Inference on One Population Mean
– Hypothesis Testing

Scenario 1. When the population is normal, and the population variance is known

Data: \( X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2) \)

Hypothesis test, for instance:
\[
\begin{align*}
H_0 &: \mu = \mu_0 \\
H_a &: \mu > \mu_0
\end{align*}
\]

Example:
\( H_0 : \mu \leq 57'' \) (null hypothesis): This is the ‘original belief’
\( H_a : \mu > 57'' \) (alternative hypothesis): This is usually your hypothesis (i.e. what you believe is true) if you are conducting the test – and in general, should be supported by your data.

The statistical hypothesis test is very similar to a law suit:
e.g) The famous O.J. Simpson trial
\[ H_0 : OJ \text{ is innocent (‘innocent unless proven guilty)} \]
\[ H_a : OJ \text{ is guilty (‘supported by the data: the evidence)} \]

<table>
<thead>
<tr>
<th>Jury’s Decision</th>
<th>The truth</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(H_0)</td>
<td>(H_0: OJ \text{ innocent})</td>
<td>Right decision</td>
</tr>
<tr>
<td>(H_a)</td>
<td>(H_a: OJ \text{ guilty})</td>
<td>Type II error</td>
</tr>
<tr>
<td></td>
<td>(H_a: OJ \text{ guilty})</td>
<td>Type I error</td>
</tr>
<tr>
<td></td>
<td>(H_a: OJ \text{ innocent})</td>
<td>Right decision</td>
</tr>
</tbody>
</table>
The **significance level** and three types of hypotheses.

\[ P(\text{Type I error}) = \alpha \quad \text{← significance level of a test (Type I error rate)} \]

1. \( H_0 : \mu = \mu_0 \quad \Leftrightarrow \quad H_0 : \mu \leq \mu_0 \)
   \[ H_a : \mu > \mu_0 \quad H_a : \mu > \mu_0 \]

2. \( H_0 : \mu = \mu_0 \quad \Leftrightarrow \quad H_0 : \mu \geq \mu_0 \)
   \[ H_a : \mu < \mu_0 \quad H_a : \mu < \mu_0 \]

3. \( H_0 : \mu = \mu_0 \quad H_a : \mu \neq \mu_0 \)

*Now we derive the hypothesis test for the first pair of hypotheses.*

\( H_0 : \mu = \mu_0 \)

\( H_a : \mu > \mu_0 \)

Data: \( X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2), \sigma^2 \) is known and the given a significance level \( \alpha \) (say, 0.05).

Let’s derive the test. (That is, derive the decision rule)

**Two approaches (**equivalent**) to derive the tests:**

- Likelihood Ratio Test
- Pivotal Quantity Method

***Now we will first demonstrate the Pivotal Quantity Method.***

1. We have already derived the PQ when we derived the C.I. for \( \mu \)

\[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ is our P.Q.} \]

2. The test statistic is the PQ with the value of the parameter of interest under the null hypothesis (\( H_0 \)) inserted:

\[ Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \text{ is our test statistic.} \]

That is, given \( H_0 : \mu = \mu_0 \) in true \( \Rightarrow Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \)
3. * Derive the decision threshold for your test based on the **Type I error rate** the significance level \( \alpha \)

(1) For the pair of hypotheses:

- \( H_0 : \mu = \mu_0 \)

- **Versus**

- \( H_a : \mu > \mu_0 \)

It is intuitive that one should reject the null hypothesis, in support of the alternative hypothesis, when the sample mean is larger than \( \mu_0 \). Equivalently, this means when the test statistic \( Z_0 \) is larger than certain positive value \( c \) - the question is what is the exact value of \( c \) -- and that can be determined based on the significance level \( \alpha \) — that is, how much Type I error we would allow ourselves to commit.

Setting:

\[
P(\text{Type I error}) = P(\text{reject } H_0 \mid H_0) = P(Z_0 \geq c \mid H_0 : \mu = \mu_0) = \alpha
\]

We will see immediately that

\[
c = z_\alpha
\]

from the pdf plot of the test statistic below.

\[
\vdash \text{At the significance level } \alpha, \text{ we will reject } H_0 \text{ in favor of } H_a \text{ if } Z_0 \geq Z_\alpha
\]
Other Hypotheses

(2)

\( H_0 : \mu = \mu_0 \) (one-sided test or one-tailed test)

\( H_a : \mu < \mu_0 \)

Test statistic (same): \( Z_0 = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1) \)

\( \alpha = P(Z_0 \leq c \mid H_0 : \mu = \mu_0) \Rightarrow c = -Z_{\alpha} \)

(3)

\( H_0 : \mu = \mu_0 \) (Two-sided or Two-tailed test)

\( H_a : \mu \neq \mu_0 \)

Test statistic (same): \( Z_0 = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1) \)

\( \alpha = P(\mid Z_0 \mid \geq c \mid H_0) = P(Z_0 \geq c \mid H_0) + P(Z_0 \leq -c \mid H_0) \)

\[ = 2 \cdot P(Z_0 \geq c \mid H_0) \]

\[ \frac{\alpha}{2} = P(Z_0 \geq c \mid H_0) \]

\[ \therefore c = Z_{\alpha/2} \]
Reject $H_0$ if $|Z_0| \geq \frac{Z_{\alpha/2}}{\sqrt{n}}$

4. We have just discussed the "rejection region" approach for decision making. There is another approach for decision making, it is "p-value" approach.

*Definition: p-value – it is the probability that we observe a test statistic value that is as extreme, or more extreme, than the one we observed, given that the null hypothesis is true.

\[
\begin{array}{|c|c|c|c|}
\hline
H_0 : \mu = \mu_0 & H_0 : \mu = \mu_0 & H_0 : \mu = \mu_0 \\
H_\alpha : \mu > \mu_0 & H_\alpha : \mu < \mu_0 & H_\alpha : \mu \neq \mu_0 \\
\hline
\end{array}
\]

Observed value of test statistic
\[
Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)
\]

p-value = \(P(Z_0 \geq z_0 \mid H_0)\)  
\hspace{1cm} p-value = \(P(Z_0 \leq z_0 \mid H_0)\)  
\hspace{1cm} p-value  
\hspace{1cm} = \(P(|Z_0| \geq |z_0| \mid H_0)\)  
\hspace{1cm} = 2 \cdot P(Z_0 \geq z_0 \mid H_0) 

(1) the area under \(N(0,1)\) pdf to the right of \(z_0\)  
(2) the area under \(N(0,1)\) pdf to the left of \(z_0\)  
(3) twice the area to the right of \(|z_0|\)
(1)
\[ H_0 : \mu = \mu_0 \]
\[ H_a : \mu > \mu_0 \]

(2)
\[ H_0 : \mu = \mu_0 \]
\[ H_a : \mu < \mu_0 \]
(3)

\[ H_0 : \mu = \mu_0 \]
\[ H_a : \mu \neq \mu_0 \]

(*That is, the p-value is the sum of the two tail areas, or equivalently, the p-value is twice the upper tail area.)

*** The way we make conclusions for the tests based on the p-value is the same for all pairs of hypotheses: **We reject \( H_0 \) in favor of \( H_a \) iff p-value < \( \alpha \)***
**Scenario 2.** The large sample scenario: Any population (*usually non-normal* – as the exact tests should be used if the population is normal), however, the sample size is large *(this usually refers to: \( n \geq 30 \))

**Theorem. The Central Limit Theorem.** Let \( X_1, X_2, \ldots, X_n \) be a random sample from a population with mean \( \mu \) and variance \( \sigma^2 \), we have \( Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \to N(0,1) \).

* Thus when \( n \) is large enough \( (n \geq 30) \), \( Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \) - by CLT.

* When \( \sigma \) is unknown, \( (n \geq 30) \), \( Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim N(0,1) \) – by CLT and the Slutsky’s Theorem

Therefore the pivotal quantities *(P.Q.’s)* for this scenario are:

\[
Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ or } Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim N(0,1)
\]

We use the first P.Q. if \( \sigma \) is known, and the second when \( \sigma \) is unknown.

The derivation of the hypothesis tests *(rejection region and the p-value)* are almost the same as the derivation of the exact Z-test discussed above.

<table>
<thead>
<tr>
<th>( H_0 ): ( \mu = \mu_0 )</th>
<th>( H_0 ): ( \mu = \mu_0 )</th>
<th>( H_0 ): ( \mu = \mu_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_a ): ( \mu &gt; \mu_0 )</td>
<td>( H_a ): ( \mu &lt; \mu_0 )</td>
<td>( H_a ): ( \mu \neq \mu_0 )</td>
</tr>
</tbody>
</table>

Test Statistic \( Z_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim N(0,1) \)

Rejection region: we reject \( H_0 \) in favor of \( H_a \) at the significance level \( \alpha \) if

| \( Z_0 \geq Z_{\alpha} \) | \( Z_0 \leq -Z_{\alpha} \) | \( |Z_0| \geq Z_{\alpha/2} \) |
|------------------------|------------------------|------------------------|

p-value = \( P(Z_0 \geq z_0 \mid H_0) \)

p-value = \( P(Z_0 \leq z_0 \mid H_0) \)

p-value = \( P(|Z_0| \geq z_0 \mid H_0) \) = \( 2 \cdot P(Z_0 \geq z_0 \mid H_0) \)

(1) the area under \( N(0,1) \) pdf to the right of \( z_0 \)

(2) the area under \( N(0,1) \) pdf to the left of \( z_0 \)

(3) twice the area to the right of \( |z_0| \)
**Scenario 3. Normal Population, but the population variance is unknown**

100 years ago – people use Z-test

This is OK for n large \( n \geq 30 \) ⇒ per the CLT (Scenario 2)

This is NOT ok if the sample size is small.

“A Student of Statistics”
– pen name of William Sealy Gosset (June 13, 1876–October 16, 1937)

*The Student’s t-test*

\[
P.Q. \quad T = \frac{X - \mu}{S / \sqrt{n}} \sim t_{n-1}
\]

(Exact t-distribution with n-1 degrees of freedom)

**Take-home Quiz due Tuesday, 4/21/2014**

Using Scenario 1 (test for one population mean \( \mu \), normal population, population variance is known), please discuss and derive the relationship between the two-sided test at the significance level \( \alpha \), and the corresponding 100(1 – \( \alpha \))% confidence interval for \( \mu \).

**Homework/Review Questions**

On **Confidence Intervals** for **One Population Mean** or **Proportion**

(* Solutions will be given soon – please try by yourselves first. These are excellent candidates for quizzes & final exam. Read Text Book §5.4, §5.5) 

5.4.3, 5.4.4, 5.4.9, 5.5.1, 5.5.2, 5.5.3, 5.5.4, 5.5.5, 5.5.6, 5.5.11, 5.5.15, 5.5.21, 5.5.22, 5.5.30, 5.5.32, 5.5.41