Tonite:

Introduction to inference (making quantitative data-based conclusions)

1. confidence intervals (not in Ch 6 Moore)
2. significance testing (Ch 6)
3. prediction intervals (not in Ch 6 Moore)

in The (irrealistish) situation where "I is known"

Sampling dist of \( \bar{X} \) is approximately normal (either because I'm willing to say the population is normal or because n is large)

\[ \frac{\bar{X} - \mu}{s / \sqrt{n}} \]

\[ \text{dsn of } \bar{X} \]

\[ M_{\bar{X}} = \mu \]

Confidence Intervals - data-based intervals intended to bracket an unknown population parameter that carry a probability-based reliability (confidence) figure

Basic idea: sampling dists for variables involving parameters can sometimes lead to interval formulas for those parameters

Example: Suppose that surveys have consistently shown that 2 BR apartment rents in Ames IA have std dev $80 at any point in time —

Suppose that this month a sample of \( n = 25 \) apartments will be taken

\[ \bar{X} = \text{sample average rent} \]

\[ \frac{\bar{X} - \mu}{s / \sqrt{n}} \]

\[ m = 2 \]
Thus is roughly a 95% chance of getting an $\bar{x}$ that is within $2 \frac{80}{125}$ of $\mu$.

An interval from $\bar{x} - 2 \frac{80}{125}$ up to $\bar{x} + 2 \frac{80}{125}$ has about a 95% chance of catching $\mu$ - so

$(\bar{x} - 2 \frac{80}{125}, \bar{x} + 2 \frac{80}{125})$ is called a 95% CI for $\mu$.

General version is:

$$\bar{x} \pm z \frac{t}{\sqrt{n}}$$

are confidence limits for $\mu$ with confidence level $p(-z < \text{a standard normal variable} < z)$.

<table>
<thead>
<tr>
<th>confidence level</th>
<th>$z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>80%</td>
<td>1.282</td>
</tr>
<tr>
<td>90%</td>
<td>1.645</td>
</tr>
<tr>
<td>95%</td>
<td>1.96</td>
</tr>
<tr>
<td>99%</td>
<td>2.576</td>
</tr>
</tbody>
</table>

Example: Vardeman's red bag

$r = 1.715$

make some 80% CI's for $\mu$ based on $n = 5$

$$\frac{\bar{x} - z}{\sqrt{n}} = 1.282 \frac{1.715}{\sqrt{5}}$$

$= 38$
<table>
<thead>
<tr>
<th>Sample</th>
<th>$\overline{x}$</th>
<th>Success?</th>
</tr>
</thead>
<tbody>
<tr>
<td>6,7,7,8,4</td>
<td>6.4</td>
<td>no 🙆</td>
</tr>
<tr>
<td>3,3,3,9,4</td>
<td>3.4</td>
<td>no 🙆</td>
</tr>
<tr>
<td>5,5,6,6,1</td>
<td>4.6</td>
<td>yes 😊</td>
</tr>
<tr>
<td>6,5,4,3,7</td>
<td>5.0</td>
<td>yes 😊</td>
</tr>
<tr>
<td>6,6,6,7,7</td>
<td>6.4</td>
<td>no 🙆</td>
</tr>
</tbody>
</table>

etc.

and of many attempts, about 80% will be successes

---

confidence level = lifetime batting average
Moore use the terminology

$$\frac{F}{\sqrt{n}} = \text{"margin of error"}$$

large confidence $\leftrightarrow$ large margin of error
it is possible to plan for a given margin and confidence level by choosing an appropriate $n$

---

To say — to — is a 90% CI for $\mu$ is to say that in obtaining it I've used a method that works in about 90% of applications — whether or not it has worked in this particular application I don't know — further, there is no way to make a probability statement about this particular realized interval

Do problem 6.6 page 308 of Moore

Example. Ames rent survey - suppose I want 99% and $\pm 2.5$ margin of error in estimating $\mu$ — $n =$ ?

$$\bar{x} = 2.576$$

$$25 = \text{margin of error} = \frac{\bar{x}}{\sqrt{n}} = 2.576 \frac{80}{\sqrt{n}}$$

$$\sqrt{n} = 2.576 \frac{80}{25}$$

$$n = (2.576 \frac{80}{25})^2 = 68$$
General Formula

\[ n = \left( \frac{z \cdot \text{margin of error}}{\text{sample mean}} \right)^2 \]

2nd type of inference: Prediction for a simple additional value - using a normal model - normal no good, neither is the following development

So there is about a 95% chance that \( z_{\text{new}} - \bar{z} \) is between

\[-2\sigma \sqrt{1 + \frac{1}{n}} \quad \text{and} \quad 2\sigma \sqrt{1 + \frac{1}{n}}\]

This is equivalent to

\[ \bar{z} - 2\sigma \sqrt{1 + \frac{1}{n}} < z_{\text{new}} < \bar{z} + 2\sigma \sqrt{1 + \frac{1}{n}}\]

which suggests that I call

\[ \bar{z} \pm 2\sigma \sqrt{1 + \frac{1}{n}} \]

95% prediction limits for \( z_{\text{new}} \)

\[ z = \text{sample mean of n observations} \]

\[ z_{\text{new}} = \text{simple additional observation from the same normal dsn} \]

Fact: The sampling dsn of \( z_{\text{new}} - \bar{z} \)

is normal with mean 0 and std dev \( \frac{1}{\sqrt{1 + \frac{1}{n}}} \)

\[ \frac{1}{\sqrt{1 + \frac{1}{n}}} \]

In general, the limits are

\[ \bar{z} \pm z \cdot \sqrt{1 + \frac{1}{n}} \]

Example: Ames rents

make a 95% prediction interval for a single additional apartment rent

\[ n = 29 \quad \sigma = 80 \quad \bar{z} = 688.20 \]

\[ 688.20 \pm 1.96(80) \sqrt{1 + \frac{1}{29}} \]

\[ 688.20 \pm 16.3 \]

much larger than \( \pm 29.7 \) from before.
**Null Hypothesis:**
A null hypothesis is a statement about a parameter of a population. A model parameter is the parameter being sampled.

**Significance Testing:**
Significance testing is a method of assessing the probability of a statement being true.

**Basic Types of Inference:**
- **Pre-Data Status:** The alternative hypothesis is a statement that embodies those departures from the null hypothesis that one wishes to detect. This takes one of two forms:
  - $H_a: \text{parameter } \neq c$
  - $H_a: \text{parameter } > c$

**Statistical Process:**
- **A Type I Error:**MSC 2/14/16
- **A Type II Error:**
- **Critical Value:**
- **P-value:**

**Sample Size:**
1. **Sample Size Formula:** $n = \frac{Z^2 \sigma^2}{\theta^2}$
2. **Type of Test:** Z-test
3. **Decision Rule:**
- **Level of Significance:** $\alpha = 0.05$
- **Critical Value:** $Z = 1.65$
- **Test Statistic:** $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

**Confidence Interval:**
- **Confidence Level:** 95%
- **Margin of Error:** $E = Z(\frac{\sigma}{\sqrt{n}})$
- **Confidence Interval Formula:** $\bar{X} - E < \mu < \bar{X} + E$

**Interpretation:**
- To say that $95\%$ of the $p$-values is to say that $95\%$ of the $p$-values are $< 0.05$.

**Example:**
- **Red Sox PIs for 2007:**
  - $1.282 < 1.75$ (using $n = 5$)

**Notes:**
- $\theta$: parameter (e.g., $n = 5$)
- $Z = 1.65$
- $\sigma$: standard deviation
- $\bar{X}$: sample mean
- $\mu$: population mean
- $E$: margin of error
to assess the plausibility of $H_a$, we use a "test statistic" and compute a "p-value."

**Def**: The **test statistic** is the data summary used.

**Def**: The **p-value** is the probability that the sampling distribution of the test statistic assigns to things as extreme as the data in hand when $H_0$ is true.

$$z = \frac{688.20 - 680}{\frac{80}{\sqrt{123}}} = 5.5$$

$$P(Z > 5.5) = 1 - .7088 = .2912$$

Example: Amy's apartment rents -

***Projecting*** based trends from last decade, we see that $\mu = 680$ - students claim this is too low. $\bar{x} = 688.20$...

How plausible is $\mu = 680$?

$H_0: \mu = 680$

$H_a: \mu > 680$

a sensible test statistic is $\bar{x}$

big p-value $\iff$ $H_0$ is not implausible

small p-value $\iff$ $H_0$ implausible

In this example, I used

$H_0: \mu =$

$H_a: \mu >$ as evidence against $H_0$

ultimately I turned $\bar{x}$ into a z-score
- I could think of the test statistic as
  \[ z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \]

  Thinking in terms of \( z \)-scores. The \( z \) sets of hypotheses and ways to get
  \( p \)-values are:

  \[ H_0: \mu = \# \quad H_0: \mu = \# \quad H_0: \mu = \# \]
  \[ H_a: \mu > \# \quad H_a: \mu < \# \quad H_a: \mu \neq \# \]

  \( z^* = z \) score for data in hand

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

  or Moore

  \( \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \)

  \( \bar{x} \)

  \( \sigma \)

  \( n \)

  **Comments**

  1. Sometimes people call

    \( p \)-value < .05 "statistically significant"

    \( p \)-value < .01 "highly statistically significant"

    be careful

  statistical significance \( \neq \) practical importance

  have enough data to see \( H_0 \) is wrong

  \( H_0 \) is far enough from right that I care

  \( \bar{x} \neq \# \)

  2. A \( p \)-value is not a "probability"

    The null hypothesis is right (or wrong)

    It's a measure of strength of evidence

    against \( H_0 \)

  3. Confidence Intervals

    a) are more informative than significance

    tests (they answer the right question)

    "What is \( \mu \)?" vs "Is \( \mu = \# \)?"

    b) carry testing information - if a 50%

    CI for \( \mu \) doesn't cover \( \# \) then \( p \)-value for

    testing \( H_0: \mu = \# \) vs \( H_a: \mu \neq \# \) is < .10

  18
The key fact in the Ch 6 introduction to inference was that $Z$ is normal, i.e. that:

$$Z = \frac{\bar{X} - \mu}{\sigma}$$

is standard normal

Ch 6 methods are usually unusable since they involve $t$ and that comes from $t$ IC like methods that don't require $\sigma$ as an input.

$t$ dsns are bell-shaped dsns, centered at 0, that are somewhat flatter/more spread out than the standard normal dsn. $t$ dsn with $\mu$ normal has a tabled $d.f. = n-1$.

Diezian Problem 35, page 58 gives 1999 annual rates of return on $n=27$ no load mutual funds -- ($\bar{X} = 13.3\%$ and $s = 4.1\%$) -- supposing that these $n=27$ can be treated as a random sample of all such funds (?????) and that annual rates of return for such funds in 1999 were approximately normal.

$$P(-2.056 < \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} < 2.056) = .95$$
i.e. 95% of sample produce

\[-2.056 < \frac{\bar{x} - \mu}{S/\sqrt{n}} < 2.056\]

equivalent to

\[-2.056 \frac{S}{\sqrt{27}} < \mu < \bar{x} + 2.056 \frac{S}{\sqrt{27}}\]

which suggests 95% limits for \(\mu\)

Example. No load mutual fund example

\(n=27\), \(\bar{x} = 13.3\), \(S = 4.1\)

95% confidence limits for \(\mu\)

\(\bar{x} \pm \frac{S}{\sqrt{n}}\)

\(13.3 \pm 2.056 \frac{4.1}{\sqrt{27}}\)

1.6

\(\bar{x} \pm 2.056 \frac{S}{\sqrt{27}}\)

which is a special case of

\(\bar{x} \pm \frac{2}{\sqrt{n}}\)

df = n-1

Note that here the "margin of error" changes "sample to sample" (since \(S\) varies)

Example. Red Bug - take some samples of size \(n=5\) and make 80% C.I.'s for \(\mu\) (know to be 5)

\(\bar{x} \pm \frac{S}{\sqrt{15}}\)

1.533

\(\bar{x} \pm 0.696\)
Problem 7.7 a) b) Page 374

Example: Predict (95%) return for 60% prediction limits

We have 12 measurements with a mean of 10.2 and a standard deviation of 2.5. We need to calculate the confidence interval for the mean.

Sample size: 12

\[ \bar{x} = 10.2 \]

Standard deviation: 2.5

95% confidence level

\[ t_{0.025, 11} = 2.201 \]

\[ s = 2.5 \]

\[ n = 12 \]

\[ \alpha = 0.05 \]

\[ t_{0.025, 11} = 2.201 \]

\[ \text{CI} = \bar{x} \pm t_{0.025, 11} \frac{s}{\sqrt{n}} \]

\[ 10.2 \pm 2.201 \frac{2.5}{\sqrt{12}} \]

\[ 10.2 \pm 0.41 \]

\[ 9.79 < \mu < 10.61 \]

This interval is the predicted range for the mean.
Example: Didmen mutual fund data

Was the mean rate of return for no load mutual funds in 1999 clearly below 15%? Use a significance test

$H_0: \mu = 15$

$H_a: \mu < 15$

$t = \frac{\bar{x} - 15}{\frac{s}{\sqrt{n}}} = \frac{13.3 - 15}{\frac{4.1}{\sqrt{27}}} = -2.15$

$p$-value is between 0.025 and 0.01.

Remember that we had 95% confidence limits for $\mu$ of $13.3 \pm 1.6$
Matched Pairs application of one-sample t methods

\[ \text{individual 1: } x_1, y_1 \]
\[ \text{individual 2: } x_2, y_2 \]
\[ \text{individual 3: } x_3, y_3 \]
\[ \text{individual n: } x_n, y_n \]

It can make sense to simply reduce x and y to \( d = y - x \) and apply methods of 7.1 to d's.

The object is comparison of MA and MB based on the 2 samples.

Basic quantity used in 7.1 is

\[ \frac{\bar{x}_A - \bar{x}_B - (M_A - M_B)}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} \]

Sadly this isn't exactly t distributed but treat it as if it was using conservative degree of freedom smaller sample size - 1 and use conservative t value.

\[ M_A = 5 \]
\[ M_B = 10 \]
\[ t_A = 1.715 \]
\[ t_B = 3.47 \]
\[ n_A = 5 \quad n_B = 4 \]

\[ 3, 3, 3, 3, 7 \quad 2, 5, 15, 15 \]

\[ \bar{x}_A = 4.9 \quad \bar{x}_B = 9.25 \]

\[ s_A = 1.95 \quad s_B = 6.75 \]

\[ \text{make a 95\% C.I. for } \mu_A - \mu_B \]

\[ (4.4 - 9.25) \pm 2.353 \sqrt{\frac{(1.95)^2}{5} + \frac{(6.75)^2}{4}} \]

\[ -4.85 \pm 8.20 \]

\[ \text{a winner (truth is } -5 = \mu_A - \mu_B) \]

\[ \frac{17.5 - 13.7}{3.8} \pm 1.660 \sqrt{\frac{(3.5)^2}{12} + \frac{(4.5)^2}{9}} \]

\[ 3.8 \pm 3.36 \]

\[ \text{Significance Testing for } \mu_A - \mu_B ? \]

\[ H_0: \mu_A - \mu_B = 0 \]

\[ H_a: \mu_A - \mu_B < 0 \]

\[ \text{Test statistic} \]

\[ T = \frac{\bar{x}_A - \bar{x}_B - \mu}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} \]

\[ = -1.39 \]

\[ \text{Problem 7.32 c) } \]

\[ \bar{x}_A = 17.5 \quad \bar{x}_B = 13.7 \]

\[ s_A = 3.5 \quad s_B = 4.5 \]

\[ n_A = 12 \quad n_B = 9 \]

\[ \text{Example: Red + Striped bags} \]

\[ \text{Do they have the same mean?} \]

\[ H_0: \mu_A - \mu_B = 0 \]

\[ H_a: \mu_A - \mu_B \neq 0 \]

\[ T = \frac{4.4 - 9.25}{-1.39} \]

\[ \text{p-values using} \]

\[ d.f. = \text{smaller} \quad \text{sample size} - 1 \]
Problem 7.32 b)

\[ T = \frac{17.5 - 13.7}{\sqrt{\frac{(3.9)^2}{12} + \frac{(4.5)^2}{5}}} = 2.10 \]

\[ H_0: \mu_A - \mu_B = 0 \]
\[ H_a: \mu_A - \mu_B > 0 \]

A = unlogged, B = logged

p-value here is between 2(0.10) and 2(0.05)

2(0.10) and 0.05

between 0.025 and 0.01