Class 16

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
Agenda:
Recap Chapter 7.3
Lecture Chapter 7.4
Discussion: Chapters
Review Chapter 6.
Recap Chapter 7.3
Sample distribution of sample means (SDSM): If all possible random samples, each of size $n$, are taken from any population with mean $\mu$ and standard deviation $\sigma$, then the sampling distribution of sample means will have the following:

1. A mean $\mu_{\bar{x}}$ equal to $\mu$
2. A standard deviation $\sigma_{\bar{x}}$ equal to $\frac{\sigma}{\sqrt{n}}$

Furthermore, if the sampled population has a normal distribution, then the sampling distribution of $\bar{x}$ will also be normal for all samples of all sizes.
The CLT is important because we are going to make probability statements about the population mean \( \mu \) based upon a sample of data \( x_1, \ldots, x_n \) and a single sample mean \( \bar{x} \).

We may not know the distribution that the individual observations come from, but we will know by CLT that the sample mean is approximately normal with

\[
\mu_{\bar{x}} = \mu, \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}.
\]
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

According to SDSM, if we had a sample $x_1, \ldots, x_n$ of size $n = 1, 2, 3, 4, 5, 15, 30, \text{ and } 50$ from Uniform(0,200) or Normal(100,(57.7)^2)

<table>
<thead>
<tr>
<th>Sample Size, $n$</th>
<th>Mean, $\mu_{\bar{x}}$</th>
<th>SD, $\sigma_{\bar{x}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>57.7350</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>40.8248</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>33.3333</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>28.8675</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>25.8199</td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>14.9071</td>
</tr>
<tr>
<td>30</td>
<td>100</td>
<td>10.5409</td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>8.1650</td>
</tr>
</tbody>
</table>

For Uniform $a$ to $b$

\[
\mu = \frac{b - a}{2}
\]

\[
\sigma^2 = \frac{(b - a)^2}{12}
\]

Theoretical values from the SDSM

\[\mu_{\bar{x}} = \mu\]

\[\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\]
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

I wrote a computer program to generate 1 million random observations $x_1, \ldots, x_{1000000}$ from the Uniform($a=0, b=200$) and also from Normal($\mu=100, \sigma^2=(57.7)^2$) distributions.

I showed you the histograms being formed!
Expanded program to generate $n$ million random observations $x_1, \ldots, x_{n\times10^6}$ from the Uniform($a=0, b=200$) and also from the Normal($\mu=100, \sigma^2=(57.7)^2$) distributions, for each of $n=1, 2, 3, 4, 5, 15, 30,$ and $50$.

8 data sets of Uniform and Normal random observations
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\( n=1 \) \( 1 \times 10^6 \) observations

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

observations are uniform

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

observations are normal

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\( n=2 \) \quad 2 \times 10^6 \text{ observations}

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

observations are uniform

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

observations are normal

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n=3 \quad 3 \times 10^6 \text{ observations} \]

\[ \mu = 100 \quad \sigma = 57.73 \]

observations are uniform

\[ \mu = 100 \quad \sigma = 57.73 \]

observations are normal

Rowe, D.B.
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

\[ n = 4 \quad 4 \times 10^6 \text{ observations} \]

- \( \mu = 100 \)
- \( \sigma = 57.73 \)

observations are uniform

- \( \mu = 100 \)
- \( \sigma = 57.73 \)

observations are normal

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n=5 \quad 5 \times 10^6 \text{ observations} \]

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

observations are uniform

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

observations are normal

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n = 15 \quad 15 \times 10^6 \text{ observations} \]

\[ \mu = 100 \quad \sigma = 57.73 \]

observations are uniform

\[ \mu = 100 \quad \sigma = 57.73 \]

observations are normal
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n = 30 \quad 30 \times 10^6 \text{ observations} \]

\[ \mu = 100 \quad \sigma = 57.73 \]

observations are uniform

observations are normal

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n = 50 \quad 50 \times 10^6 \text{ observations} \]

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

observations are uniform

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

observations are normal
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

Sample means and standard deviations from each of the \( n \) million observations from the Uniform\((a=0, b=200)\) and Normal\((\mu=100, \sigma^2=(57.7)^2)\) distributions.

i.e. \( n=5, 5\times10^6 \)

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( \ldots )</th>
<th>( x_{5000000} )</th>
</tr>
</thead>
</table>

Groups of \( n=5 \)

\( x_1, x_2, x_3, x_4, x_5 \)

\( x_6, x_7, x_8, x_9, x_{10} \)

\( \ldots \)

\( x_{4999996}, \ldots, x_{5000000} \)

Mean of groups

\( \bar{x}_1 \)

\( \bar{x}_2 \)

\( \ldots \)

\( \bar{x}_{10^6} \)

8 data sets of Uniform and Normal

Histogram of \( \bar{x}'s \)

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

Computed sample means and standard deviations from the one million means $\bar{x}_1, \ldots, \bar{x}_{10^6}$

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Mean $\mu_{\bar{x}}$</th>
<th>Mean U $\bar{x}_{\bar{x}}$</th>
<th>Mean N $\bar{x}_{\bar{x}}$</th>
<th>SD $\sigma_{\bar{x}}$</th>
<th>SD U $s_{\bar{x}}$</th>
<th>SD N $s_{\bar{x}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>100.0642</td>
<td>100.0077</td>
<td>57.7350</td>
<td>57.7071</td>
<td>57.7888</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>99.9828</td>
<td>100.0037</td>
<td>40.8248</td>
<td>40.8418</td>
<td>40.8206</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>99.9909</td>
<td>99.9627</td>
<td>33.3333</td>
<td>33.3418</td>
<td>33.2984</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>99.9559</td>
<td>100.0642</td>
<td>28.8675</td>
<td>28.8946</td>
<td>28.8126</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>100.0074</td>
<td>100.0320</td>
<td>25.8199</td>
<td>25.7865</td>
<td>25.8397</td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>100.0134</td>
<td>99.9517</td>
<td>14.9071</td>
<td>14.9035</td>
<td>14.8918</td>
</tr>
<tr>
<td>30</td>
<td>100</td>
<td>99.9934</td>
<td>99.9836</td>
<td>10.5409</td>
<td>10.5335</td>
<td>10.5352</td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>99.9918</td>
<td>99.9890</td>
<td>8.1650</td>
<td>8.1605</td>
<td>8.1709</td>
</tr>
</tbody>
</table>

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n = 1 \quad \times 10^6 \text{means} \]

\[ \mu = 100 \]

\[ \sigma = 57.73 \]

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

$n = 2 \quad 1 \times 10^6$ means

- Histogram of means from uniform
- Histogram of means from normal

$\mu = 100$
$\sigma = 57.73$

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ \mu = 100 \]

\[ \sigma = 57.73 \]

Histogram of means from uniform

Histogram of means from normal

\[ S_{\bar{X}} \]

\[ \bar{X} \]

Rowe, D.B.
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

\( n = 4 \quad 1 \times 10^6 \text{ means} \)

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

Histogram of means from uniform

Histogram of means from normal

\( S_{\bar{X}} \)

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n = 5 \quad 1 \times 10^6 \text{ means} \]

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

Histogram of means from uniform

Histogram of means from normal

\[ \overline{X} \]

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n = 15 \quad 1 \times 10^6 \text{ means} \]

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

Histogram of means from uniform

Histogram of means from normal
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n = 30 \quad 1 \times 10^6 \text{ means} \]

\[ \mu = 100 \]

\[ \sigma = 57.73 \]

Histogram of means from uniform

Histogram of means from normal

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

\[ n = 50 \quad 1 \times 10^6 \text{ means} \]

\[ \mu = 100 \]
\[ \sigma = 57.73 \]

Histogram of means from uniform

Histogram of means from normal

Rowe, D.B.
7: Sample Variability

7.3 The Sampling Distribution of Sample Means

With a population mean \( \mu \) and standard deviation \( \sigma \).

Random samples of size \( n \) with replacement, for “large” \( n \), the distribution of the sample means quickly becomes normally distributed with

\[
\mu_{\bar{x}} = \mu, \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}
\]

Generally \( n \geq 30 \) is sufficiently “large,” but can be as small as 15 or as big as 50 depending upon the shape of distribution!
7: Sample Variability

Questions?

Homework: Chapter 7 # 1, 6, 21, 23, 29, 35
Lecture Chapter 7.4
Chapter 7: Sample Variability (continued)

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
Recall: When we take a sample of data $x_1, \ldots, x_n$ from a population, then compute an estimate of a parameter it is called a sample statistic. i.e. $\bar{x}$ for $\mu$.

**Sampling Distribution of a sample statistic:** The distribution of values for a sample statistic obtained from repeated samples, all of the same size and all drawn from the same population.
# 7: Sample Variability

## 7.2 Sampling Distributions

### The Sampling Distribution of Sample Means

| Statistical population being studied | Repeated sampling is needed to form the sampling distribution. | All possible samples of size \( n \) | One value of the sample statistic (\( \bar{x} \) in this case) corresponding to the parameter of interest (\( \mu \) in this case) is obtained from each sample | Then all of these values of the sample statistic, \( \bar{x} \), are used to form the sampling distribution.
|--------------------------------------|---------------------------------------------------------------|-----------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|

As the number of samples increases the empirical dist. turns into theoretical dist.

---

**Figure from Johnson & Kuby, 2008.**

\[
\mu_{\bar{x}} = \mu \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}
\]
3: Descriptive Analysis and Bivariate Data

3.2 Bivariate Data: Scatter Diagram

Our data.

Find yourself. If not here then removed or your response incomplete.

Height vs. Weight

108 responses
7: Sample Variability

7.4 Application of the Sampling Distribution of Sample Means

Assume that this is a population of data.

Does this population look normally distributed?

\[ \mu = 67.0 \]
\[ \sigma = 3.7 \]

\( N = 108 \) values

Put normal with same \( \mu \) and \( \sigma \).
If I’m interested in the mean $\mu_x$ of a sample of size $n=15$, then by SDSM $\mu_x = \mu$ and $\sigma_x = \frac{\sigma}{\sqrt{n}}$, in addition, by the CLT $\bar{x}$ is (hopefully) normally distributed.

I wrote a computer program to take a sample of $n=1,3,5,15$ from the population of $N$ heights with replacement $10^6$ times. $N=108$ values.
7: Sample Variability

7.4 Application of the Sampling Distribution of Sample Means

By SDSM, $\mu_\bar{x} = 67.0$ and $\sigma_\bar{x} = \frac{3.7}{\sqrt{1}} = 3.7$.

Histogram of the 1 million means

$n=1$

Put normal with same $\mu_\bar{x}$ and $\sigma_\bar{x}$.

$N=108$ values

$\mu = 67.0$

$\sigma = 3.7$
7: Sample Variability

7.4 Application of the Sampling Distribution of Sample Means

By SDSM, $\mu_\bar{x} = 67.0$ and $\sigma_\bar{x} = \frac{3.7}{\sqrt{3}} = 2.1$.

$N=108$ values

$\mu = 67.0$

$\sigma = 3.7$

Histogram of the 1 million means

$n=3$

Put normal with same $\mu_\bar{x}$ and $\sigma_\bar{x}$. 

Rowe, D.B.
7: Sample Variability

7.4 Application of the Sampling Distribution of Sample Means

By SDSM, \( \mu_x = 67.0 \) and \( \sigma_x = \frac{3.7}{\sqrt{5}} = 1.7 \).

Histogram of the 1 million means

\( N = 108 \) values

\( \mu = 67.0 \)

\( \sigma = 3.7 \)

Put normal with same \( \mu_x \) and \( \sigma_x \).
7: Sample Variability

7.4 Application of the Sampling Distribution of Sample Means

By SDSM, \( \mu_x = 67.0 \) and \( \sigma_x = \frac{3.7}{\sqrt{15}} = 1.0 \).

Histogram of the 1 million means

Put normal with same \( \mu_x \) and \( \sigma_x \).

By CLT becomes normal

\( N=108 \) values

\( \mu = 67.0 \)

\( \sigma = 3.7 \)
Now that we believe that the mean $\bar{x}$ from a sample of $n=15$ is normally distributed with mean $\mu_{\bar{x}} = \mu$ and standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, we can find probabilities.
To find these probabilities, we first convert to $z$ scores:

\[
P(a < \bar{x} < b)
\]

\[
P(c < z < d)
\]

\[
P(b < \bar{x})
\]

\[
P(d < z)
\]

\[
P(\bar{x} < a)
\]

\[
P(z < c)
\]

\[
z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}},\quad c = \frac{a - \mu_{\bar{x}}}{\sigma_{\bar{x}}},\quad d = \frac{b - \mu_{\bar{x}}}{\sigma_{\bar{x}}}
\]

and use the table in book.
Example:
What is probability that sample mean $\bar{x}$ from a random sample of $n=15$ heights is greater than 69" when $\mu = 67.0$ and $\sigma = 3.7$?
7: Sample Variability

Questions?

Homework: Chapter 7 # 1, 6, 21, 23, 29, 35
Discussion: Chapters
Discussion on Course

We’re moving into a new phase of the course…

Part III on Inferential Statistics.

Parts I and II were all foundational material for Part III.

Before we discussed …
Discussion on Course

Part I: Descriptive Statistics

Chapter 1: Statistics
  Background material. Definitions.

Chapter 2: Descriptive Analysis and Presentation
  single variable data
  Graphs, Central Tendency, Dispersion, Position

Chapter 3: Descriptive Analysis and Presentation
  bivariate data
  Scatter plot, Correlation, Regression
Discussion on Course

Part II: Probability

Chapter 4: Probability
  Conditional, Rules, Mutually Exclusive, Independent

Chapter 5: Probability Distributions (Discrete)
  Random variables, Probability Distributions, Mean & Variance, Binomial Distribution with Mean & Variance

Chapter 6: Probability Distributions (Continuous)
  Normal Distribution, Standard Normal, Applications, Notation

Chapter 7: Sample Variability
  Sampling Distributions, SDSM, CLT
Discussion on Course

Part III: Inferential Statistics
Chapter 8: Introduction to Statistical Inferences
   Hypothesis testing

Chapter 9: Inferences Involving One Population
   Mean $\mu$ ($\sigma$ unknown), proportion $p$, variance $\sigma^2$

Chapter 10: Inferences Involving Two Populations
   Difference in means $\mu_1-\mu_2$, proportions $p_1-p_2$, variances $\sigma_1^2 / \sigma_2^2$

Part IV: More Inferential Statistics
Chapter 11: Applications of Chi-Square
   Chi-square statistics. .... We will discuss later.
Discussion on Course

Next Lecture will be on Chapter 8.

Chapter 8: Introduction to Statistical Inferences
Hypothesis testing
Review Chapters 6 (Exam 4 Chapters)

Just the highlights!
The mathematical formula for the normal distribution is (p 315):

\[ f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \]

where
\[ e = 2.718281828459046... \]
\[ \pi = 3.141592653589793... \]
\[ \mu = \text{population mean} \]
\[ \sigma = \text{population std. deviation} \]

We will not use this formula.
Example: Here is a normal distribution with $\mu = 5$ and $\sigma^2 = 4$.

Let’s say we want to know the red area under the normal distribution between $x_1 = 3$ and $x_2 = 7$.

What is the area under the normal distribution between these two values?
6: Normal Probability Distributions

6.3 The Standard Normal Probability Distributions

Normal distribution $\mu=5$ & $\sigma^2=4$ and standard normal distribution.

Area between $x_1$ and $x_2$ is same as the area between $z_1$ and $z_2$.

If $x_1 = 3$ and $x_2 = 7$ then $z_1 = (x_1 - \mu)/\sigma = -1$ and $z_2 = (x_2 - \mu)/\sigma = 1$.

Rowe, D.B.
6: Normal Probability Distributions

6.3 The Standard Normal Probability Distributions

Now we can simply look up the $z$ areas in a table.

Appendix B Table 3 page 662.
# 6: Normal Probability Distributions

## Appendix B

### Table 3

Areas of the Standard Normal Distribution

The entries in this table are the probabilities that a random variable, with a standard normal distribution, assumes a value between 0 and \( z \); the probability is represented by the shaded area under the curve in the accompanying figure. Areas for negative values of \( z \) are obtained by symmetry.

<table>
<thead>
<tr>
<th>( z )</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0000</td>
<td>0.0040</td>
<td>0.0080</td>
<td>0.0120</td>
<td>0.0160</td>
<td>0.0199</td>
<td>0.0239</td>
<td>0.0279</td>
<td>0.0319</td>
<td>0.0359</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0398</td>
<td>0.0438</td>
<td>0.0478</td>
<td>0.0517</td>
<td>0.0557</td>
<td>0.0596</td>
<td>0.0635</td>
<td>0.0674</td>
<td>0.0714</td>
<td>0.0753</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0793</td>
<td>0.0832</td>
<td>0.0871</td>
<td>0.0910</td>
<td>0.0948</td>
<td>0.0987</td>
<td>0.1026</td>
<td>0.1064</td>
<td>0.1103</td>
<td>0.1141</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1179</td>
<td>0.1217</td>
<td>0.1255</td>
<td>0.1293</td>
<td>0.1331</td>
<td>0.1368</td>
<td>0.1406</td>
<td>0.1443</td>
<td>0.1480</td>
<td>0.1517</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1554</td>
<td>0.1591</td>
<td>0.1628</td>
<td>0.1664</td>
<td>0.1700</td>
<td>0.1736</td>
<td>0.1772</td>
<td>0.1808</td>
<td>0.1844</td>
<td>0.1879</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1915</td>
<td>0.1950</td>
<td>0.1985</td>
<td>0.2020</td>
<td>0.2054</td>
<td>0.2088</td>
<td>0.2123</td>
<td>0.2157</td>
<td>0.2190</td>
<td>0.2224</td>
</tr>
<tr>
<td>0.6</td>
<td>0.2257</td>
<td>0.2291</td>
<td>0.2324</td>
<td>0.2357</td>
<td>0.2389</td>
<td>0.2422</td>
<td>0.2454</td>
<td>0.2486</td>
<td>0.2517</td>
<td>0.2549</td>
</tr>
<tr>
<td>0.7</td>
<td>0.2580</td>
<td>0.2611</td>
<td>0.2642</td>
<td>0.2673</td>
<td>0.2704</td>
<td>0.2734</td>
<td>0.2764</td>
<td>0.2794</td>
<td>0.2823</td>
<td>0.2852</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2881</td>
<td>0.2910</td>
<td>0.2939</td>
<td>0.2967</td>
<td>0.2995</td>
<td>0.3023</td>
<td>0.3051</td>
<td>0.3078</td>
<td>0.3106</td>
<td>0.3133</td>
</tr>
<tr>
<td>0.9</td>
<td>0.3159</td>
<td>0.3186</td>
<td>0.3212</td>
<td>0.3238</td>
<td>0.3264</td>
<td>0.3289</td>
<td>0.3315</td>
<td>0.3340</td>
<td>0.3365</td>
<td>0.3389</td>
</tr>
<tr>
<td>1.0</td>
<td>0.3413</td>
<td>0.3438</td>
<td>0.3461</td>
<td>0.3485</td>
<td>0.3508</td>
<td>0.3531</td>
<td>0.3554</td>
<td>0.3577</td>
<td>0.3599</td>
<td>0.3621</td>
</tr>
<tr>
<td>1.1</td>
<td>0.3643</td>
<td>0.3665</td>
<td>0.3686</td>
<td>0.3708</td>
<td>0.3729</td>
<td>0.3749</td>
<td>0.3770</td>
<td>0.3790</td>
<td>0.3810</td>
<td>0.3830</td>
</tr>
<tr>
<td>1.2</td>
<td>0.3849</td>
<td>0.3869</td>
<td>0.3888</td>
<td>0.3907</td>
<td>0.3925</td>
<td>0.3944</td>
<td>0.3962</td>
<td>0.3980</td>
<td>0.3997</td>
<td>0.4015</td>
</tr>
<tr>
<td>1.3</td>
<td>0.4032</td>
<td>0.4049</td>
<td>0.4066</td>
<td>0.4082</td>
<td>0.4099</td>
<td>0.4115</td>
<td>0.4131</td>
<td>0.4147</td>
<td>0.4162</td>
<td>0.4177</td>
</tr>
<tr>
<td>1.4</td>
<td>0.4192</td>
<td>0.4207</td>
<td>0.4222</td>
<td>0.4236</td>
<td>0.4251</td>
<td>0.4265</td>
<td>0.4279</td>
<td>0.4292</td>
<td>0.4306</td>
<td>0.4319</td>
</tr>
<tr>
<td>1.5</td>
<td>0.4332</td>
<td>0.4345</td>
<td>0.4357</td>
<td>0.4370</td>
<td>0.4382</td>
<td>0.4394</td>
<td>0.4406</td>
<td>0.4418</td>
<td>0.4429</td>
<td>0.4441</td>
</tr>
<tr>
<td>1.6</td>
<td>0.4452</td>
<td>0.4463</td>
<td>0.4474</td>
<td>0.4484</td>
<td>0.4495</td>
<td>0.4505</td>
<td>0.4515</td>
<td>0.4525</td>
<td>0.4535</td>
<td>0.4545</td>
</tr>
<tr>
<td>1.7</td>
<td>0.4554</td>
<td>0.4564</td>
<td>0.4573</td>
<td>0.4582</td>
<td>0.4591</td>
<td>0.4599</td>
<td>0.4608</td>
<td>0.4616</td>
<td>0.4625</td>
<td>0.4633</td>
</tr>
<tr>
<td>1.8</td>
<td>0.4641</td>
<td>0.4649</td>
<td>0.4656</td>
<td>0.4664</td>
<td>0.4671</td>
<td>0.4678</td>
<td>0.4686</td>
<td>0.4693</td>
<td>0.4699</td>
<td>0.4706</td>
</tr>
<tr>
<td>1.9</td>
<td>0.4713</td>
<td>0.4719</td>
<td>0.4726</td>
<td>0.4732</td>
<td>0.4738</td>
<td>0.4744</td>
<td>0.4750</td>
<td>0.4756</td>
<td>0.4761</td>
<td>0.4767</td>
</tr>
<tr>
<td>2.0</td>
<td>0.4772</td>
<td>0.4778</td>
<td>0.4783</td>
<td>0.4788</td>
<td>0.4793</td>
<td>0.4798</td>
<td>0.4803</td>
<td>0.4808</td>
<td>0.4812</td>
<td>0.4817</td>
</tr>
<tr>
<td>2.1</td>
<td>0.4821</td>
<td>0.4826</td>
<td>0.4830</td>
<td>0.4834</td>
<td>0.4838</td>
<td>0.4842</td>
<td>0.4846</td>
<td>0.4850</td>
<td>0.4854</td>
<td>0.4857</td>
</tr>
<tr>
<td>2.2</td>
<td>0.4861</td>
<td>0.4864</td>
<td>0.4868</td>
<td>0.4871</td>
<td>0.4875</td>
<td>0.4878</td>
<td>0.4881</td>
<td>0.4884</td>
<td>0.4887</td>
<td>0.4890</td>
</tr>
<tr>
<td>2.3</td>
<td>0.4893</td>
<td>0.4896</td>
<td>0.4898</td>
<td>0.4901</td>
<td>0.4904</td>
<td>0.4906</td>
<td>0.4909</td>
<td>0.4911</td>
<td>0.4913</td>
<td>0.4916</td>
</tr>
<tr>
<td>2.4</td>
<td>0.4918</td>
<td>0.4920</td>
<td>0.4922</td>
<td>0.4925</td>
<td>0.4927</td>
<td>0.4929</td>
<td>0.4931</td>
<td>0.4932</td>
<td>0.4934</td>
<td>0.4936</td>
</tr>
</tbody>
</table>
6: Normal Probability Distributions

Appendix B, Table 3, Page 662

The entries in this table are the probabilities that a random variable, with a standard normal distribution, assumes a value between 0 and $z$; the probability is represented by the shaded area under the curve in the accompanying figure. Areas for negative values of $z$ are obtained by symmetry.

- Total area under curve is 1. \( f(z) \)
- The area less than 0 is 0.5.
- The area greater than 0 is .5.
- The area from 0 to $z$ is the same as the area from $-z$ to 0.
- The area greater than $z$ is 0.5 minus the area between 0 and $z$.
- The area less than $z$ is 0.5 plus the area between 0 and $z$. 
6: Normal Probability Distributions

6.3 The Standard Normal Probability Distributions

Now we can simply look up the $z$ areas in a table.

Appendix B Table 3 page 662.

Area between 0 and 1 is 0.3413.

Area between -1 and 1 is 0.6826
Assume that IQ scores are normally distributed with a mean $\mu$ of 100 and a standard deviation $\sigma$ of 16.

If a person is picked at random, what is the probability that his or her IQ is between 100 and 115? i.e. $P(100 < x < 115)$ ?

Figures from Johnson & Kuby, 2008.
6: Normal Probability Distributions

6.4 Applications of Normal Distributions

IQ scores normally distributed

\( \mu = 100 \) and \( \sigma = 16 \).

\[
P(100 < x < 115)
\]

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

\[
z_1 = \frac{x_1 - \mu}{\sigma} = \frac{100 - 100}{16} = 0
\]

\[
z_2 = \frac{x_2 - \mu}{\sigma} = \frac{115 - 100}{16} = 0.94
\]

Figures from Johnson & Kuby, 2008.
6: Normal Probability Distributions

6.4 Applications of Normal Distributions

IQ scores normally distributed \( \mu = 100 \) and \( \sigma = 16 \).

\[ P(100 < x < 115) \]

Now we can use the table.

\[ P(100 < x < 115) = P(0 < z < 0.94) \]

Figures from Johnson & Kuby, 2008.

Rowe, D.B.
6: Normal Probability Distributions
6.5 Notation

We can use the table in reverse.

**Example:**
Let $\alpha=0.05$. Find $P(z>z(0.05))=0.05$.

Same as finding $P(0<z<z(0.05))=0.45$.

Figures from Johnson & Kuby, 2008.
6: Normal Probability Distributions

6.5 Notation

Example:
Same as finding
\[ P(0 < z < z(0.05)) = 0.45. \]

Table shows this area (0.4500)

<table>
<thead>
<tr>
<th>z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0000</td>
<td>0.0040</td>
<td>0.0080</td>
<td>0.0120</td>
<td>0.0160</td>
<td>0.0199</td>
<td>0.0239</td>
<td>0.0279</td>
<td>0.0319</td>
<td>0.0359</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0398</td>
<td>0.0438</td>
<td>0.0478</td>
<td>0.0517</td>
<td>0.0557</td>
<td>0.0596</td>
<td>0.0636</td>
<td>0.0675</td>
<td>0.0714</td>
<td>0.0753</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0793</td>
<td>0.0832</td>
<td>0.0871</td>
<td>0.0910</td>
<td>0.0948</td>
<td>0.0987</td>
<td>0.1026</td>
<td>0.1064</td>
<td>0.1103</td>
<td>0.1141</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1179</td>
<td>0.1217</td>
<td>0.1255</td>
<td>0.1293</td>
<td>0.1331</td>
<td>0.1368</td>
<td>0.1406</td>
<td>0.1443</td>
<td>0.1480</td>
<td>0.1517</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1554</td>
<td>0.1591</td>
<td>0.1628</td>
<td>0.1664</td>
<td>0.1700</td>
<td>0.1736</td>
<td>0.1772</td>
<td>0.1808</td>
<td>0.1844</td>
<td>0.1879</td>
</tr>
</tbody>
</table>

... 

1.5 0.4332 0.4345 0.4357 0.4370 0.4382 0.4394 0.4406 0.4418 0.4429 0.4441
1.6 0.4452 0.4463 0.4474 0.4484 0.4495 0.4505 0.4515 0.4525 0.4535 0.4545
1.7 0.4554 0.4564 0.4573 0.4582 0.4591 0.4599 0.4608 0.4616 0.4625 0.4633
1.8 0.4641 0.4649 0.4656 0.4664 0.4671 0.4678 0.4686 0.4693 0.4699 0.4706
1.9 0.4713 0.4719 0.4726 0.4732 0.4738 0.4744 0.4750 0.4756 0.4761 0.4767

Figures from Johnson & Kuby, 2008.

Rowe, D.B.
6: Normal Probability Distributions
6.6 Normal Approximation of the Binomial Distribution

Approximate binomial probabilities with normal areas.
Use a normal with $\mu = np$, $\sigma^2 = np(1 - p)$

$\mu = (14)(.5) = 7$
$\sigma^2 = (14)(.5)(1 - .5) = 3.5$

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1 - p)^{n-x}$$
6: Normal Probability Distributions

6.6 Normal Approximation of the Binomial Distribution

From the binomial formula

\[ P(4) = \frac{14!}{4!(14-4)!} (0.5)^4 (1-0.5)^{14-4} \]

\[ P(x = 4) = 0.061 \]

\[ P(1.34 < z < 1.87) = 0.0594 \]

From the Normal Distribution

\[ P(3.5 < x < 4.5) \quad \mu = 7, \quad \sigma^2 = 3.5 \]

\[ z_1 = \frac{x_1 - \mu}{\sigma} = \frac{3.5 - 7}{\sqrt{3.5}} = -1.87 \]

\[ z_2 = \frac{x_2 - \mu}{\sigma} = \frac{4.5 - 7}{\sqrt{3.5}} = -1.34 \]

\[ n=14, \quad p=1/2 \]