

where \( 1/\lambda \) is the mean time between events

Note that if the number of occurrences per time period is Poisson with mean \( \lambda \), then the time between occurrences is exponential with mean time \( 1/\lambda \).
Exponential Distribution

- Shape of the exponential distribution

Example

Example: Customers arrive at the claims counter at the rate of 15 per hour (Poisson distributed). What is the probability that the arrival time between consecutive customers is less than five minutes?

- Time between arrivals is exponentially distributed with mean time between arrivals of 4 minutes (15 per 60 minutes, on average)
- $1/\lambda = 4.0$, so $\lambda = .25$
- $P(x < 5) = 1 - e^{-\lambda x} = 1 - e^{-(.25)(5)} = 0.7135$

Chapter Summary

- Reviewed key discrete distributions: binomial, poisson, hypergeometric
- Reviewed key continuous distributions: normal, uniform, exponential
- Found probabilities using formulas and tables
- Recognized when to apply different distributions
- Applied distributions to decision problems