Hypothesis Test-Confidence Interval connection

Hypothesis tests for mean
Tell whether observed data are consistent with $\mu = \mu_0$.

More specifically...
An hypothesis test with significance level $\alpha$ will reject the hypothesis $\mu = \mu_0$ if and only if
$$P(\text{observed data} \mid \mu = \mu_0) < \alpha.$$ 

Confidence Intervals for mean
Give a set of values for true mean, $\mu$, that are not inconsistent with the observed data.

More specifically...
If $\mu_0$ is a value outside the $(1-\alpha)%$ confidence interval, then
$$P(\text{observed data} \mid \mu = \mu_0) < \alpha.$$ 

From the above we can infer...

If $\mu_0$ is a value outside the $(1-\alpha)%$ confidence interval, then an hypothesis test with significance level $\alpha$ will reject the null hypothesis $H_0: \mu = \mu_0$.

Similarly, if $\mu_0$ is inside the confidence interval then $H_0: \mu = \mu_0$ will not be rejected.
Example 1: Facial Height Data

The average change in facial height in the sample was -1.7 mm, with $s = 6.8$ mm. The null hypothesis, $H_0: \mu = 0$, was rejected at the $\alpha = 0.05$ level.

The 95% confidence interval for the mean change in facial height is

$$-1.7 \pm 1.99 \cdot 6.8/\sqrt{84} = -1.7 \pm 1.5 = (-3.2, -0.2),$$

which does not contain 0.

Example 2: Chewing Gum Data

The average two-year change in DMFS for those in group A was -0.72 DMFS. The 95% confidence interval was (-2.92, 1.48), which contains 0.

The null hypothesis $H_0: \mu = 0$ was not rejected ($p = 0.51$).

Note that the 95% confidence interval shows that neither a true average decrease of more than 2 DMFS nor an increase of more than 1 DMFS would be inconsistent with these data.

Thus, not rejecting a null hypothesis does not necessarily indicate that we should accept the null hypothesis as the truth!
Hypothesis Test for Proportions:

The one-sample test of proportions is based on the statistic

\[
Z = \frac{\hat{p} - p_0}{\sqrt{p_0 \cdot (1 - p_0) / n}}.
\]

- If \(|Z|\) is large (\(Z\) is far from zero), it suggests that \(p \neq p_0\).

- If \(p = p_0\) and \(np(1-p) \geq 5\), then \(Z\) will have an approximate \(N(0,1)\) distribution.

- If \(Z\) is larger value than one would expect to come from a \(N(0,1)\) distribution, then we will have evidence that \(p \neq p_0\).
Example: Presidential Election Poll

Presidential preference poll results
(7/5/2016)*

<table>
<thead>
<tr>
<th></th>
<th>Trump</th>
<th>Clinton</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>400</td>
<td>380</td>
<td>780</td>
</tr>
</tbody>
</table>

*among those who favored either Trump or Clinton

Do these data provide evidence that Trump has a lead in the population sampled?

To answer this question we can compute the hypothesis test

\[ H_0: p \leq 0.50, \text{ and } H_1: p > 0.50, \]

where \( p \) is the proportion preferring Trump.

- Note that the form of \( H_0 \) differs from previously introduced tests.

- With these hypotheses, a rejection of \( H_0 \) will indicate evidence that Trump has a lead.

- If we used \( H_0: p = 0.50, \text{ and } H_1: p \neq 0.50, \) a rejection of \( H_0 \) would indicate only that the candidates are not tied.

- This is an example of a “one-sided” hypothesis test.
Example: Presidential Election poll ...continued

To test the one-sided hypotheses

\[ H_0: p \leq 0.50, \text{ and } H_1: p > 0.50, \]

Compute the Z statistic:

\[ \hat{p} = \frac{400}{780} = 0.513 \]

\[ Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} = \frac{0.513 - 0.50}{\sqrt{0.50 \cdot 0.50/780}} = 0.726. \]

Note that for \( p_0 \) we use the “border” probability, 0.50. This would be the “worst case” in terms of disproving \( H_0 \).

For one-sided tests the p-value is also one-sided. For this example the p-value is:

\[ \text{p-value} = P( N(0,1) > 0.726 ) = 0.234 \]

- The probability is on only one side of the distribution.

- The side it is on should be the side that indicates that \( H_0 \) is not true.
For this example, let $\alpha = 0.05$, then since the p-value $= 0.234$ is greater than 0.05 we do not reject $H_0$.

Suppose we cannot compute the p-value easily we can also compare the statistic to a critical value to see if the p-value is less than or greater than $\alpha$.

In the case of a one sided test at significance level $\alpha$:

- If $H_1$: $p > p_0$, then reject $H_0$ if $Z > Z_{1-\alpha}$
- If $H_1$: $p < p_0$, then reject $H_0$ if $Z < Z_{\alpha}$

Note that we do not halve the $\alpha$ for one-sided tests.

So for this example we would compare $Z = 0.726$ to $Z_{0.95} = 1.645$.

Since $0.726 < 1.645$, we do not reject $H_0$. 
Final notes on the Presidential Poll Example:

- Though we did not reject $H_0$, this does NOT necessarily indicate that $H_1$ is true.

- The 95% confidence interval for $p$ is (47.8%, 54.8%), which indicates that there is NOT strong evidence of either candidate having the majority.

- Tests for proportions do not have to be one-sided. Two-sided tests for proportions can also be formulated.

- One-sided tests can be used for comparisons other than proportions (e.g. one can use them with t-tests).
One-Sample statistics summary

- **Binary data** (a.k.a. dichotomous, or “yes/no” data)
  - Large sample, $np(1-p) > 5$
    - Confidence interval for proportion, §8.3
    - One-sample test for proportion, §9.3
  - Small sample, $np(1-p) \leq 5$
    - Exact Binomial confidence intervals, HW #6
    - Exact one-sample Binomial Test*

- **Continuous data** (a.k.a. Interval data)
  - Large sample
    - Confidence interval for mean §8.2
    - One-sample t test §9.2
  - Small sample
    - data from Normal distribution
      - Confidence interval for mean §8.2
      - One-sample t test §9.2
    - non-Normal data
      - Exact confidence interval for median*
      - One-sample test for median (similar to sign test §13.1)

*not covered in course topics
One-sample T-test

1. Decide on hypotheses:
   Two sided: $H_0: \mu = \mu_0$, vs. $H_1: \mu \neq \mu_0$
   
   One sided:
   
   $H_0: \mu \leq \mu_0$, vs. $H_1: \mu > \mu_0$, or
   
   $H_0: \mu \geq \mu_0$, vs. $H_1: \mu < \mu_0$

2. Decide on significance level $\alpha$

3. Compute the statistic $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$

4. Rejection criteria
   
   Two sided test: reject $H_0$ if $|T| > t_{n-1,1-\alpha/2}$
   
   One sided test:
   
   If $H_1: \mu > \mu_0$, reject $H_0$ if $T > t_{n-1,1-\alpha}$
   
   If $H_1: \mu < \mu_0$, reject $H_0$ if $T < t_{n-1,\alpha}$

5. Compute p-value
   
   Two sided test: $p$-value = $P(|t_{n-1}| > |T|)$
   
   One sided test:
   
   If $H_1: \mu > \mu_0$, then $p$-value = $P(t_{n-1} > T)$
   
   If $H_1: \mu < \mu_0$, then $p$-value = $P(t_{n-1} < T)$

6. Confidence interval for $\mu$
   
   \[
   \left( \bar{X} - t_{n-1,1-\alpha/2} \cdot s/\sqrt{n} , \bar{X} + t_{n-1,1-\alpha/2} \cdot s/\sqrt{n} \right)
   \]
One-sample Z-test for proportions

Only valid if \( np(1-p) > 5 \). Estimate \( p \) using \( \hat{p} \).

1. Decide on hypotheses:
   
   Two sided: \( H_0: p = p_0, \) vs. \( H_1: p \neq p_0 \)
   
   One sided:
   
   \( H_0: p \leq p_0, \) vs. \( H_1: p > p_0 \), or
   
   \( H_0: p \geq p_0, \) vs. \( H_1: p < p_0 \)

2. Decide on significance level \( \alpha \)

3. Compute the statistic
   
   \[ Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} \]

4. Rejection criteria
   
   Two sided test: reject \( H_0 \) if \( |Z| > Z_{1-\alpha/2} \)
   
   One sided test:
   
   If \( H_1: p > p_0, \) then reject \( H_0 \) if \( Z > Z_{1-\alpha} \)
   
   If \( H_1: p < p_0, \) then reject \( H_0 \) if \( Z < Z_{\alpha} \)

5. Compute p-value
   
   Two sided test: p-value = \( P(|N(0,1)| > |Z|) \)
   
   One sided test:
   
   If \( H_1: p > p_0, \) then p-value = \( P(N(0,1) > Z) \)
   
   If \( H_1: p < p_0, \) then p-value = \( P(N(0,1) < Z) \)

6. Confidence interval for \( p \):
   
   \[ \left( \hat{p} - Z_{1-\alpha/2} \cdot \sqrt{\hat{p}(1 - \hat{p})/n}, \hat{p} + Z_{1-\alpha/2} \cdot \sqrt{\hat{p}(1 - \hat{p})/n} \right) \]
Using Table 4 to estimate t probabilities

Because Table 4 in the coursepack only gives percentiles of selected t distributions, we cannot get exact t probabilities from them. We can get bounds on the probabilities.

Example: Use table 4 to estimate \( P(t_{83} > 2.29) \).

- Table 4 implies that 1.99 and 2.37, are the 97.5\(^{th}\) and 99\(^{th}\) percentiles of the \( t_{83} \) distribution.

- Which says:
  - \( P(t_{83} > 1.99) = .025 \), and
  - \( P(t_{83} > 2.37) = .01 \).

- Since 2.29 is between the 1.99 and 2.37, it follows that \( P(t_{83} > 2.29) \) is between \( P(t_{83} > 1.99) \) and \( P(t_{83} > 2.37) \).

- So \( P(t_{83} > 2.29) \) is between .01 and .025.