ME 498 Manufacturing Data and Quality Systems

Inferences about Process Quality II
Recap: confidence interval

\[(1) \bar{X} - Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} \]

\[(2) \sigma^2 \geq \frac{(n-1)S^2}{\chi^2_{\alpha,n-1}} \]

\[(3) \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha,n-1}} \]

\[(4) \mu \geq \bar{X} - t_{\alpha,n-1} \times \frac{S}{\sqrt{n}} \]

\[(5) \mu \leq \bar{X} + t_{\alpha,n-1} \times \frac{S}{\sqrt{n}} \]

\[(6) \bar{X} - t_{\alpha/2,n-1} \times \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2,n-1} \times \frac{S}{\sqrt{n}} \]

\[(7) \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}} \]

\[(8) \mu \geq \bar{X} - Z_{\alpha} \times \frac{\sigma}{\sqrt{n}} \]

\[(9) \mu \leq \bar{X} + Z_{\alpha} \times \frac{\sigma}{\sqrt{n}} \]
Hypothesis testing

• I am interested in whether the grade difference between undergraduate students and graduate students is statistically significant.

• Data collection: grading homework, projects, exams, etc.

Fall 2018 data

<table>
<thead>
<tr>
<th>Undergraduate</th>
<th>94.43</th>
<th>93.16</th>
<th>92.02</th>
<th>91.08</th>
<th>90.58</th>
<th>88.41</th>
<th>87.56</th>
<th>84.83</th>
<th>84.73</th>
<th>83.26</th>
<th>82.06</th>
<th>80.28</th>
<th>78.07</th>
<th>65.57</th>
<th>63.52</th>
<th>27.40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduate</td>
<td>98.22</td>
<td>97.21</td>
<td>96.83</td>
<td>96.72</td>
<td>96.06</td>
<td>95.96</td>
<td>95.38</td>
<td>94.82</td>
<td>94.81</td>
<td>93.61</td>
<td>92.98</td>
<td>92.98</td>
<td>92.96</td>
<td>92.71</td>
<td>90.43</td>
<td>89.42</td>
</tr>
</tbody>
</table>

• Undergraduate: average = 83.96; std = 16.73
• Graduate: average = 94.44; std = 2.45
• Conclusion?
Design:
1) State the null and alternative hypotheses, and specify sample size $n$ and significance level $\alpha$.
2) Define the test statistic.
3) Find the distribution of the test statistic and the rejection region of $H_0$.

**Example:** Test if the mean of a Normal distribution equals to $1.5$ with $\sigma^2 = 4$.

\[
H_0 : \mu = 1.5 \\
H_1 : \mu \neq 1.5 \\
\]
\[
n = 5 \\
\alpha = 0.05 \\
\]
\[
Z_0 = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \\
Z_0 \sim N(0,1) \\
|Z_0| > Z_{\alpha/2} = Z_{0.025} = 1.96
\]

Test Statistic  Rejection Region
Hypothesis testing

Use:
4) Collect sample data and calculate the test statistic using the sample.
5) Compare the test statistic of the sample with the rejection region.
6) Make the decision and assess the risk.

Example (cont’d): Test on the mean of a Normal distribution with $\sigma^2 = 4$

\[
X_1 = 0.5 \\
X_2 = 1.5 \\
X_3 = 1 \quad \rightarrow \quad \bar{X} = 1.3 \quad \rightarrow \quad Z_0 = \frac{1.3 - 1.5}{2/\sqrt{5}} = -0.22 \quad \quad |Z_0| \neq Z_{\alpha/2} = 1.96
\]
\[
X_4 = 1.5 \\
X_5 = 2
\]

$H_0$ cannot be rejected!
Court system and hypothesis testing

Hypothesis testing in science is very much like the criminal court system in the United States. How do we decide guilt?

- Assume innocence
  - $H_0$: assumed parameter
- Evidence is presented at a trial
  - Calculate testing statistics based on sampling data
- Decide if guilty is “proven”, i.e., require strong evidence to “prove” guilty
  - Reject $H_0$ if the testing statistic is beyond reasonable doubt, i.e., falling in its rejection region

A jury’s possible decision:

- Guilty (reject $H_0$)
- Not guilty (cannot reject $H_0$) but it is wrong to say “Accept $H_0$”

Note that a jury cannot declare somebody “innocent”, A jury can only say “not guilty.”

Do you see the difference?