Please download the following files:

1. iowacorn.dat
2. poverty.dat

SAS for one-sample t-tests
- SAS automatically does a two-sided test
  
  \[ H_0: \mu = \mu_0 \]
  \[ H_a: \mu \neq \mu_0 \]

We will use the “iowacorn.dat” data on annual precipitation, temperature, corn production, and acres harvested in corn in each year of a 10 year period, to test two sets of hypotheses. The first is for annual precipitation in inches:

\[ H_0: \mu_{\text{precip}} = 35 \]
\[ H_a: \mu_{\text{precip}} \neq 35 \]

The second is for annual average temperature:

\[ H_0: \mu_{\text{temp}} = 45 \]
\[ H_a: \mu_{\text{temp}} \neq 45 \]

We will test both null hypotheses at the .10 significance level. We will first use the confidence-interval method.

```
data corn;
  * infile 'c:\temp\iowacorn.dat';
  input precip temp corn acres;
datalines;
* note: copy and paste data in here ;
;
run;
```

```
proc means data = corn n mean stddev stderr clm alpha = .10 ;
var precip temp ;
run ;
```

### Lower 90%

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Std Error</th>
<th>CL for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>precip</td>
<td>10</td>
<td>33.9100000</td>
<td>6.6430331</td>
<td>2.1007115</td>
<td>30.0591585</td>
</tr>
<tr>
<td>temp</td>
<td>10</td>
<td>48.7610000</td>
<td>1.6182463</td>
<td>0.5117344</td>
<td>47.8229330</td>
</tr>
</tbody>
</table>

### Upper 90%

<table>
<thead>
<tr>
<th>Variable</th>
<th>CL for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>precip</td>
<td>37.7608415</td>
</tr>
<tr>
<td>temp</td>
<td>49.6990670</td>
</tr>
</tbody>
</table>

1. What assumptions are necessary to justify computing t confidence intervals and using them to do t hypothesis tests?

2. The first confidence interval that SAS produced above was (30.06, 37.76). We are 90% confident that ______ is in this interval.

3. What can we conclude from this confidence interval about the hypothesis test regarding \( \mu_{\text{precip}} \)?

4. What can we conclude from the other confidence interval about the hypothesis test regarding \( \mu_{\text{temp}} \)?

### One-sample t-tests using proc univariate

Proc univariate knows how to do only one kind of t-test:

- one-sample
- two-sided

```
proc univariate mu0 = 35 45 data = corn ;
var precip temp ;
run ;
```
The UNIVARIATE Procedure
Variable: precip

Tests for Location: Mu0=35

Test  -Statistic- -----p Value------
Student’s t t -0.51887 Pr > |t| 0.6164
Sign M 0 Pr >= |M| 1.0000
Signed Rank S -2.5 Pr >= |S| 0.8457

The UNIVARIATE Procedure
Variable: temp

Tests for Location: Mu0=45

Test  -Statistic- -----p Value------
Student’s t t 7.349515 Pr > |t| <.0001
Sign M 5 Pr >= |M| 0.0020
Signed Rank S 27.5 Pr >= |S| 0.0020

Paired t-test

To carry out the hypothesis test of interest, we apply one-sample procedures to the differences between values measured on members of each pair.

We are interested in whether life expectancy is the same for males as for females. We have a dataset containing various demographic and public health variables on 97 countries in the world in the early 1990s. Two variables reported on each country are the life expectancy at birth for males and the life expectancy at birth for females. Our null hypothesis is that the mean life expectancy for males in the population of all countries in the world is the same as the mean life expectancy for females in the population of all countries.

We will do a two-sided test, because we do not know in advance whether to expect \( \mu_1 \) (mean male life expectancy) to be higher or lower than \( \mu_2 \) (mean female life expectancy).

\[ H_0 : \mu_1 = \mu_2 \]
\[ H_a : \mu_1 \neq \mu_2 \]

or equivalently:

\[ H_0 : \mu_1 - \mu_2 = 0 \]
\[ H_a : \mu_1 - \mu_2 \neq 0 \]

where \( \delta \) denotes \( \mu_1 - \mu_2 \).

We will use the observed differences between the male and female life expectancies observed on each country as our data to carry out the hypothesis test regarding \( \delta \) at the .05 significance level.

Note that by default, proc univariate tests the null hypothesis that \( \mu = 0 \), so in this case we don’t have to give it a value for \( \mu_0 \).
data poverty;
* infile 'c:\temp\poverty.dat';
length country $20. ;
input livebrth death infdeath mlifeexp flifeexp pcgdp group @53 country ;
diff = mlifeexp - flifeexp;
datalines;
* note: copy and paste data in here;
;
run;

proc means data = poverty n mean stddev stderr clm alpha = .05;
var diff;
run;

proc univariate data = poverty;
var diff;
run;

The MEANS Procedure
Analysis Variable : diff

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Std Error</th>
<th>CL Lower 95%</th>
<th>CL Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>97</td>
<td>-4.6655670</td>
<td>2.3711209</td>
<td>0.2407509</td>
<td>-5.1434537</td>
<td>-4.1876803</td>
</tr>
</tbody>
</table>

The UNIVARIATE Procedure
Variable: diff

Tests for Location: Mu0=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student's t</td>
<td>t</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Sign</td>
<td>M</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Signed Rank</td>
<td>S</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>