Today:

Chapter 11, confidence intervals for means

Announcements

- Monday discussion is for credit.
- Useful summary tables:

  Sampling distributions: p. 382-383

  Confidence intervals: p. 483

  Hypothesis tests: p. 586-587

Homework: (Due Mon, Nov 29 because of Fri holiday)

Chapter 11: #40, 53, 78 (Use R Commander, counts double – data file linked to website)

Chapter 11

Estimating Means with Confidence

Recall:

- A **parameter** is a population characteristic – value is usually unknown. We estimate the parameter using sample information. Chapter 11: C.I.s for *means*.
- A **statistic,** or **estimate,** is a characteristic of a sample. A statistic estimates a parameter.
- A **confidence interval** is an interval of values computed from sample data that is likely to include the true population value.
- The **confidence level** for an interval describes our confidence in the procedure we used. *We are confident* that most of the confidence intervals we compute using a procedure will contain the true population value.

Recall from Chapter 10 (p. 408)

A **confidence interval** or **interval estimate** for any of the five parameters can be expressed as

Sample estimate ± multiplier × standard error

where the multiplier is a number based on the confidence level desired and determined from the standard normal distribution (for proportions) or Student's *t*-distribution (for means).

Estimation Situations Involving Means

Situation 1. ***Estimating the mean*** *of a quantitative variable.*

Example research questions:

What is the mean number of emails students get per day?
What is the mean number of words a 2-year old knows?

Population parameter: μ (spelled "*mu*" and pronounced "mew") = population mean for the variable

Sample estimate**:** $\overline{x}$ = the sample mean for the variable, based on a sample of size n.

Estimating the Population Mean of Paired Differences

Situation 2. ***Data measured in pairs, take differences, estimate the mean of the population of differences:***

What is the mean difference in blood pressure before and after meditation? (d_i = *difference for person i*)

What is the mean difference in hrs/day studying and online (for non-study reasons) for college students?

Population parameter: μ_d (called "*mu*" *d*)

Sample estimate: $\bar{d}$ = the sample mean for the differences, based on a sample of *n* pairs, where the difference is computed for each pair.

Difference in two means

Situation 3. ***Estimating the difference between two populations*** *with regard to the* ***mean*** *of a quantitative variable.*

Example research questions:

How much difference is there in average weight loss for those who diet compared to those who exercise to lose weight?

How much difference is there in the mean IQs for children whose moms smoked and didn't during pregnancy?

Population parameter: $\mu_1 - \mu_2$ = difference between the two population means.

Sample estimate: $\bar{x}_1 - \bar{x}_2$ = difference between the two sample means. This requires *independent* samples.

Recall from Chapter 10 (p. 408)

A **confidence interval** or **interval estimate** for any of the five parameters can be expressed as

Sample estimate ± multiplier × standard error

where the multiplier is a number based on the confidence level desired and determined from the standard normal distribution (for proportions) or Student's *t*-distribution (for means).

Standard errors (in general)

Rough Definition: The **standard error** of a sample statistic measures, **roughly, the average difference** between the statistic and the population parameter. This "***average difference***" is over all possible random samples of a given size that can be taken from the population.

Technical Definition: The **standard error** of a sample statistic is the *estimated standard deviation of the sampling distribution* for the statistic.

Standard Error of the Mean

In practice, the population standard deviation σ is rarely known, so we cannot compute the standard deviation of $\bar{x}$,

$$\text{s.d.}(\bar{x}) = \frac{\sigma}{\sqrt{n}}.$$

In practice, we only take one random sample, so we only have the sample mean $\bar{x}$ and the sample standard deviation s. Replacing σ with s in the standard deviation expression gives us an estimate that is called the **standard error of** $\bar{x}$.

$$\text{s.e.}(\bar{x}) = \frac{s}{\sqrt{n}}.$$

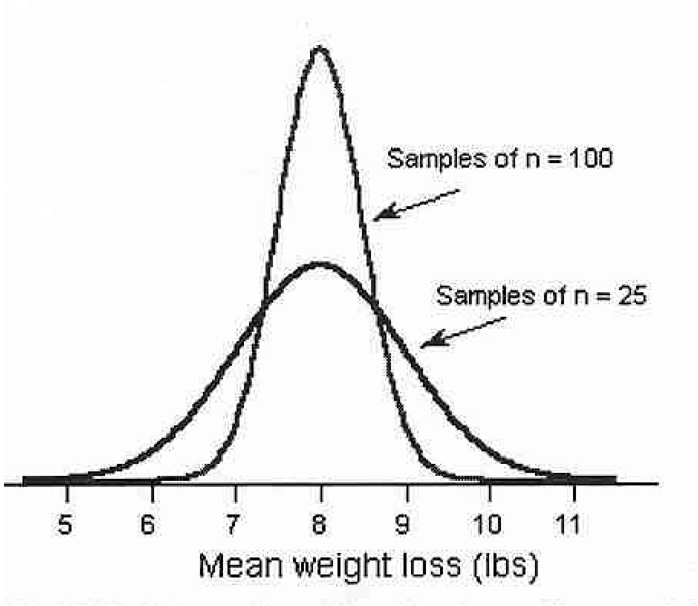
From Chapter 9 example of a sample of $n = 25$ weight losses, the sample standard deviation is $s = 4.74$ pounds.
So the standard error of the mean is 4.74/5 = 0.948 pounds.

Increasing the Size of the Sample

Suppose we take n = 100 people instead of just 25. The standard deviation of the mean would be

$$\text{s.d.}(\bar{x}) = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{100}} = 0.5 \text{ pounds.}$$

- For samples of ***n* = 25**, sample means are likely to range between 8 ± 3 pounds => 5 to 11 pounds.
- For samples of ***n* = 100**, sample means are likely to range only between 8 ± 1.5 pounds => 6.5 to 9.5 pounds.



Larger samples tend to result in ***more accurate*** estimates of population values than smaller samples.

Standard Error of a Sample Mean

$$s.e.(\bar{x}) = \frac{s}{\sqrt{n}}, \quad s = \text{sample standard deviation}$$

Example 11.2 *Mean Hours Watching TV*

Poll: Class of 175 students. In a typical day, about how much time do you spend watching television?

Minitab provides **s.e.**, R Commander doesn't, but provides ***s***

Statistics → Summaries → Numerical summaries, check Standard Deviation

Variable	N	Mean	Median	TrMean	StDev	SE Mean
TV	175	2.09	2.000	1.950	1.644	0.124

$$s.e.(\bar{x}) = \frac{s}{\sqrt{n}} = \frac{1.644}{\sqrt{175}} = .124$$

Standard Error for the mean of paired differences

$$s.e.(\bar{d}) = \frac{s_d}{\sqrt{n}}$$

where s_d = sample standard deviation for the *differences*

Example: How much taller (or shorter) are daughters than their mothers these days? $\boldsymbol{s_d}$ **= 3.14** (for individuals)

n = 93 pairs (daughter – mother) $\bar{d}$ = 1.3 inches

$$s.e.(\bar{d}) = \frac{s_d}{\sqrt{n}} = \frac{3.14}{\sqrt{93}} = .33$$

Standard Error of the Difference

Between Two Sample Means (*unpooled*)

$$s.e.(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Example 11.3 *Lose More Weight by Diet or Exercise?*

Study: $n_1 = 42$ men on diet, $n_2 = 27$ men on exercise routine

Diet: Lost an average of 7.2 kg with std dev of 3.7 kg
Exercise: Lost an average of 4.0 kg with std dev of 3.9 kg

So, $\bar{x}_1 - \bar{x}_2 = 7.2 - 4.0 = 3.2$ kg

and $$s.e.(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{(3.7)^2}{42} + \frac{(3.9)^2}{47}} = 0.81$$

Recall from Chapter 10 (p. 408)

A **confidence interval** or **interval estimate** for any of the five parameters can be expressed as

Sample estimate ± multiplier × standard error

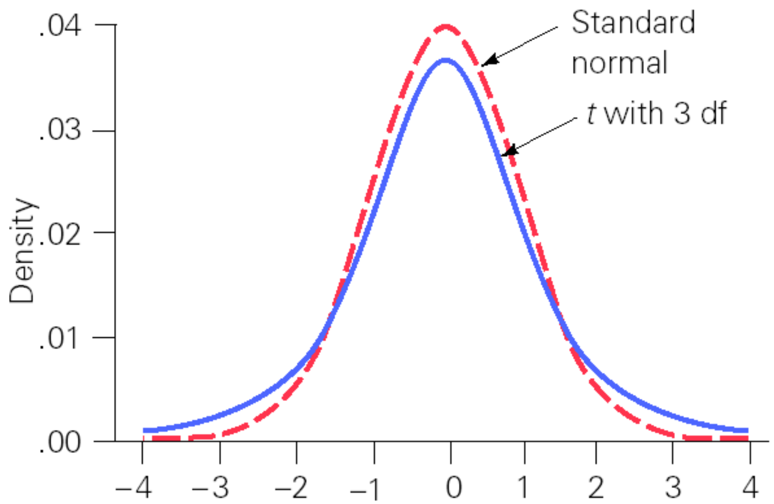
The **multiplier** is a number based on the confidence level desired and determined from the standard normal distribution (for proportions) or Student's *t*-distribution (for means).

Student's *t*-Distribution: Replacing σ with *s*

Dilemma: we generally don't know σ. Using *s* we have:

$$t = \frac{\bar{x} - \mu}{s.e.(\bar{x})} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{\sqrt{n}(\bar{x} - \mu)}{s}$$

If the sample size *n* is small, this standardized statistic will not have a $N(0,1)$ distribution but rather a ***t*-distribution** with ***n* – 1 *degrees of freedom* (df).**

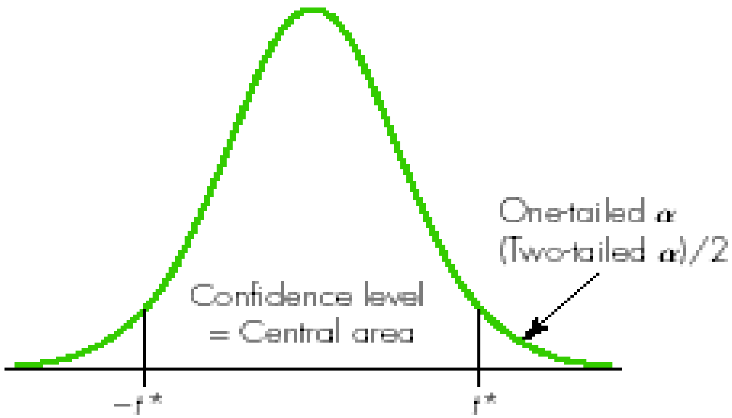


Finding the *t*-multiplier

- Excel: See page 452.
- R Commander:

  Distributions → Continuous distributions → t distribution → t quantiles
  - Probabilities: $\alpha/2$ (for 95%, use .025)
  - Degrees of freedom
  - Lower tail
  - Gives negative of the t-multiplier
  - Ex: .025, 25, lower tail → −2.059539, multiplier ≈ 2.06
- Table A2, page 728 (easy to use compared to z!)

Table A.2 t^* Multipliers for Confidence Intervals and Rejection Region Critical Values



	Confidence Level						
df	*.80*	*.90*	*.95*	*.98*	*.99*	*.998*	*.999*
1	3.08	6.31	12.71	31.82	63.66	318.31	636.62
2	1.89	2.92	4.30	6.96	9.92	22.33	31.60
3	1.64	2.35	3.18	4.54	5.84	10.21	12.92
4	1.53	2.13	2.78	3.75	4.60	7.17	8.61
5	1.48	2.02	2.57	3.36	4.03	5.89	6.87
6	1.44	1.94	2.45	3.14	3.71	5.21	5.96
7	1.41	1.89	2.36	3.00	3.50	4.79	5.41
8	1.40	1.86	2.31	2.90	3.36	4.50	5.04
9	1.38	1.83	2.26	2.82	3.25	4.30	4.78
10	1.37	1.81	2.23	2.76	3.17	4.14	4.59

Confidence Interval for One Mean *or* Paired Data

A Confidence Interval for a Population Mean

$$\bar{x} \pm t^* \times s.e.(\bar{x}) \Rightarrow \bar{x} \pm t^* \times \frac{s}{\sqrt{n}}$$

where the **multiplier *t****** is the value in a *t*-distribution with degrees of freedom = df = *n* - 1 such that the area between -*t** and *t** equals the desired confidence level. (Found from Excel, R Commander or Table A.2.)

Conditions:

- Population of measurements **is bell-shaped** (no major skewness or outliers) and r.s. of any size > 2; OR
- Population of measurements is ***not*** bell-shaped, **but** a ***large*** random **sample** is measured, $n \geq 30$.

Example: 95% C.I. for Mean Body Temperature

Data from a blood bank near Seattle (from friend of Dr. Utts).

n = 101 blood donors, ages 17 to 84 (*large* sample)

Sample mean = 97.89 degrees

Sample standard deviation = 0.73 degrees

Standard error of the mean = $\frac{0.73}{\sqrt{101}} = 0.073$

Multiplier = t^* with df of 100 = 1.98 (from Table A.2)

Sample estimate ± multiplier × standard error
97.89 ± 1.98 × 0.073
97.89 ± 0.14
97.75 to 98.03 degrees (does *not* cover 98.6)

Paired Data Confidence Interval

Data: two variables for n individuals or pairs; use the difference $d = x_1 - x_2$.

Population parameter: μ_d = mean of differences for the population (same as $\mu_1 - \mu_2$).

Sample estimate: $\bar{d}$ = sample mean of the differences

Standard deviation and standard error:

s_d = standard deviation of the sample of differences;

$$s.e.(\bar{d}) = \frac{s_d}{\sqrt{n}}$$

Confidence interval for μ_d: $\bar{d} \pm t^* \times s.e.(\bar{d})$, where df = $n - 1$ for the multiplier t^*.

Find 90% C.I. for difference: (daughter – mother) height difference

How much taller (or shorter) are daughters than their mothers these days?

n = 93 pairs (daughter – mother), $\bar{d}$ = 1.3 inches

s_d = 3.14 inches, so $s.e.(\bar{d}) = \frac{s_d}{\sqrt{n}} = \frac{3.14}{\sqrt{93}} = .33$

Multiplier = t^* with 92 df for 90% C.I. = 1.66 (use df=90)

Sample estimate ± multiplier × standard error

1.3 ± 1.66 × 0.33

1.3 ± 0.55

0.75 to 1.85 inches (does *not* cover 0)

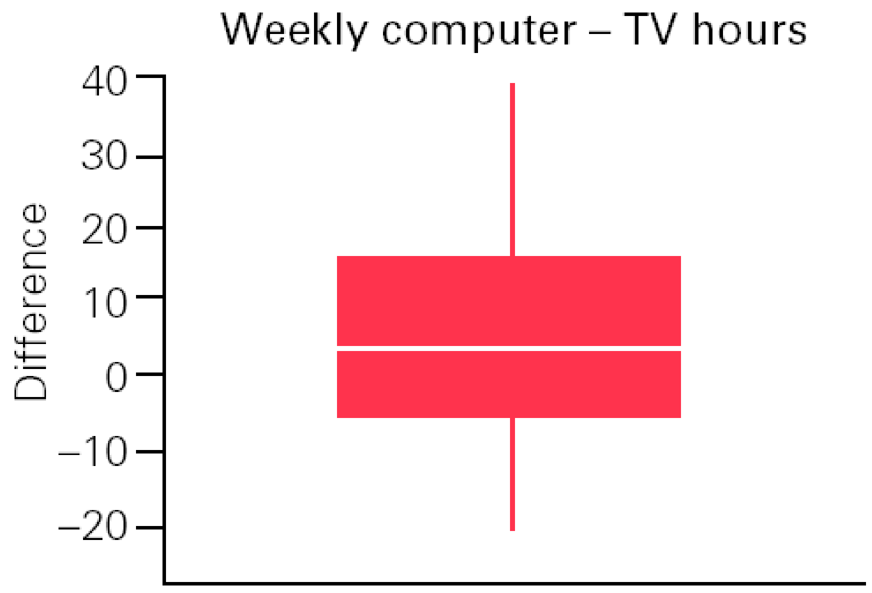
Confidence interval interpretations

- We are 95% confident that the mean body temperature for healthy adults (who would donate blood??) is between 97.75 degrees and 98.03 degrees.
- We are 90% confident that the mean height difference between female college students and their mothers is between 0.75 and 1.85 inches, with students being taller than their mothers.

Example 11.9: Small sample, so check for outliers

Data: Hours spent watching TV and hours spent on computer per week for $n = 25$ students.

Create a 90% CI for the ***mean difference*** in hours spent using computer versus watching TV.



Note: Boxplot shows no obvious skewness and no outliers.

Example 11.9 *Screen Time: Computer vs TV*

Results:

$$\bar{d} = 5.36,\ s_d = 15.24, \text{ and } s.e.(\bar{d}) = \frac{s_d}{\sqrt{n}} = \frac{15.24}{\sqrt{25}} = 3.05$$

Multiplier t^* from Table A.2 with df = 24 is $t^* = 1.71$

90% Confidence Interval:
5.36 ± 1.71(3.05) => 5.36 ± 5.22 => 0.14 to 10.58 hours

Interpretation: We are **90% confident** that the average difference between computer usage and television viewing for students represented by this sample is covered by the interval from 0.14 to 10.58 hours per week, with more hours spent on computer usage than on television viewing.

11.4 CI for Difference Between Two Means (Independent samples)

A CI for the Difference Between Two Means (Independent Samples, unpooled case):

$$\bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where ***t**** is the value in a *t*-distribution with area between -*t** and *t** equal to the desired confidence level.

Necessary Conditions

- Two samples must be **independent**.

Either …

- Populations of measurements both **bell-shaped**, and random samples of any size are measured.

or …

- **Large** ($n \geq 30$) random samples are measured.

Degrees of Freedom

The *t*-distribution is only approximately correct and **df formula** is complicated (Welch's approx):

$$\mathrm{df} \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

Statistical software can use the above approximation, but if done ***by hand*** then use a conservative *df* = smaller of $(n_1 - 1)$ and $(n_2 - 1)$.

Example 11.11 *Effect of a Stare on Driving*

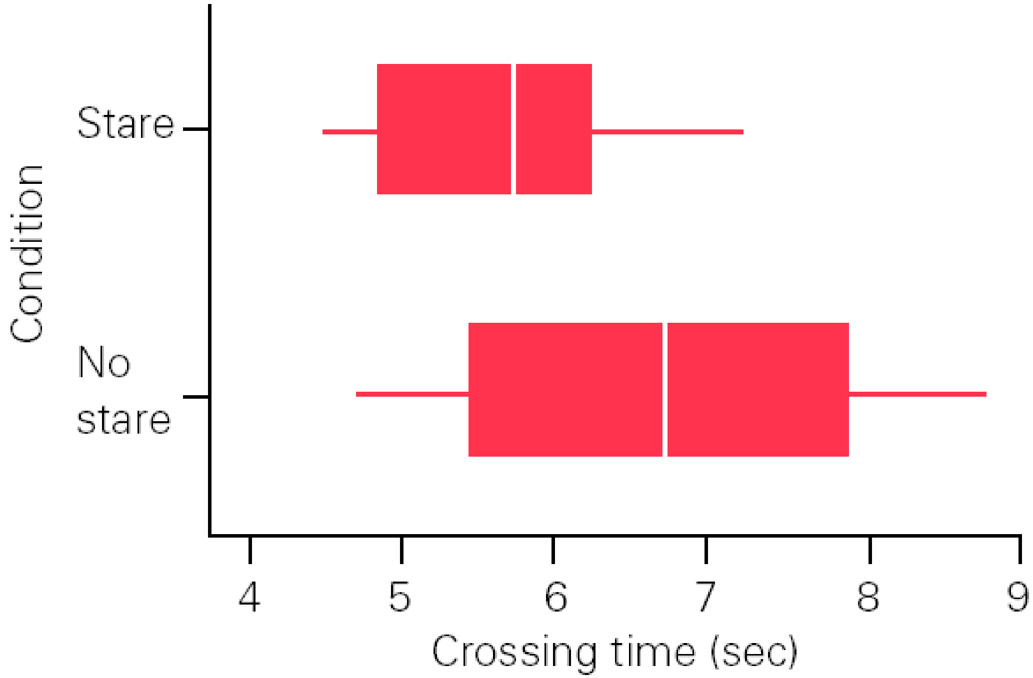
Randomized experiment: Researchers either stared or did not stare at drivers stopped at a campus stop sign; Timed how long (sec) it took driver to proceed from sign to a mark on other side of the intersection.

No Stare Group (n = 14):	8.3, 5.5, 6.0, 8.1, 8.8, 7.5, 7.8, 7.1, 5.7, 6.5, 4.7, 6.9, 5.2, 4.7
Stare Group (n = 13):	5.6, 5.0, 5.7, 6.3, 6.5, 5.8, 4.5, 6.1, 4.8, 4.9, 4.5, 7.2, 5.8

Task: Make a 95% CI for the ***difference between the mean*** crossing times for the two populations represented by these two independent samples.

Example 12.8 *Effect of a Stare on Driving*

Checking Conditions:



Boxplots show …

- No outliers and no strong skewness.
- Crossing times in **stare** group generally faster and less variable.

Example 12.8 *Effect of a Stare on Driving*

Two-sample T for CrossTime

Group	N	Mean	StDev	SE Mean
NoStare	14	6.63	1.36	0.36
Stare	13	5.59	0.822	0.23

95% CI for mu (NoStare) − mu (Stare): (0.14, 1.93)

T-Test mu (NoStare) = mu (Stare) (vs not =): T = 2.41 P = 0.025 DF = 21

Note: The df = 21 was reported by the computer package based on the Welch's approximation formula.

The **95% confidence interval** for the difference between the population means is 0.14 seconds to 1.93 seconds .

Approximate 95% CI

For sufficiently large samples, the interval

Sample estimate ± 2 × Standard error

is an **approximate 95% confidence interval** for a population parameter.

Note: Except for very small degrees of freedom, the multiplier t^* for 95% confidence interval is close to 2. So, 2 is often used to approximate, rather than finding degrees of freedom. For instance, for 95% C.I.:

$t^*(30) = 2.04$, $t^*(60) = 2.00$, $t^*(90) = 1.99$, $t^*(\infty) = z^* = 1.96$

Example 11.13 *Diet vs Exercise*

Study: $n_1 = 42$ men on diet, $n_2 = 47$ men exercise

Diet: Lost an average of 7.2 kg with std dev of 3.7 kg
Exercise: Lost an average of 4.0 kg with std dev of 3.9 kg

So, $\bar{x}_1 - \bar{x}_2 = 7.2 - 4.0 = 3.2$ kg and $s.e.(\bar{x}_1 - \bar{x}_2) = 0.81$ kg

Approximate 95% Confidence Interval:
3.2 ± 2(.81) => 3.2 ± 1.62 => 1.58 to 4.82 kg

Note: We are 95% confident the interval 1.58 to 4.82 kg covers the higher mean weight loss for dieters compared to those who exercised. The ***interval does not cover 0***, so a real difference is likely to hold for the population as well.

Using R Commander

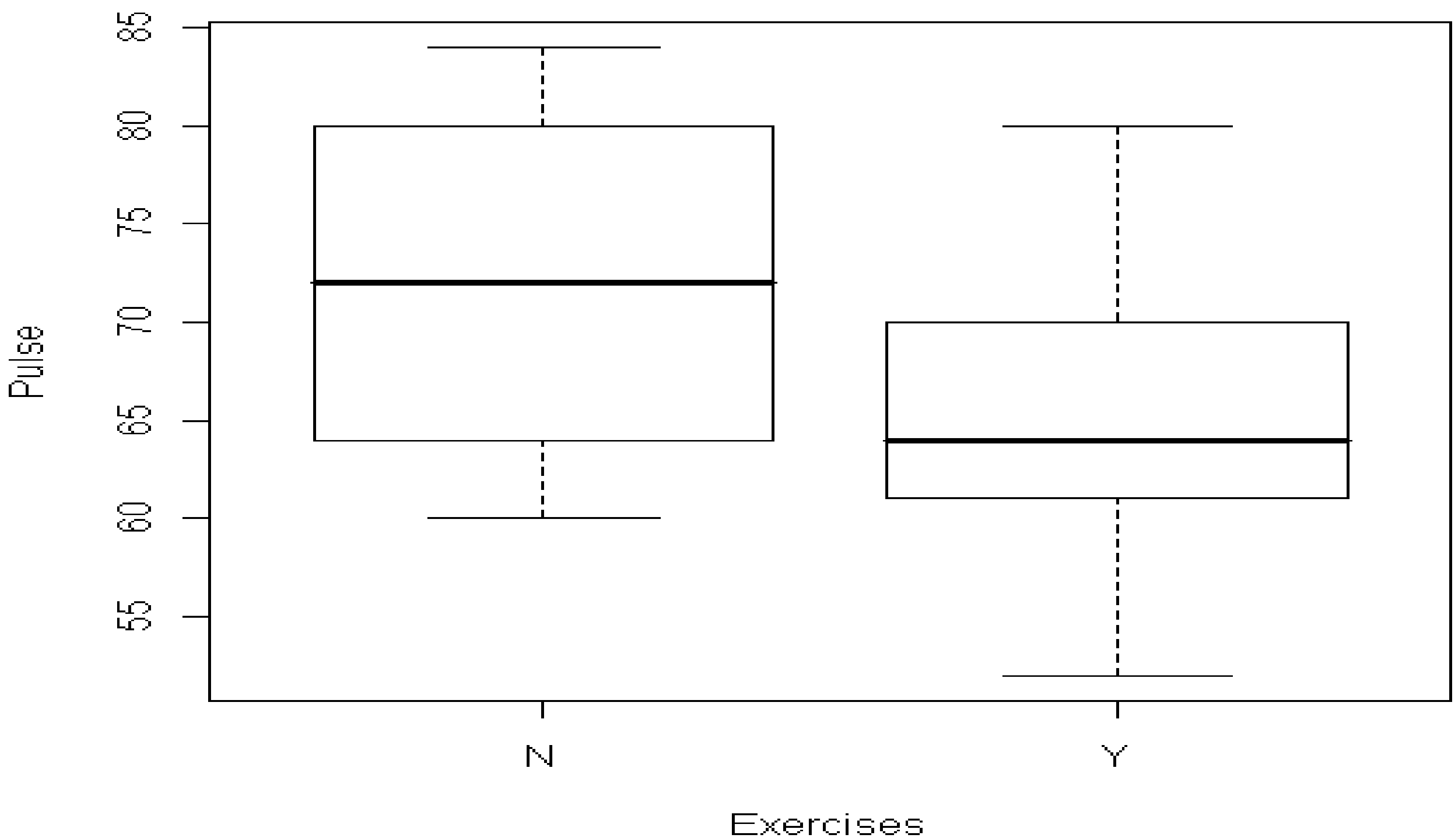
- Read in or enter data set
- Statistics → Means →
  - Single sample t-test
  - Independent samples t-test (requires data in one column, and group code in another)
  - Paired t-test (requires the original data in two separate columns)

Example: Exercise 11.51
Two-sample t-test to compare pulse

- Data → New data set – give name, enter data
- One column for Exercise (Y,N) one for pulse
- Statistics → Means →Independent samples t-test
- Choose the alternative (≠,>,<) and conf. level

```
data:  Pulse by Exercises
t = 1.7156, df = 13.014, p-value = 0.05496
alternative hypothesis: true difference in means is
  not equal to 0
95 percent confidence interval:
 -1.727387 15.060720
sample estimates:
mean in group N mean in group Y
       72.00000        65.33333
```

Should check for outliers with small sample(s)
Graphs→Boxplot→Plot by Group
This looks okay – no outliers or major skew



Confidence interval applet

http://www.rossmanchance.com/applets/Confsim/Confsim.html