\[
\frac{\bar{x}}{\sqrt{n}} = \frac{s}{\sqrt{n}} = \frac{5.3}{\sqrt{30}} = .83
\]

c) \( P(\bar{x} \geq 21) = ? \)

\[ z = \frac{21 - 18.6}{.83} = 2.89 \]

2.89 look-up gives \(.0019 \leq \alpha < .0025\)

Confidence Intervals are data-based intervals meant to bracket some unknown population or model parameter that carry a probability-based reliability or confidence figure.

Basic idea used: sampling distributions for variables that involve parameters of interest can lead to formulas for intervals for those parameters.

\[ z = \frac{\bar{x} - \mu}{s/\sqrt{n}} \]

Tenets:
- Use sampling distributions to introduce basics of (probability-based) "statistical inference"—Ch 6 of Moore
- plus some:
  1. Confidence Intervals
  2. Prediction Intervals
  3. Significance/Hypothesis Testing

Formulas for inference are mathematically OK, but useless (involve \( \bar{x} \))

We know that (either because a population is normal or \( n \) is large) \( \bar{x} \) can often be treated as normal.
95% of \( \bar{x} \)'s within \( 2 \frac{S}{\sqrt{n}} \) of \( \mu \)

means that limits not usually available

\[ \bar{x} \pm 2 \frac{S}{\sqrt{n}} \]

will catch \( \mu \) "95% of the time".

So we will call an interval with these end-points a "95% CI for \( \mu \)".

Example: Suppose that I'm interested in average 2BR apartment rent in Ames, IA - perhaps historical info.

The general version of this is

\[ \bar{x} \pm z \frac{S}{\sqrt{n}} \]

where \( z \) is chosen so that

\[ \Pr(-z < z\text{-score} < z) = \text{desired confidence level} \]

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

says that \( S = 80 \) for the population of rental rates - if a sample today of \( n = 25 \) gives \( \bar{x} = 688.20 \)

Then a 95% CI for mean rate today is

\[ \bar{x} \pm 2 \frac{S}{\sqrt{n}} \]

\[ 688.20 \pm \left( 2 \frac{80}{\sqrt{25}} \right) \]

interval is

\[ (658.5, 717.90) \]

Demonstrated: Red Bag \( \nu = 1.715 \)
population/universe is (roughly) normal
\( n = 5 \) \( \bar{x} \) is normal
make some 80% CI's for \( \mu \)

\[ \bar{x} \pm 1.282 \frac{1.715}{\sqrt{5}} \]

\( .98 \)
| Sample | $z$ | Does interval work?
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7, 7, 5, 4</td>
<td>5.6</td>
<td>Yes</td>
</tr>
<tr>
<td>3, 3, 7, 5, 7</td>
<td>6.6</td>
<td>No</td>
</tr>
<tr>
<td>1, 4, 5, 3, 2</td>
<td>3.0</td>
<td>No</td>
</tr>
<tr>
<td>4, 7, 6, 4, 7</td>
<td>5.6</td>
<td>Yes</td>
</tr>
<tr>
<td>9, 3, 5, 4</td>
<td>5.0</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Lifetime batting average should be 80%.

---

This demonstration is meant to give you the right understanding of how a 95% CI should be interpreted.

To say — to — is a 95% CI for $\mu$ is to say that in obtaining it I've used a method that works in about 95% of applications — whether it has or has not worked in my particular application is typically unknown — regardless there is no probability statement left to make.

Example: Anus 2BR apartment survey — I want to plan a sample size to produce 95% confidence and $\pm 25$.

\[
\frac{\sigma}{\sqrt{n}} = \text{"margin of error"}
\]

\[
\frac{25}{2.576} = \frac{\sigma}{\sqrt{n}} \Rightarrow \sigma = 25 \times \frac{2.576}{25} = 6.4
\]

Do 6.6 of Moore page 308.

Moore calls

\[
z = \frac{X - \mu}{\sigma/\sqrt{n}} = \text{"margin of error"}
\]

larger $n$ gets you smaller "margin of error".

larger confidence gets you larger "margin of error".

you can choose a sample size to get a desired combination of confidence and margin of error.
General Formula

\[ n = \left( \frac{z \times \text{margin of error}}{\text{standard deviation}} \right)^2 \]

**2nd type of Inference: Prediction Intervals**

Object: bracket \( x_{\text{new}} \) (not \( x \)!

notice that if I have complete knowledge about the population, this is a probability problem.

e.g. for red bag \( x \pm z(1.75) \) has about a 95% chance of bracketing \( x_{\text{new}} \)

\[ z_{\text{new}} = z \]

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

so 95% of differences \( x_{\text{new}} - \bar{x} \) will fall between

\[ 0 - 2z\sqrt{\frac{1}{n}} \text{ and } 0 + 2z\sqrt{\frac{1}{n}} \]

for 95% of experiences

but now, what if I don't know \( x \)?

(idea: if I don't know either \( \mu \) or \( \sigma \))

- use \( \bar{x} \) to approximate \( x \) and
- "judge" properly — this is based on another sampling dsn, one for

\[ \bar{x} \]

an additional observation

\[ \bar{x} \]

If the population is normal, I can tell you about this sampling dsn

\[ -2\sqrt{\frac{1}{n}} < x_{\text{new}} - \bar{x} < 2\sqrt{\frac{1}{n}} \]

i.e. for 95% of experiences I'll have

\[ \bar{x} - 2\sqrt{\frac{1}{n}} < x_{\text{new}} < \bar{x} + 2\sqrt{\frac{1}{n}} \]

so, I'll call

\[ \bar{x} \pm 2\sqrt{\frac{1}{n}} \]

95% prediction limits for \( x_{\text{new}} \)
Example: Amos rents ZBR apartment

\[ n = 25 \text{ and normal} \]

Given \[ \bar{x} = 688.20 \] make a 95% prediction interval for an additional apartment rent.

\[ 688.20 \pm 1.96 \sqrt{\frac{15}{25}} \]

\[ 688.20 \pm 162.70 \]

Price is more than \[ 2\sigma(20) \text{ no surprise - } 95\% \text{ have to be wider than } CI \]

Here my lifetime batting average

For whole business of selecting the sample of \( n = 5 \) computing interval and then selecting \( x_{\text{new}} \) is supposed to be about 80%.

General formula:

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

Demonstration: Red Bar make some 80% P.I.'s for \( x_{\text{new}} \) (\( n = 5 \))

\[ \bar{x} = \frac{\sum x}{n} = 4 \]

\[ z = 1.96 (1.715) \sqrt{1 + \frac{1}{5}} \]

\[ z = 2.41 \]

\[ 2.41 = \sqrt{\frac{1.282(1.715)}{1 + \frac{1}{5}}} \]

Sample | \( \bar{x} \) | \( x_{\text{new}} \) | Interval work?
|-------|--------|-----------------|
| 3, 6, 1, 4, 6 | 5.0 | 4 | yes
| 6, 6, 1, 5, 6 | 5.6 | 3 | no
| 2, 6, 4, 1, 3 | 3.6 | 6 | yes

For situation of Moore's 6.6 page 308 make a 90% prediction interval for tomorrow's potassium level based on \( n = 3 \)

3rd type of "standard" inference is significance/hypothesis testing

The idea is one of assessing the plausibility of a statement about a parameter.
A null hypothesis is a statement about a parameter of the form 
\[ H_0: \text{parameter} = \# \]
That represents a status quo/pre-data view.

The alternative hypothesis is a statement about the parameter that embodies those departures from \( H_0 \) that we want to detect - this is of the form 
\[ H_a: \text{parameter} \neq \# \]

A test statistic is the data summary to be used.

A p-value is the probability that the sampling distribution of the test statistic assigns to values that are more extreme than the one actually observed if \( H_0 \) is in fact true.

An example: Assume rent - suppose last year's mean was $660 and by standards of CPI a "fair" increase would be to $680. If I'm a consumer advocate, I might use
\[ H_0: \mu = 680 \]
\[ H_a: \mu > 680 \]
(laboring for price gouging in Amus)

Example: Amus rent. \( z \) is test statistic.

To get the p-value, I need a normal area - so I need a z-score.
\[
z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{688.2 - 680}{\sigma/\sqrt{40}} = 38
\]
look-up gives me .7088 as tabled
so
\[p\text{-value} = 1 - .7088 = .2912\]

I.e. I'll see results this extreme nearly 30% of the time even if \(H_0\) is true.
This is therefore only very weak evidence against \(H_0\).

I.e. the test statistic could be
\[
z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}
\]
and using this the 3 sets of hypotheses and \(p\)-values are as follows:

- \(H_0: \mu = \#\)
  - \(H_0: \mu = \#\)
  - \(H_0: \mu = \#\)
- \(H_0: \mu > \#\)
- \(H_0: \mu < \#\)

Small \(p\)-values \(\rightarrow\) strong evidence against \(H_0\)

Large \(p\)-values \(\rightarrow\) weak evidence against \(H_0\)

In our example we used
\(H_0: \mu = \#\) is a right tail area
\(H_0: \mu = \#\) for \(p\)-value
to get the \(p\)-value we had to go through
\(z\)-score ... we might call the \(z\)-score that...;

Problem 6.36 page 333 Moore
also reed the problem from the point
of view of a production supervisor
who is concerned about any deviation
from 300 ml, up or down

Comments/Philosophy Regarding Testing

1. Sometimes people use terminology
   \(p\)-value < .05 \(\leftrightarrow\) "statistically significant"
   \(p\)-value < .01 \(\leftrightarrow\) "highly statistically significant"
Beware that you don't hear more than you should.

Statistical significance ≠ Practical importance.

You have enough data to say that Ho='Wrong'... but is it wrong?

If a 90% CI for μ fails to include μ, p-value to test Ho:μ=μ vs Hα:μ≠μ is less than 10%.

3. A p-value is not a "probability that Ho is true" (or a "probability that Ho is false")... rather it is a measure of strength of evidence.

By my standards CIs are much more informative than tests.

- They attempt to answer "The right question"... namely "What is the parameter?" and answer the "Wrong question"... is the parameter ≠ ?
- Besides, they carry significance testing information any way, i.e. if a 90% CI for μ includes μ then for testing Ho:μ=μ vs Hα:μ≠μ p-value exceeds 10%.

Title: "Real" Inference / Ch 7 & More
(get rid of the "known σ"requirement)

What we did last was built primarily on the fact

\[ z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \]  This propagates through formulas.

is (at least approximately) std normal.
It would be nice if I could replace $t$ with $z$ and still have something to work with... in a way we can do that—when sampling from a normal population:

$$ t = \frac{\bar{z} - \mu}{\frac{SE}{\sqrt{n}}} $$

not normal, but it does have a table den—
the so-called "t" den with $t = n - 1$

$\text{desired value}$

$-2.056 < \bar{z}$

$\text{from the t table}$

$\P(-2.056 < \bar{z} < 2.056) = .95$

Diehlman Problem 35 page 58 gives 1995 annual rates of return for $n = 27$ no load mutual funds—$\bar{z} = 13.3\%$ $s = 4.1\%$

Suppose 1) 1995 rates of return were normal and 2) this was a random sample—based on these, we’ll do inference:

Then,

$$ -2.056 < \frac{\bar{z} - \mu}{\frac{s}{\sqrt{27}}} < 2.056 $$

is the same as

$$ \frac{\bar{z} - 2.056s}{\sqrt{27}} < \mu < \bar{z} + 2.056\frac{s}{\sqrt{27}} $$
That is, I can use
\[ \bar{x} \pm 2.056 \frac{s}{\sqrt{n}} \]
as 95% confidence limits for \( \mu \).
In this example that's
\[ 13.3 \pm 2.056 \frac{4.1}{\sqrt{127}} = 1.6 \]

General Formula
\[ \bar{x} \pm t \frac{s}{\sqrt{n}} \]  
where \( t = t_{n-1} \)

"Standard error" = estimated standard deviation
\[ \frac{s}{\sqrt{n}} = \text{standard error of the mean} \]
\[ \frac{s}{\sqrt{n}} = \text{std deviation of } \bar{x} \]

Problem 7.7 a), b) page 374
\[ \bar{x} \pm t \frac{s}{\sqrt{n}} \]
\[ 1.75 \pm 2.353 \frac{1.35}{\sqrt{4}} = 1.15 \]

Comments: 1) Moore talks about "robustness" of this method, as its nominal confidence levels won't be ridiculously wrong as long as the universe sampled is not terribly non-normal.
Example: Dilemma of load mutual fund
rate of return - yields 55% per year.
For a single additional mutual fund: $n = 27$
$z = 3.3, s = 4.1$
$\sigma = 2.056(4.1) / (1 + 27)$

\[ z = \frac{v - \mu}{\sigma} \]

Formulas called a large sample

Formula:

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

Sometimes called a large sample

Real: Prediction Intervals - I can do

This leads to prediction limits

Population

\[ \bar{x} \pm z \frac{\sigma}{\sqrt{n}} \]

Has a t-distribution

with df = n - 1

\[ s \]

95% CI: When sampling a normal
Problem 7.7 page 374 make a 90% PI for a single additional rate
\[
\frac{3 \pm 2.333(0.125)\sqrt{1 + \frac{1}{n}}}{0.35}
\]
1.75 \pm 2.353 (0.125) \sqrt{1 + \frac{1}{n}}

significance testing goes the same way I use
\[
t = \frac{\bar{x} - \mu}{s / \sqrt{n}}
\]
and \( t \) goes to \( t \)-table with \( n-1 \) df to get \( p \)-values

Example

Diehman no load mutual funds. Was the average rate of return for no load mutual funds clearly below 15 in 1999?

\( H_0: \mu = 15 \)
\( H_a: \mu < 15 \)

\[
t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{13.3 - 15}{4.1/\sqrt{127}} = -2.5
\]

Problem 7.7 part c

\( H_0: \mu = 1.3 \)
\( H_a: \mu > 1.3 \)

\[
t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{1.75 - 1.3}{0.125/\sqrt{14}} = 6.38
\]

\( t \)-dist

p-value between 0.005 and 0.0025

p-value \( = \frac{1}{14} \)
An important application of these methods for one sample is to situations where we have one sample of "pairs"—say, I have "before" and "after" values or I have "with" and "without" treatment on the same set of individuals.

Two sample problems

```
<table>
<thead>
<tr>
<th>1st sample</th>
<th>2nd sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>x₁, y₁</td>
<td>z₁, y₂</td>
</tr>
<tr>
<td>x₂, y₂</td>
<td>z₂, y₃</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>xₙ, yₙ</td>
<td>zₙ, yₙ</td>
</tr>
</tbody>
</table>
```

A standard method of analysis is to take differences:

\[ x - y = d \]

and do inference for \( d \).

This is a different scenario than having \( z \)'s on one set of individuals and \( y \)'s on another. 2 samples

Mann uses the variable

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

Not really \( t \)-distributed (even if populations are normal) but it turns out that treating it as if it were using

d.f. = smaller sample size - 1

gives "conservative" methods.
Testing that variable as a t-variable gives confidence limits for $\overline{M_A} - \overline{M_B}$

$$\overline{X_A} - \overline{X_B} \pm t \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$

$d.f.$ = smaller of 2 sample sizes - 1

**Demonstration**

Red $(M_A = 5, \overline{X_A} = 1.715)$ $n_A = 4$

Striped $(M_B = 10, \overline{X_B} = 3.47)$ $n_B = 5$

Problem 7.32c page 393

$n_A = 12$, $(\overline{X_A} = 17.5)$ $s_A = 3.5$

$n_B = 12$, $(\overline{X_B} = 13.7)$ $s_B = 4.5$

$t = \frac{(17.5 - 13.7)}{1.86} = \frac{3.8}{1.86} = 2.038$  

$df = 23$  

$t = 1.995$  

This is a winner.

Use 95% confidence

Red $6, 4, 7, 4 \overline{X_A} = 5.25$ $s_A = 1.5$

Striped $7, 13, 8, 15, 10 \overline{X_B} = 11.0$ $s_B = 2.92$

$$\overline{X_A} - \overline{X_B} \pm t \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$

$$(5.25 - 11.0) = \pm 2.52 \sqrt{\frac{(1.5)^2}{12} + \frac{(2.92)^2}{12}}$$

$d.f = 23$  

$t = 1.995$  

$df = 1.995$

2. Sample significance tests for $\overline{M_A} - \overline{M_B}$

$H_0: M_A - M_B = \# \text{ from 0}$

$H_a: M_A - M_B < \#$

and test statistic

$$t = \frac{\overline{X_A} - \overline{X_B} - \#}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

and get p-values from t-dsn with $d.f.$ = smaller sample size - 1
Problem 7.32

\[ H_0: \mu_{\text{unipped}} - \mu_{\text{topped}} = 0 \]

\[ H_1: \mu_{\text{unipped}} - \mu_{\text{topped}} > 0 \]

\[ t = \frac{\bar{z}_A - \bar{z}_B - \delta}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} \]

\[ = \frac{17.5 - 13.7}{\sqrt{\frac{(3.8^2)}{12} + \frac{(4.5^2)}{5}}} = 2.10 \]

with \( df = 2.0 \) confidence limits for \( \mu_A - \mu_B \)

\[ \bar{z}_A - \bar{z}_B \pm 2 \cdot \text{pooled} \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}} \]

\[ \text{diff.} = n_A + n_B - 2 \]

Test \( H_0: \mu_A - \mu_B = \# \) using

\[ t = \frac{(\bar{z}_A - \bar{z}_B - \#)}{\text{pooled} \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \]
Problem 7.32—supposing \( \frac{A}{B} = \frac{B}{B} \)

\[
S_{\text{pooled}} = \sqrt{\frac{\frac{(6.5)^2}{12} + \frac{(7.5)^2}{12}}{13}}
\]

\[= 3.95\]

90\% confidence limits for \( \mu_A - \mu_B \)

\[
\bar{x}_A - \bar{x}_B \pm t \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}
\]

\[17.5 - 13.7 \pm (1.729) 3.95 \sqrt{\frac{1}{12} + \frac{1}{12}}\]