Class 22

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
Agenda:

Recap Chapter 8.5

Lecture Chapter 9.1

Go over Exam 5
Recap Chapter 8.5
8: Introduction to Statistical Inference
8.4 Hypothesis Test of Mean (σ Known): Probability Approach

Step 1 The Set-Up: Null (H₀) and alternative (Hₐ) hypotheses
H₀: μ = 69" vs. Hₐ: μ ≠ 69"

Step 2 The Hypothesis Test Criteria: Test statistic.

\[ z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \]

σ known, n is “large” so by CLT \( \bar{x} \) is normal
\( z^* \) is normal

Step 3 The Sample Evidence: Calculate test statistic.

\[ z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{67.2 - 69}{4 / \sqrt{15}} = -1.74 \]

Step 4 The Probability Distribution:

\[ P(z > |z^*|) = p - value \rightarrow 0.0819 \]

Step 5 The Results:

\( p - value \leq \alpha \), reject \( H_0 \), \( p - value > \alpha \) fail to reject \( H_0 \)

\( \alpha = 0.05 \)
8: Introduction to Statistical Inference
8.5 Hypothesis Test of Mean (σ Known): Classical Approach

**Step 1 The Set-Up:** Null ($H_0$) and alternative ($H_a$) hypotheses

$H_0: \mu = 69$ vs. $H_a: \mu \neq 69$

**Step 2 The Hypothesis Test Criteria:** Test statistic.

$z^* = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

σ known, $n$ is “large” so by CLT $\bar{x}$ is normal

$z^*$ is normal

**Step 3 The Sample Evidence:** Calculate test statistic.

$z^* = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{67.2 - 69}{\frac{4}{\sqrt{15}}} = -1.74$

$n = 15$, $\bar{x} = 67.2$, $\sigma = 4$

**Step 4 The Probability Distribution:**

$\alpha = 0.05$, $z(\alpha/2) = 1.96$

**Step 5 The Results:**

$|z^*| > z(\alpha/2)$, **reject $H_0$**, $|z^*| \leq z(\alpha/2)$ **fail to reject $H_0$**
8: Introduction to Statistical Inference
8.5 Hypothesis Test of Mean ($\sigma$ Known): Classical Approach

There are three possible hypothesis pairs for the mean.

$$H_0: \mu \geq \mu_0 \text{ vs. } H_a: \mu < \mu_0$$

Reject $H_0$ if $$z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \text{ is less than } -z(\alpha)$$

data indicates $\mu < \mu_0$ because $\bar{x}$ is “a lot” smaller than $\mu_0$
8: Introduction to Statistical Inference
8.5 Hypothesis Test of Mean ($\sigma$ Known): Classical Approach

There are three possible hypothesis pairs for the mean.

$H_0: \mu \leq \mu_0$ vs. $H_a: \mu > \mu_0$

Reject $H_0$ if greater then

$$z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = z(\alpha)$$

data indicates $\mu > \mu_0$

because $\bar{x}$ is “a lot” larger than $\mu_0$
8: Introduction to Statistical Inference
8.5 Hypothesis Test of Mean ($\sigma$ Known): Classical Approach

There are three possible hypothesis pairs for the mean.

$$H_0: \mu = \mu_0 \text{ vs. } H_a: \mu \neq \mu_0$$

Reject $H_0$ if less than $-z(\alpha / 2)$

or if is greater than $z(\alpha / 2)$

$$z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

data indicates $\mu \neq \mu_0$, $\bar{x}$ far from $\mu_0$
Let’s examine the hypothesis test

\[ H_0: \mu \leq 69'' \quad \text{vs.} \quad H_a: \mu > 69'' \]

with \( \alpha = 0.05 \) for the heights of Math 1700 students.

Generate random data values.
Generated $15 \times 10^6$ normal data values from $\mu = 69$" and $\sigma = 4$".

Calculated $1 \times 10^6$ means with $n = 15$.
(Will repeat for $\mu = 72$".)
Hypothesis Test of Mean (\(\sigma\) Known): Classical Approach

\[ H_0: \mu \leq 69'' \] vs. \[ H_a: \mu > 69'' \]

\(\alpha = 0.05\)

When the true mean \(\mu = 69''\), we reject \(H_0\) \(\alpha\) fraction of the time.

Commit a Type I Error.

Given \(\alpha\), we want \(\mu_{\text{critical}}\).

\(n = 15\)

\(\bar{x}'s\)
Instead of $\mu_{\text{critical}}$ we find critical $z$, $z_{\text{critical}} = z(\alpha)$.

Do this by assuming that $H_0: \mu = 69$" is true, then calculate

$$z = \frac{\bar{x} - 69}{\frac{4}{\sqrt{15}}}$$
8: Introduction to Statistical Inference
8.5 Hypothesis Test of Mean (σ Known): Classical Approach

\[ H_0: \mu \leq 69 \text{ vs. } H_a: \mu > 69 \]
\[ \alpha = 0.05 \]

When the true mean \( \mu = 72 \), we do not reject \( H_0 \) \( \beta \) fraction of the time.

Commit a Type II Error

Rowe, D.B.
8: Introduction to Statistical Inference

8.5 Hypothesis Test of Mean ($\sigma$ Known): Classical Approach

$H_0: \mu \leq \mu_0$ vs. $H_a: \mu > \mu_0$

$\bar{x}'s$

$n=15$

$1 \times 10^6$

$\bar{x}'s$

Fail to Reject

Reject

$H_0$

$\bar{x}'s$

$\mu$ critical

1-\( \alpha \)

1-\( \beta \)

\( \alpha \)

\( \beta \)

Power of the test:

$1 - \beta = P(\text{Reject } H_0 | H_0 \text{ False})$

Discrimination ability. Ability to detect difference.

Rowe, D.B.
8: Introduction to Statistical Inference

8.5 Hypothesis Test of Mean (σ Known): Classical Approach

\[ H_0: \mu \leq \mu_0 \text{ vs. } H_a: \mu > \mu_0 \]

\[ n = 15 \]
\[ \bar{x} = 1 \times 10^6 \]

\[ z = \frac{\bar{x} - 69}{4 / \sqrt{15}} \]

Rowe, D.B.

\[ P(\text{Reject } H_0 | H_0 \text{ False}) = 1 - \beta \]

Discrimination ability.

Ability to detect difference.

\[ 1 - \beta \]

\[ P(\text{Reject } H_0 | H_0 \text{ False}) \]
8: Introduction to Statistical Inference

8.5 Hypothesis Test of Mean (σ Known): Classical Approach

We want our \( \alpha \), Prob of Type I Error to be small.

So why not just decrease \( \alpha \)?

Decreasing \( \alpha \) increases \( \beta \).

And vice versa.
What is the solution?

Increase $n$.

Figure shows $n$ increased to $n = 30$ from $n = 15$.

Note $\alpha$ and $\beta$ both smaller with larger $n$. 

Rowe, D.B.
What is the solution?

Increase $n$.

Figure shows $n$ increased to $n = 30$ from $n = 15$.

Note $\alpha$ and $\beta$ both smaller with larger $n$. 

Rowe, D.B.
Chapter 8: Introduction to Statistical Inference

Questions?

Homework: Chapter 8 # 5, 15, 19, 22, 23, 24, 25, 35, 47, 51, 57, 59, 81, 87, 91, 93, 97, 105, 106, 107, 109, 119, 139, 140, 145, 149, 157, 159
Lecture Chapter 9.1
Chapter 9: Inferences Involving One Population

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
9: Inferences Involving One Population

9.2 Inference about the Mean $\mu$ ($\sigma$ Unknown)

In Chapter 8, we performed hypothesis tests on the mean by

1) assuming that $\bar{x}$ was normally distributed ($n$ “large”),

2) assuming the hypothesized mean $\mu_0$ were true,

3) assuming that $\sigma$ was known, so that we could form

$$z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

which with 1) – 3) has standard normal dist.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

However, in real life, we never know $\sigma$ for

$$z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

so we would like to estimate $\sigma$ by $s$, then use

$$t^* = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}.$$  

But $t^*$ does not have a standard normal distribution.

It has what is called a Student $t$-distribution.

---

Rowe, D.B.
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

What is the Student $t$-distribution and how do we get it?

Background Information

If the data comes from normally distributed population, then

$$x \sim N(\mu, \sigma^2)$$

mean variance

$$\bar{x} \sim N(\mu, \sigma^2 / n)$$

mean variance

generate

$\mu = 100$

$\sigma = 57.7$

$n = 5$

generate $5 \times 10^6$

random values

sample means from each 5

$n=5$

$5 \times 10^6$

$n=5$

$1 \times 10^6$
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

If we know the true population mean $\mu$, then

$$\bar{x} - \mu \sim N(0, \sigma^2 / n)$$

Given the variance of the mean $\sigma^2/n$, the distribution of

$$\frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

Rowe, D.B.
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

It turns out that with the variance $\sigma^2$ known, the distribution of

\[
\frac{(n-1)s^2}{\sigma^2}
\]

has a chi-square distribution with $n-1$ degrees of freedom.

$(\chi^2$ distribution on Pages 453-454)
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

The ratio
\[ t = \left( \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \right) \sqrt{\frac{(n-1)s^2}{\sigma^2}} \left( n-1 \right) \]
with simplification
\[ t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \]

is
\[ t = z \sqrt{\frac{\chi^2}{df}} \]
and has a Student $t$-distribution with $n-1$ df.

\[
\begin{align*}
\mu &= 100 \\
\sigma &= 57.7 \\
n &= 5
\end{align*}
\]

Rowe, D.B.
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

Student $t$-distribution has heavier tails than standard normal.

\[ t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \]

\[ z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \]

$\mu = 100$

$\sigma = 57.7$

$n = 5$

Rowe, D.B.
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ (\(\sigma\) Unknown)

The standard normal dist is:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

The Student-t distribution is:

$$f(t \mid \nu) = \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \sqrt{\nu\pi} \left(1 + \frac{1}{\nu \, t^2}\right)^{-\frac{\nu+1}{2}}$$

$$\nu = df = n - 1$$

Rowe, D.B.
9: Inferences Involving One Population
9.1 Inference about the Mean \( \mu \) (\( \sigma \) Unknown)

The standard normal dist. is:

\[
f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}
\]

The Student-t distribution is:

\[
f(t | \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu \pi}} \frac{1}{\left(1 + \frac{1}{\nu} t^2\right)^{\frac{\nu+1}{2}}}
\]

as \( n \) increases, so does \( \nu = df = n-1 \) and Student \( t \) becomes standard normal

Rowe, D.B.
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

The standard normal dist. is:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

The Student-t distribution is:

$$f(t \mid \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \frac{1}{\left(1 + \frac{1}{\nu}t^2\right)^{\frac{\nu+1}{2}}}$$

as $n$ increases, so does $\nu=\text{df}=n-1$ and Student $t$ becomes standard normal

Rowe, D.B.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)
Using the $t$-Distribution Table

Finding critical value from a Student $t$-distribution, $df=n-1$

$t(df,\alpha)$, $t$ value with $\alpha$ area larger than it

with $df$ degrees
of freedom

Table 6
Appendix B
Page 719.

Figure from Johnson & Kuby, 2012.
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

**Example:** Find the value of $t(10,0.05)$, $df=10$, $\alpha=0.05$.

Area in One Tail

<table>
<thead>
<tr>
<th>Area in Two Tails</th>
<th>0.25</th>
<th>0.10</th>
<th>0.05</th>
<th>0.025</th>
<th>0.01</th>
<th>0.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
<td>0.50</td>
<td>0.20</td>
<td>0.10</td>
<td>0.05</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>3</td>
<td>0.765</td>
<td>1.64</td>
<td>2.35</td>
<td>3.18</td>
<td>4.54</td>
<td>5.84</td>
</tr>
<tr>
<td>4</td>
<td>0.741</td>
<td>1.53</td>
<td>2.13</td>
<td>2.78</td>
<td>3.75</td>
<td>4.60</td>
</tr>
<tr>
<td>5</td>
<td>0.727</td>
<td>1.48</td>
<td>2.02</td>
<td>2.57</td>
<td>3.36</td>
<td>4.03</td>
</tr>
<tr>
<td>6</td>
<td>0.718</td>
<td>1.44</td>
<td>1.94</td>
<td>2.45</td>
<td>3.14</td>
<td>3.71</td>
</tr>
<tr>
<td>7</td>
<td>0.711</td>
<td>1.41</td>
<td>1.89</td>
<td>2.36</td>
<td>3.00</td>
<td>3.50</td>
</tr>
<tr>
<td>8</td>
<td>0.706</td>
<td>1.40</td>
<td>1.86</td>
<td>2.31</td>
<td>2.90</td>
<td>3.36</td>
</tr>
<tr>
<td>9</td>
<td>0.703</td>
<td>1.38</td>
<td>1.83</td>
<td>2.26</td>
<td>2.82</td>
<td>3.25</td>
</tr>
<tr>
<td>10</td>
<td>0.700</td>
<td>1.37</td>
<td>1.81</td>
<td>2.23</td>
<td>2.76</td>
<td>3.17</td>
</tr>
</tbody>
</table>

Table 6: Appendix B Page 719.

Go to 0.05 One Tail column and down to 10 $df$ row.

Figures from Johnson & Kuby, 2012.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

Confidence Interval Procedure

When making a confidence interval for $\mu$ when $\sigma$ unknown, we assume that the population is normal, not just mean, but when $n$ is “large,” can use for nonnormal distributions.

The assumption for inferences about the mean $\mu$ when $\sigma$ is unknown: The sampled population is normally distributed.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ (\(\sigma\) Unknown)

Confidence Interval Procedure

Discussed a confidence interval for the $\mu$ when $\sigma$ was known,

Confidence Interval for Mean:

$$\bar{x} - z(\alpha / 2) \frac{\sigma}{\sqrt{n}} \text{ to } \bar{x} + z(\alpha / 2) \frac{\sigma}{\sqrt{n}}$$

(8.1)

now, with sigma unknown, the CI for the mean is

Confidence Interval for Mean:

$$\bar{x} - t(df,\alpha / 2) \frac{s}{\sqrt{n}} \text{ to } \bar{x} + t(df,\alpha / 2) \frac{s}{\sqrt{n}}$$

(9.1)
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS

Step 1 The Set-Up:

Step 2 The Hypothesis Test Criteria:

Step 3 The Sample Evidence:

Step 4 The Probability Distribution:

Step 5 The Results:
THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS

Step 1 The Set-Up:

a. Describe the population parameter of interest.

The population parameter of interest is the mean $\mu$, the height of Math 1700 students.
THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS

Step 1 The Set-Up:
- b. State the null hypothesis ($H_0$) and the alternative hypotheses ($H_a$).

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Alternative Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Greater than or equal to ($\geq$)</td>
<td>Less than ($&lt;$)</td>
</tr>
<tr>
<td>2. Less than or equal to ($\leq$)</td>
<td>Greater than ($&gt;$)</td>
</tr>
<tr>
<td>3. Equal to ($=$)</td>
<td>Not equal to ($\neq$)</td>
</tr>
</tbody>
</table>

$H_0: \mu = 69''$ vs. $H_a: \mu \neq 69''$

Figure from Johnson & Kuby, 2012.
THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS

Step 2 The Hypothesis Test Criteria:

a. Check the assumptions.
Assume from past experience we know population is normal.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS
Step 2 The Hypothesis Test Criteria:
   b. Identify the probability distribution and the test statistic to be used.

$\sigma$ is unknown, therefore $t$-distribution with $n-1$ df.

Test Statistic for Mean:

$$t^* = \frac{\overline{x} - \mu_0}{s / \sqrt{n}}$$

where $\sigma$ is assumed to be unknown  \hspace{1cm} (9.2)
THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS

Step 3 The Sample Evidence:

a. Collect a sample of information.
   Take a random sample from a population with mean \( \mu \) that being questioned.

b. Calculate the value of the test statistic.

\[
t^* = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{67.2 - 69}{3.5 / \sqrt{15}} = -1.99
\]

Assuming \( n=15 \) , \( \bar{x} = 67.2 \) , and \( s = 3.5 \).
9: Inferences Involving One Population
9.1 Inference about the Mean \( \mu \) (\( \sigma \) Unknown)

THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS
Step 4 The Probability Distribution:
   a. Determine the critical region and critical value(s).

Critical Region: The set of values for the test statistic that will cause us to reject the null hypothesis.

Critical value(s): The “first” or “boundary” value(s) of the critical region(s).
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

**THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS**

**Step 4 The Probability Distribution:**

a. Determine the critical region and critical value(s).

b. Determine whether or not the calculated test statistic is in the critical region.

$H_0: \mu = 69''$ vs. $H_a: \mu \neq 69''$

$P(t > t(df, \alpha / 2)) = \alpha / 2$

$t(14,.025) = 2.14$,

since two sided test.

$\sigma$ known was 1.96

Rowe, D.B.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS

Step 5 The Results:

a. State the decision about $H_0$.
   Need a decision rule.

Decision rule:

a. If the test statistic falls within the critical region, then the decision must be reject $H_0$.

b. If the test statistic is not in the critical region, then the decision must be fail to reject $H_0$. 
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

THE PROBABILITY-VALUE HYPOTHESIS TEST: 5 STEPS
Step 5 The Results:
b. State the conclusion about $H_a$.
With $\alpha = 0.05$,

there is not sufficient evidence to reject $H_0$.

Fail to reject $H_0$.

Figure adapted from Johnson & Kuby, 2012.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

Recap 9.1:
Essentially have new critical value, $t(df,\alpha)$ to look up in a table when $\sigma$ is unknown. Used same way as before.

For $\sigma$ assumed known:
\[
\bar{x} \pm z(\alpha / 2) \frac{\sigma}{\sqrt{n}}
\]
\[
z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}
\]

For $\sigma$ assumed unknown:
\[
\bar{x} \pm t(df,\alpha / 2) \frac{s}{\sqrt{n}}
\]
\[
t^* = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}
\]
Chapter 9: Inferences Involving One Population

Questions?

Homework: Chapter 9 # 5, 7, 9, 21, 23, 27, 28, 35, 37, 43, 49, 55, 71, 85, 89, 91, 93, 95, 97, 99, 105, 109, 119, 121, 129, 131, 135, 139, 145
Go over Exam 5