Class 22

Daniel B. Rowe, Ph.D.

Department of Mathematical and Statistical Sciences

MARQUETTE UNIVERSITY

Be The Difference.

Agenda:

Recap Chapter 10.2-10.3

Lecture Chapter 10.4-10.5

Recap Chapter 10.2-10.3

10: Inferences Involving Two Populations

10.2 Inference for Mean Difference Two Dependent Samples

Confidence Interval Procedure

Paired Difference $$d = x_1 - x_2 \qquad (10.1)$$

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i \qquad s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2 \qquad \mu_{\bar{d}} = \mu_d \quad \sigma_{\bar{d}} = \frac{\sigma_d}{\sqrt{n}}$$

With σ_d unknown, a 1-α confidence interval for μ_d=(μ_1-μ_2) is:

Confidence Interval for Mean Difference (Dependent Samples)

$$\bar{d} - t(df, \alpha/2)\frac{s_d}{\sqrt{n}} \quad \text{to} \quad \bar{d} + t(df, \alpha/2)\frac{s_d}{\sqrt{n}} \quad \text{where } df = n-1 \qquad (10.2)$$

10: Inferences Involving Two Populations

10.2 Inference for Mean Difference Two Dependent Samples

Car	1	2	3	4	5	6
Brand A	125	64	94	38	90	106
Brand B	133	65	103	37	102	115

Example:

Construct a 95% CI for mean difference in Brand B – A tire wear.

d_i's: $8,\ 1,\ 9,\ -1,\ 12,\ 9$

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

$n = 6$ $\quad df = 5$ $\quad t(df,\alpha/2) = 2.57$

$\bar{d} = 6.3$ $\quad \alpha = 0.05$

$$s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

$s_d = 5.1$

$$\bar{d} \pm t(df,\alpha/2)\frac{s_d}{\sqrt{n}} \longrightarrow (0.090, 11.7)$$

Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.2 Inference for Mean Difference Two Dependent Samples

$n=6 \quad 8,\ 1,\ 9,\ -1,\ 12,\ 9$

Car	1	2	3	4	5	6
Brand A	125	64	94	38	90	106
Brand B	133	65	103	37	102	115

Example:

Test mean difference of Brand B minus Brand A is zero.

Step 1 H_0: $\mu_d=0$ vs. H_a: $\mu_d \neq 0$

Step 2

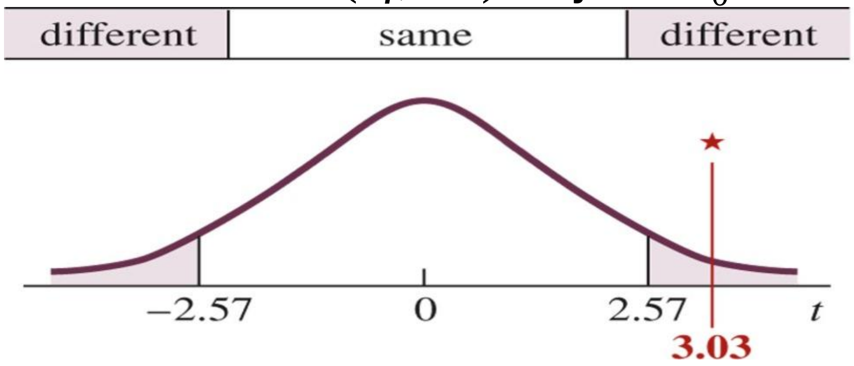
$$df = 5 \qquad t^* = \frac{\bar{d} - \mu_{d0}}{s_d / \sqrt{n}}$$
$$\alpha = .05$$

Step 3 $\bar{d} = 6.3$, $s_d = 5.1$

$$t^* = \frac{6.3 - 0}{5.1/\sqrt{6}} = 3.03$$

Step 4 $t(df, \alpha/2) = 2.57$

Step 5 Since $t^* > t(df,\alpha/2)$, reject H_0



Conclusion: Significant difference in tread wear at .05 level.

Figures from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.3 Inference for Mean Difference Two Independent Samples

Confidence Interval Procedure

With σ_1 and σ_2 unknown, a 1-α confidence interval for $\mu_1 - \mu_2$ is:

Confidence Interval for Mean Difference (Independent Samples)

$$(\bar{x}_1 - \bar{x}_2) - t(df, \alpha/2)\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)} \text{ to } (\bar{x}_1 - \bar{x}_2) + t(df, \alpha/2)\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}$$

where df is either calculated or smaller of df_1, or df_2 (10.8)

Actually, this is for $\sigma_1 \neq \sigma_2$.

Next larger number than

$$df = \left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2 \Bigg/ \left(\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}\right)$$

If using a computer program.

If not using a computer program.

9: Inferences Involving One Population

9.2 Inference about the Binomial Probability of Success

There are three possible hypothesis pairs for the difference in means.

H_0: μ_1-μ_2 ≥ 0 vs. H_a: μ_1-μ_2 < 0

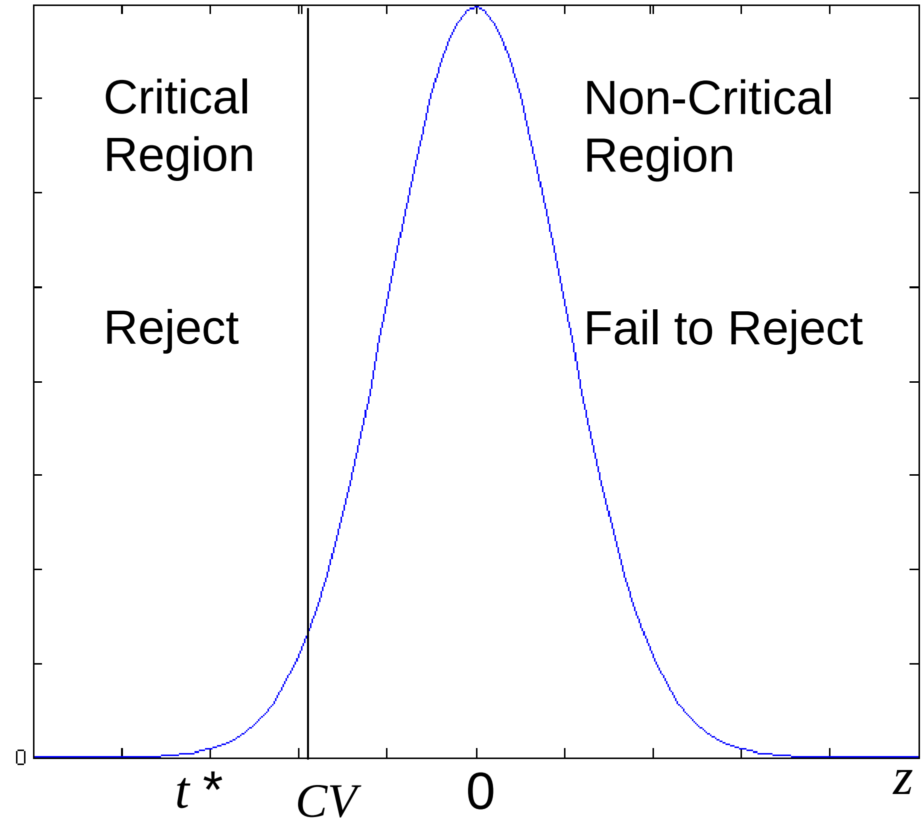
Reject H_0 if

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}}$$

less than

$$-t(df, \alpha)$$

data indicates μ_1-μ_2 < 0 because $\bar{x}_1 - \bar{x}_2$ is “a lot” smaller than 0.



9: Inferences Involving One Population

9.2 Inference about the Binomial Probability of Success

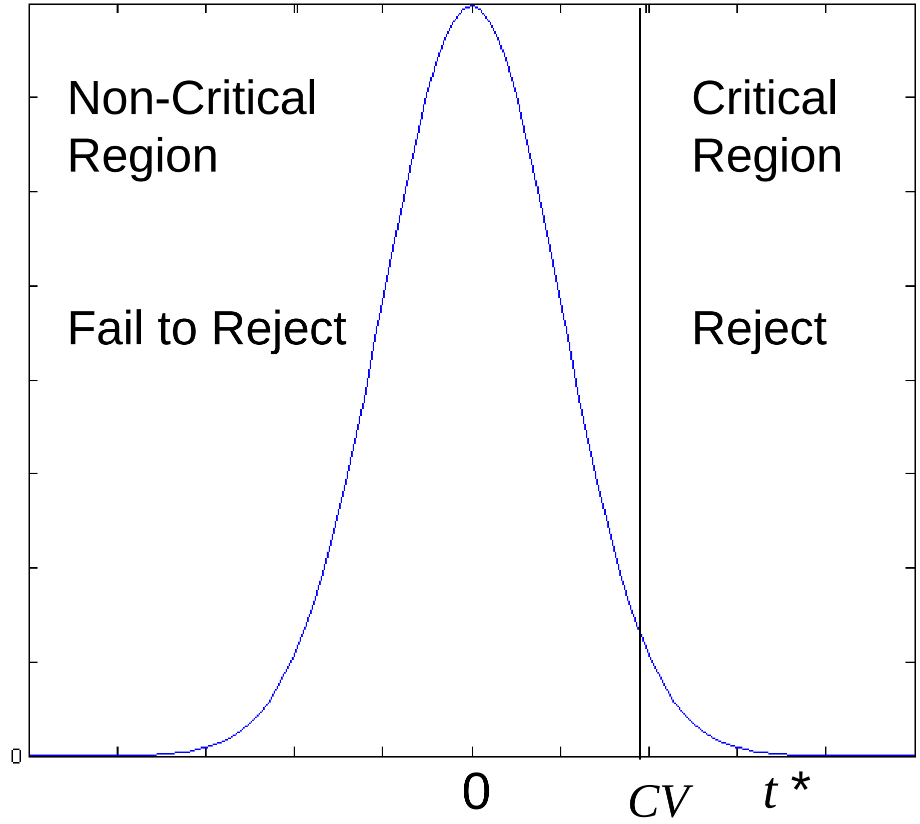
There are three possible hypothesis pairs for the difference in means.

H_0: μ_1-μ_2 ≤ 0 vs. H_a: μ_1-μ_2 > 0

Reject H_0 if greater then

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\left(s_1^2/n_1\right) + \left(s_2^2/n_2\right)}} \qquad t(df, \alpha)$$

data indicates μ_1-μ_2 > 0 because $\bar{x}_1 - \bar{x}_2$ is "a lot" larger than 0.



Rowe, D.B.

9: Inferences Involving One Population

9.2 Inference about the Binomial Probability of Success

There are three possible hypothesis pairs for the difference in means.

H_0: μ_1-μ_2 = 0 vs. H_a: μ_1-$\mu_2 \neq 0$

Reject H_0 if

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\left(s_1^2/n_1\right) + \left(s_2^2/n_2\right)}}$$

less than

$$-t(df, \alpha/2)$$

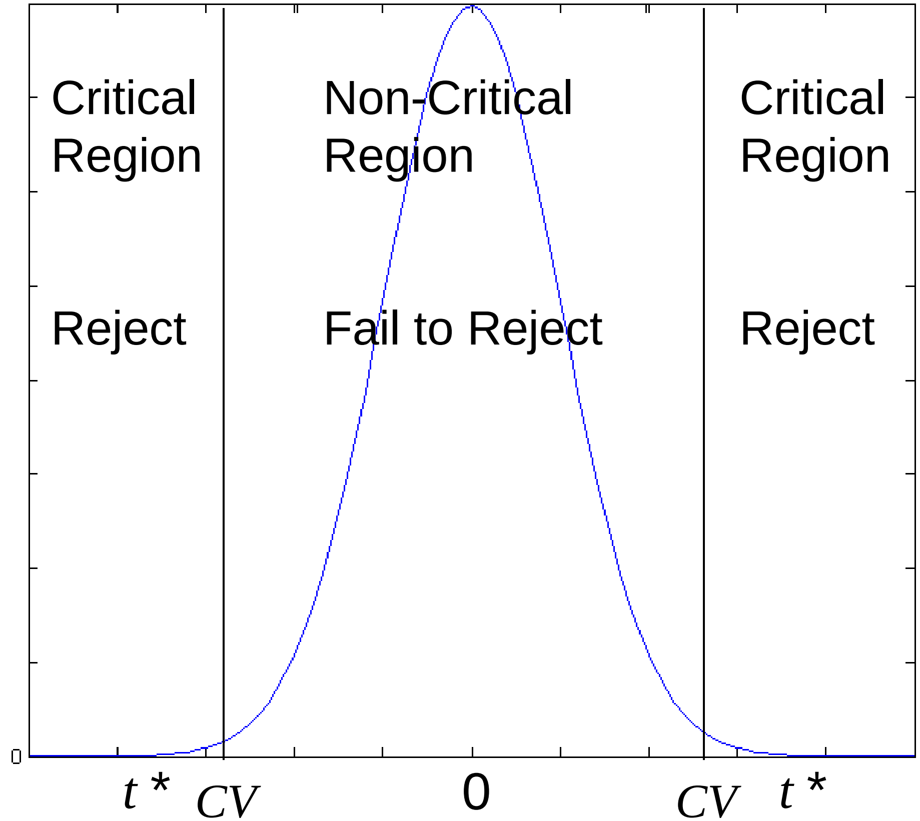
or if

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\left(s_1^2/n_1\right) + \left(s_2^2/n_2\right)}}$$

is greater than

$$t(df, \alpha/2)$$

data indicates μ_1-$\mu_2 \neq 0$, $\bar{x}_1 - \bar{x}_2$ far from 0.



10: Inferences Involving Two Populations

10.3 Inference Mean Difference Confidence Interval

Sample	Number	Mean	Standard Deviation
Female (f)	$n_f = 20$	$\bar{x}_f = 63.8$	$s_f = 2.18$
Male (m)	$n_m = 30$	$\bar{x}_m = 69.8$	$s_m = 1.92$

Example:

Interested in difference in mean heights between men and women. The heights of 20 females and 30 males is measured. Construct a 95% confidence interval for $\mu_m - \mu_f$, σ_m & σ_f unknown

$$(\bar{x}_m - \bar{x}_f) \pm t(df, \alpha/2)\sqrt{\left(\frac{s_m^2}{n_m}\right) + \left(\frac{s_f^2}{n_f}\right)}$$

$$\alpha = 0.05$$

$$t(19, .025) = 2.09$$

$$(69.8 - 63.8) \pm 2.09\sqrt{\left(\frac{(1.92)^2}{30}\right) + \left(\frac{(2.18)^2}{20}\right)}$$

therefore 4.75 to 7.25

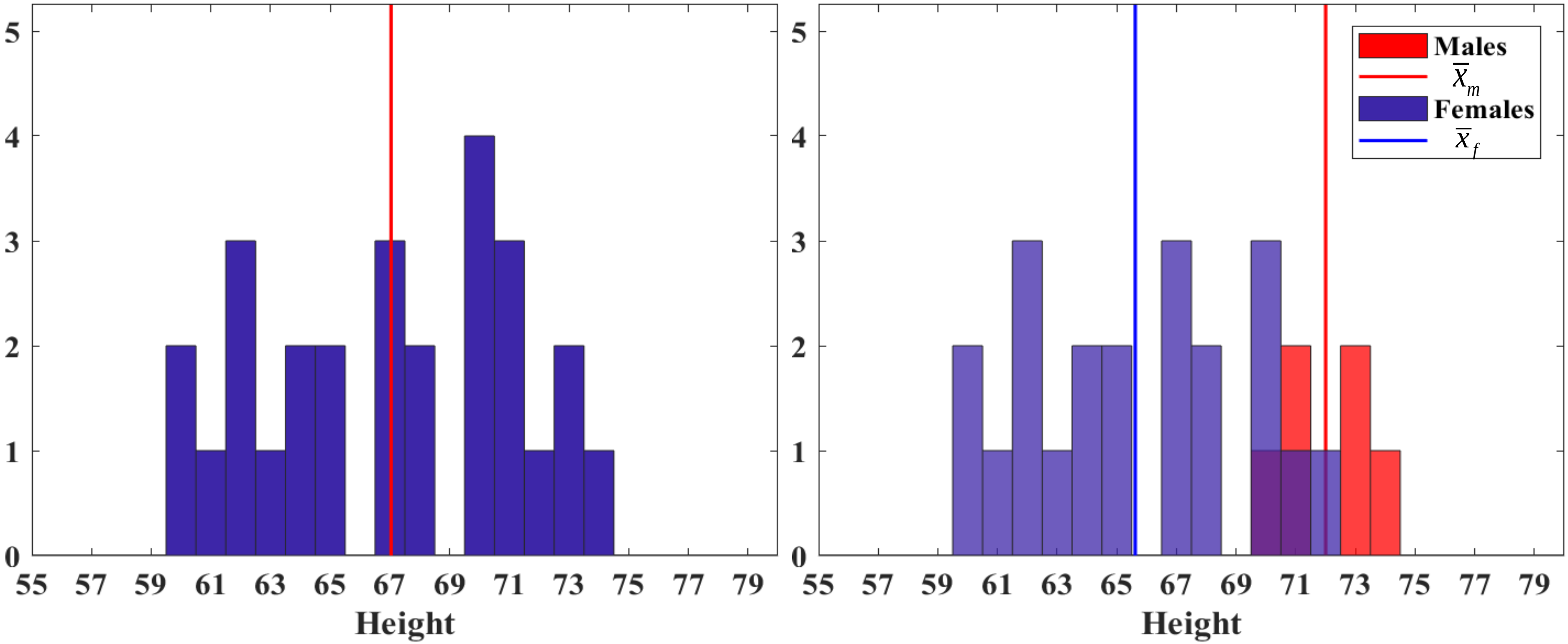
Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.3 Inference for Mean Difference Two Independent Samples

Hypothesis Testing Procedure

27 values



Is the height of males = height of females at α=.05?

10: Inferences Involving Two Populations

10.3 Inference for Mean Difference Two Independent Samples

Hypothesis Testing Procedure

27 values

Step 1
H_0: $\mu_f = \mu_m$ vs. H_a: $\mu_f \neq \mu_m$

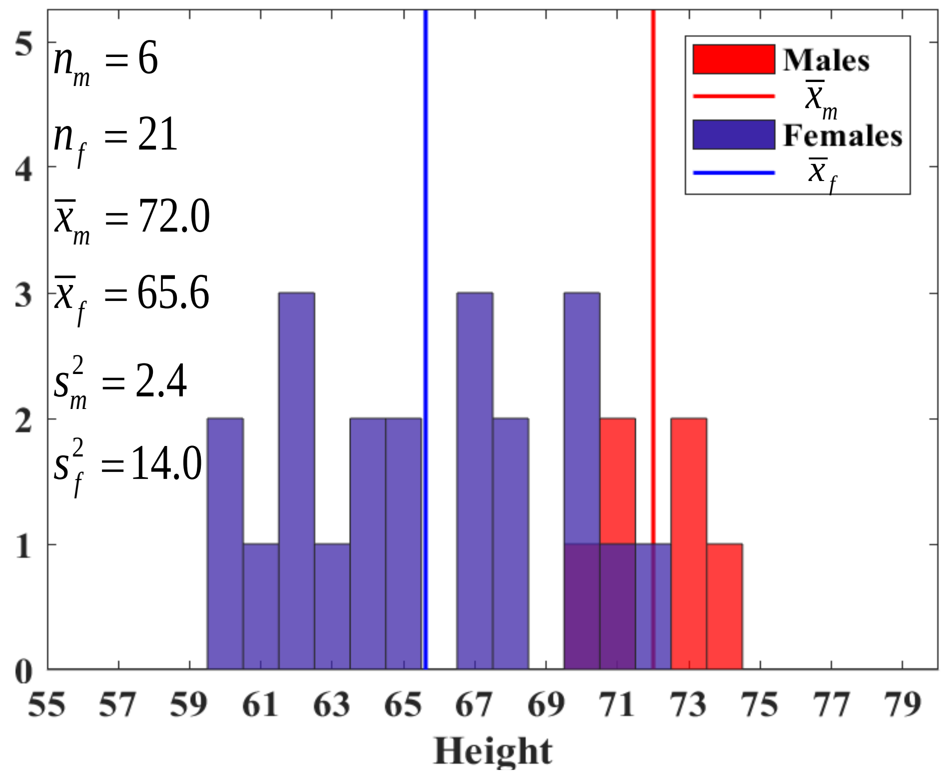
Step 2

$$t^* = \frac{(\bar{x}_m - \bar{x}_f) - (\mu_{0,m} - \mu_{0,f})}{\sqrt{\left(\frac{s_m^2}{n_m}\right) + \left(\frac{s_f^2}{n_f}\right)}}$$

$df = 5$
$\alpha = .05$

Step 3

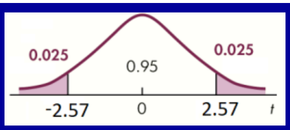
$$t^* = \frac{(72.0 - 65.6) - (0)}{\sqrt{\left(\frac{2.4}{6}\right) + \left(\frac{14.0}{21}\right)}} = 6.17$$



Classical
Step 4

$t(df, \alpha/2) = 2.57$

Classical
Step 5 Reject H_0 $6.17 > 2.57$, height males≠height females



10: Inferences Involving Two Populations

10.3 Inference for Mean Difference Two Independent Samples

Hypothesis Testing Procedure

27 values

Step 1

H_0: $\mu_f = \mu_m$ vs. H_a: $\mu_f \neq \mu_m$

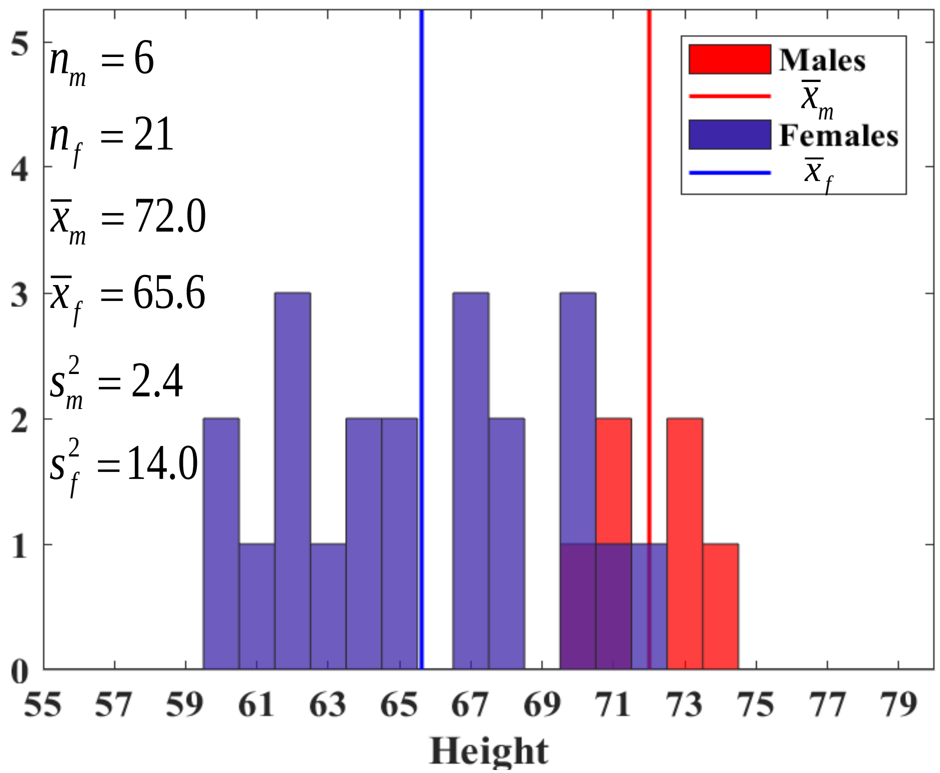
Step 2

$$t^* = \frac{(\bar{x}_m - \bar{x}_f) - (\mu_{0,m} - \mu_{0,f})}{\sqrt{\left(\frac{s_m^2}{n_m}\right) + \left(\frac{s_f^2}{n_f}\right)}}$$

$df = 5$
$\alpha = .05$

Step 3

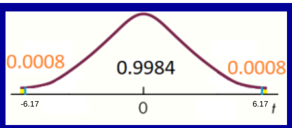
$$t^* = \frac{(72.0 - 65.6) - (0)}{\sqrt{\left(\frac{2.4}{6}\right) + \left(\frac{14.0}{21}\right)}} = 6.17$$



p-value
Step 4

$P(|t^*| > 6.17) = .0016$

p-value
Step 5 Reject H_0 $.002 < .05$, height males≠height females



Chapter 10: Inferences Involving Two Populations

Questions?

Homework: Read Chapter 10.3

WebAssign

Chapter 10 # 41, 45, 53, 57, 58, 59, 63

Lecture Chapter 10.4-10.5

9: Inferences Involving One Population

9.2 Inference about the Binomial Probability of Success

Chapter 5

We talked about a Binomial experiment with two outcomes.
Each performance of the experiment is called a trial.
Each trial is independent.

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$n = 1, 2, 3, \ldots$$
$$0 \le p \le 1$$
$$x = 0, 1 \ldots, n$$

n = number of trials or times we repeat the experiment.
x = the number of successes out of n trials.
p = the probability of success on an individual trial.

9: Inferences Involving One Population

9.2 Inference about the Binomial Probability of Success

When we perform a binomial experiment we can estimate the probability of heads as

Sample Binomial Probability

i.e. number of H out of n flips

$$p' = \frac{x}{n} \qquad (9.3)$$

where x is the number of successes in n trials.

This is a point estimate. Recall the rule for a CI is

point estimate $\pm$ some amount

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Background

For Binomial, where x is number of successes out of n trials.

We said that $\text{mean}(cx) = cnp$ and $\text{variance}(cx) = c^2 npq$.

$\longrightarrow$ $\text{mean}(x/n) = p$ and $\text{variance}(x/n) = pq/n$. q=1-p

We are often interested in comparisons between proportions $p_1 - p_2$. There is another rule that says that if x_1 and x_2 are random variables, then $\text{mean}(x_1 \pm x_2) = \text{mean}(x_1) \pm \text{mean}(x_2)$

further, $\text{mean}\left(\frac{x_1}{n_1} \pm \frac{x_2}{n_2}\right) = \text{mean}\left(\frac{x_1}{n_1}\right) \pm \text{mean}\left(\frac{x_2}{n_2}\right)$

and $\text{variance}\left(\frac{x_1}{n_1} \pm \frac{x_2}{n_2}\right) = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$.

if x_1 & x_2 independent

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

That is where 1. and 2. in the green box below come from

If independent samples of size n_1 and n_2 are drawn … with p_1=P_1(success) and p_2=P_2(success) , then the sampling distribution of $p_1' - p_2'$ has these properties:

1. mean $\mu_{p_1'-p_2'} = p_1 - p_2$

2. standard error $\sigma_{p_1'-p_2'} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$ (10.10)

3. approximately normal dist if n_1 and n_2 are sufficiently large.

ie **I** n_1,n_2>20 **II** n_1p_1, n_1q_1, n_2p_2, n_2q_2>5 **III** sample<10% of pop

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Confidence Interval Procedure

Assumptions for … difference between two proportions p_1-p_2: The n_1 … and n_2 random observations … are selected independently from two populations that are not changing

Confidence Interval for the Difference between Two Proportions p_1 - p_2

$$(p_1' - p_2') - z(\alpha/2)\sqrt{\frac{p_1'q_1'}{n_1} + \frac{p_2'q_2'}{n_2}} \quad \text{to} \quad (p_1' - p_2') + z(\alpha/2)\sqrt{\frac{p_1'q_1'}{n_1} + \frac{p_2'q_2'}{n_2}}$$

where $p_1' = \frac{x_1}{n_1}$ and $p_2' = \frac{x_2}{n_2}$. (10.11)

$q_1' = 1 - p_1' \quad q_2' = 1 - p_2'$

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Confidence Interval Procedure

Example: Another Semester

Construct a 99% CI for proportion of female A's minus male A's difference $p_f - p_m$. Fill in.

120 values

$n_m = 52$

$n_f = 68$

$x_m = 21$

$x_f = 43$

$$z(\alpha/2) =$$

$$p'_f = \frac{x_f}{n_f} =$$

$$p'_m = \frac{x_m}{n_m} =$$

$$(p'_f - p'_m) \pm z(\alpha/2)\sqrt{\frac{p'_f q'_f}{n_f} + \frac{p'_m q'_m}{n_m}}$$

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Confidence Interval Procedure

$\alpha = 0.01$

Example:

Construct a 99% CI for proportion of female A's minus male A's difference $p_f - p_m$.

120 values

$n_m = 52$

$n_f = 68$

$x_m = 21$

$x_f = 43$

$$z(\alpha/2) = 2.58$$

$$p'_f = \frac{x_f}{n_f} = \frac{43}{68} = .62$$

$$p'_m = \frac{x_m}{n_m} = \frac{21}{52} = .40$$

$$(p'_f - p'_m) \pm z(\alpha/2)\sqrt{\frac{p'_f q'_f}{n_f} + \frac{p'_m q'_m}{n_m}}$$

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Confidence Interval Procedure

$\alpha = 0.01$

Top 5 of 6 exams.

Example:

Construct a 99% CI for proportion of female A's minus male A's difference $p_f - p_m$ for a previous class.

120 values

$n_m = 52$

$n_f = 68$

$x_m = 21$

$x_f = 43$

$z(\alpha/2) = 2.58$

$$p'_f = \frac{x_f}{n_f} = \frac{43}{68} = .62$$

$$p'_m = \frac{x_m}{n_m} = \frac{21}{52} = .40$$

$$(p'_f - p'_m) \pm z(\alpha/2)\sqrt{\frac{p'_f q'_f}{n_f} + \frac{p'_m q'_m}{n_m}}$$

$$(.62 - .40) \pm 2.58\sqrt{\frac{(.62)(.38)}{68} + \frac{(.40)(.60)}{52}}$$

$$-.003 \text{ to } .460$$

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Hypothesis Testing Procedure

We can perform hypothesis tests on the proportion

H_0: $p_1 \geq p_2$ vs. H_a: $p_1 < p_2$

H_0: $p_1 \leq p_2$ vs. H_a: $p_1 > p_2$

H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$

$$\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2} = pq\left[\frac{1}{n_1} + \frac{1}{n_2}\right]$$

when $p_1 = p_2 = p$.

Test Statistic for the Difference between two Proportions-
Population Proportions **Known**

$$z^* = \frac{(p_1' - p_2') - \overbrace{(p_{0,1} - p_{0,2})}^{0}}{\sqrt{\underbrace{pq}_{p \text{ known}}\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

$$p_1' = \frac{x_1}{n_1} \quad p_2' = \frac{x_2}{n_2} \qquad (10.12)$$

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Hypothesis Testing Procedure

Test Statistic for the Difference between two Proportions-Population Proportions **UnKnown**

$$z^* = \frac{(p_1' - p_2') - \overbrace{(p_{0,1} - p_{0,2})}^{0}}{\sqrt{\underbrace{p_p'}_{p_p \text{ estimated}} q_p' \left[\dfrac{1}{n_1} + \dfrac{1}{n_2}\right]}} \qquad (10.15)$$

where we assume $p_1=p_2$ and use pooled estimate of proportion

$$p_1' = \frac{x_1}{n_1} \quad p_2' = \frac{x_2}{n_2} \qquad \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2} = pq\left[\frac{1}{n_1} + \frac{1}{n_2}\right] \qquad p_p' = \frac{x_1 + x_2}{n_1 + n_2} \qquad q_p' = 1 - p_p'$$

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Is proportion of Salesman's defectives less than Competitor's? $\alpha = .05$

Step 1 Fill in.

Product	Number Defective	Number Checked
Salesperson's	$x_s = 15$	$n_s = 150$
Competitor's	$x_c = 6$	$n_c = 150$

Step 2

Step 3

Step 4

Step 5

Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Is proportion of Salesman's defectives less than Competitor's? $\alpha = .05$

Step 1

H_0: p_s-p_c≤0 vs. H_a: p_s-p_c>0

Step 2

Product	Number Defective	Number Checked
Salesperson's	$x_s = 15$	$n_s = 150$
Competitor's	$x_c = 6$	$n_c = 150$

Step 3

Step 4

Step 5

Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Is proportion of Salesman's defectives less than Competitor's? $\alpha = .05$

Product	Number Defective	Number Checked
Salesperson's	$x_s = 15$	$n_s = 150$
Competitor's	$x_c = 6$	$n_c = 150$

Step 1

H_0: p_s-p_c≤0 vs. H_a: p_s-p_c>0

Step 2

$$z^* = \frac{(p'_s - p'_c) - (p_{0,s} - p_{0,c})}{\sqrt{p'_p q'_p \left[\frac{1}{n_s} + \frac{1}{n_c}\right]}}$$

$\alpha = .05$

Step 3

Step 4

Step 5

Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Is proportion of Salesman's defectives less than Competitor's? $\alpha = .05$

Step 1

H_0: p_s-p_c≤0 vs. H_a: p_s-p_c>0

Product	Number Defective	Number Checked
Salesperson's	$x_s = 15$	$n_s = 150$
Competitor's	$x_c = 6$	$n_c = 150$

Step 2

$$z^* = \frac{(p'_s - p'_c) - (p_{0,s} - p_{0,c})}{\sqrt{p'_p q'_p \left[\frac{1}{n_s} + \frac{1}{n_c}\right]}}$$

$\alpha = .05$

$$p'_s = \frac{x_s}{n_s} = \frac{15}{150} \qquad p'_c = \frac{x_c}{n_c} = \frac{6}{150}$$

$$p'_p = \frac{x_s + x_c}{n_s + n_c} = \frac{15+6}{150+150}$$

Step 3

$$z^* = \frac{(.10 - .04) - (0)}{\sqrt{(.07)(.93)\left[\frac{1}{150} + \frac{1}{150}\right]}} = 2.04$$

Step 4

Step 5

Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Is proportion of Salesman's defectives less than Competitor's? $\alpha = .05$

Step 1
H_0: p_s-p_c≤0 vs. H_a: p_s-p_c>0

Step 2

$$z^* = \frac{(p'_s - p'_c) - (p_{0,s} - p_{0,c})}{\sqrt{p'_p q'_p \left[\frac{1}{n_s} + \frac{1}{n_c} \right]}}$$

$\alpha = .05$

Product	Number Defective	Number Checked
Salesperson's	$x_s = 15$	$n_s = 150$
Competitor's	$x_c = 6$	$n_c = 150$

$$p'_s = \frac{x_s}{n_s} = \frac{15}{150} \qquad p'_c = \frac{x_c}{n_c} = \frac{6}{150}$$

$$p'_p = \frac{x_s + x_c}{n_s + n_c} = \frac{15+6}{150+150}$$

Step 3

$$z^* = \frac{(.10 - .04) - (0)}{\sqrt{(.07)(.93)\left[\frac{1}{150} + \frac{1}{150} \right]}} = 2.04$$

Step 4

$z(\alpha) = 1.65$

Step 5

Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Is proportion of Salesman's defectives less than Competitor's? $\alpha = .05$

Product	Number Defective	Number Checked
Salesperson's	$x_s = 15$	$n_s = 150$
Competitor's	$x_c = 6$	$n_c = 150$

Step 1

H_0: p_s-p_c≤0 vs. H_a: p_s-p_c>0

Step 2

$$z^* = \frac{(p'_s - p'_c) - (p_{0,s} - p_{0,c})}{\sqrt{p'_p q'_p \left[\frac{1}{n_s} + \frac{1}{n_c} \right]}}$$

$\alpha = .05$

$$p'_s = \frac{x_s}{n_s} = \frac{15}{150} \qquad p'_c = \frac{x_c}{n_c} = \frac{6}{150}$$

$$p'_p = \frac{x_s + x_c}{n_s + n_c} = \frac{15+6}{150+150}$$

Step 3

$$z^* = \frac{(.10 - .04) - (0)}{\sqrt{(.07)(.93)\left[\frac{1}{150} + \frac{1}{150} \right]}} = 2.04$$

Step 4

$z(\alpha) = 1.65$

Step 5 Reject H_0 $\;<.05$

$.02 < p-value < .023$ or $2.04 > 1.65$

Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Hypothesis Testing Procedure

We can perform hypothesis tests on two variances

$H_0 : \sigma_1^2 \geq \sigma_2^2$ vs. $H_a : \sigma_1^2 < \sigma_2^2$

$H_0 : \sigma_1^2 \leq \sigma_2^2$ vs. $H_a : \sigma_1^2 > \sigma_2^2$

$H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_a : \sigma_1^2 \neq \sigma_2^2$

Assumptions: Independent samples from normal distribution

Actually $\longleftarrow$ ignore

$$F* = \frac{\left[(n_n - 1)s_n^2 / \sigma^2\right] / (n_n - 1)}{\left[(n_d - 1)s_d^2 / \sigma^2\right] / (n_d - 1)}$$

Test Statistic for Equality of Variances

$$F* = \frac{s_n^2}{s_d^2} \quad \text{with } df_n = n_n - 1 \text{ and } df_d = n_d - 1. \qquad (10.16)$$

Use new table to find areas for new statistic.

10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Properties of F distribution

1. F is non-negative
2. F is nonsymmetrical
3. F is a family of dists.

$df_n = \nu_n = n_n - 1, df_d = \nu_d = n_d - 1.$

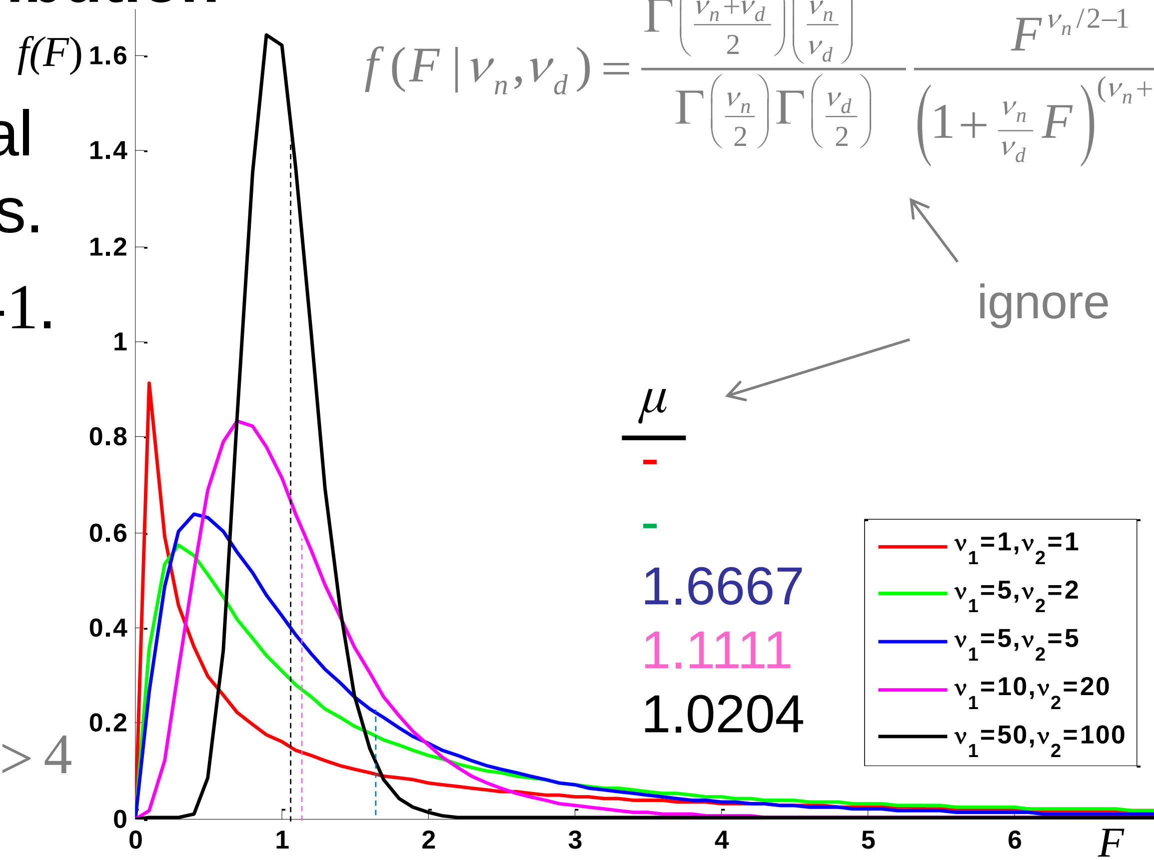
$$f(F \mid \nu_n, \nu_d) = \frac{\Gamma\left(\frac{\nu_n+\nu_d}{2}\right)\left(\frac{\nu_n}{\nu_d}\right)^{\nu_n/2}}{\Gamma\left(\frac{\nu_n}{2}\right)\Gamma\left(\frac{\nu_d}{2}\right)} \frac{F^{\nu_n/2-1}}{\left(1+\frac{\nu_n}{\nu_d}F\right)^{(\nu_n+\nu_d)/2}}$$

ignore

ignore

$$\mu = \frac{\nu_d}{\nu_d - 2} \quad , \nu_d > 2$$

$$\sigma^2 = \frac{2\nu_d^2(\nu_n + \nu_d - 2)}{\nu_n(\nu_d - 2)^2(\nu_d - 4)} , \nu_2 > 4$$



10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Hypothesis Testing Procedure

Test Statistic for Equality of Variances

$$F^* = \frac{s_n^2}{s_d^2} \quad \text{with } df_n = n_n - 1 \text{ and } df_d = n_d - 1 \;. \qquad (10.16)$$

Will also need critical values.

$$P\left(F > F(df_n, df_d, \alpha)\right) = \alpha$$

Table 9
Appendix B
Page 722

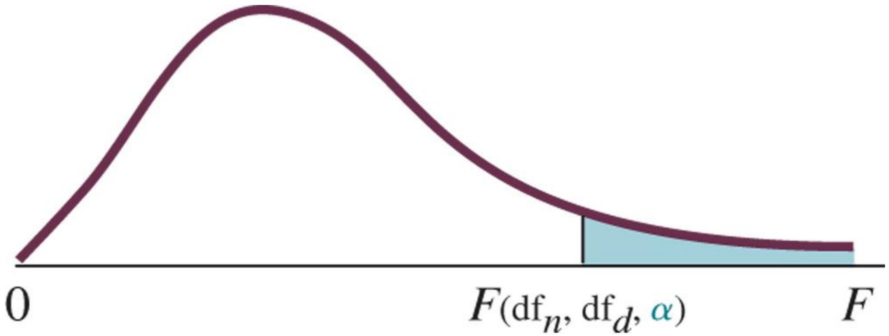
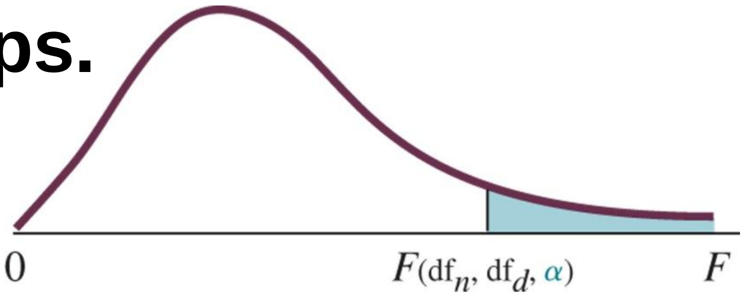


Figure from Johnson & Kuby, 2012.

10: Inferences Involving Two Pops.

10.5 Inference Ratio of Two Variances



Example: Find F(5,8,0.05).

$df_n = n_n - 1$ $df_d = n_d - 1$

Table 9, Appendix B, Page 722.

$\alpha = 0.05$

Degrees of Freedom for Numerator df_n

Degrees of Freedom for Denominator df_d	1	2	3	4	5	6	7	8	9	10
1	161.	200.	216.	225.	230.	234.	237.	239.	241.	242.
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98

Figures from Johnson & Kuby, 2012.

10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Hypothesis Testing Procedure

★

***One tailed tests*:** Arrange H_0 & H_a so H_a is always "greater than"

H_0: $\sigma_1^2 \geq \sigma_2^2$ vs. H_a: $\sigma_1^2 < \sigma_2^2$ $\longrightarrow$ H_0: $\sigma_2^2 / \sigma_1^2 \leq 1$ vs. H_a: $\sigma_2^2 / \sigma_1^2 > 1$ $\quad F^* = s_2^2 / s_1^2$

H_0: $\sigma_1^2 \leq \sigma_2^2$ vs. H_a: $\sigma_1^2 > \sigma_2^2$ $\longrightarrow$ H_0: $\sigma_1^2 / \sigma_2^2 \leq 1$ vs. H_a: $\sigma_1^2 / \sigma_2^2 > 1$ $\quad F^* = s_1^2 / s_2^2$

Reject H_0 if $F^* = s_n^2 / s_d^2 > F(df_n, df_d, \alpha)$.

***Two tailed tests*:** put larger sample variance s^2 in numerator

H_0: $\sigma_1^2 = \sigma_2^2$ vs. H_a: $\sigma_1^2 \neq \sigma_2^2$ $\longrightarrow$ H_0: $\sigma_n^2 / \sigma_d^2 = 1$ vs. H_a: $\sigma_n^2 / \sigma_d^2 \neq 1$

$\sigma_n^2 = \sigma_1^2$ if $s_1^2 > s_2^2$, $\sigma_n^2 = \sigma_2^2$ if $s_2^2 > s_1^2$

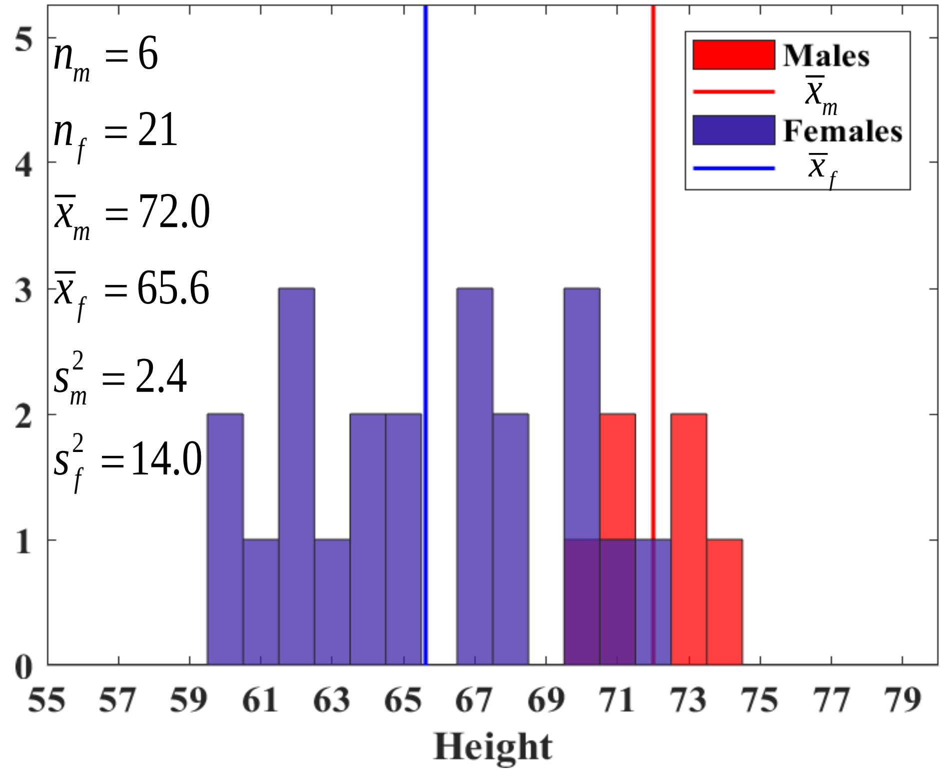
Reject H_0 if $F^* = s_n^2 / s_d^2 > F(df_n, df_d, \alpha/2)$.

10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Is variance of female heights greater than that of males? $\alpha = .01$ 27 values

Step 1 Fill in.



Step 2

Step 3

Step 4

Step 5

10: Inferences Involving Two Populations

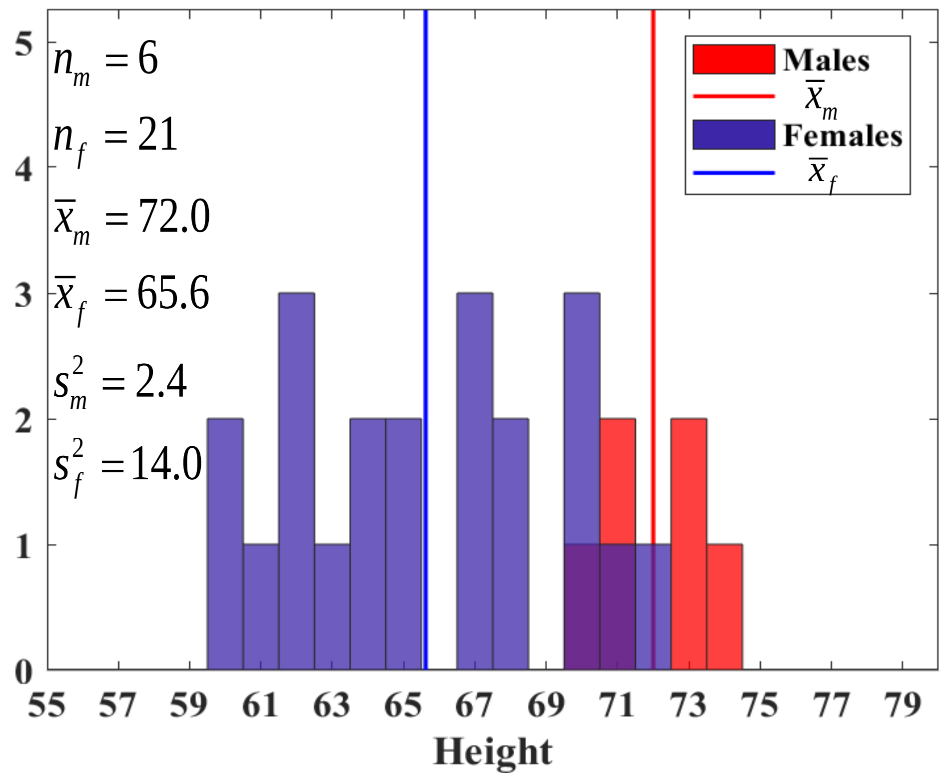
10.5 Inference for Ratio of Two Variances Two Ind. Samples

Is variance of female heights greater than that of males? $\alpha = .01$ 27 values

Step 1

$H_0: \sigma_f^2 \le \sigma_m^2$ vs. $H_a: \sigma_f^2 > \sigma_m^2$

$H_0: \sigma_f^2 / \sigma_m^2 \le 1$ vs. $H_a: \sigma_f^2 / \sigma_m^2 > 1$



Step 2

Step 3

Step 4

Step 5

10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Is variance of female heights greater than that of males? $\alpha = .01$ 27 values

Step 1

$H_0: \sigma_f^2 \le \sigma_m^2$ vs. $H_a: \sigma_f^2 > \sigma_m^2$

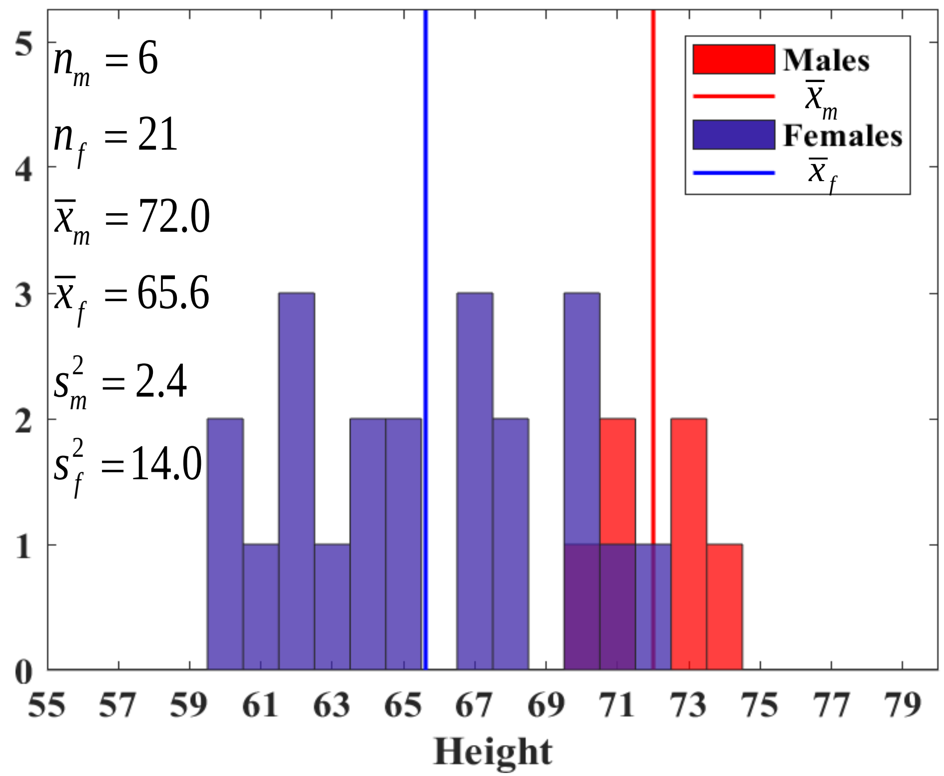
$H_0: \sigma_f^2 / \sigma_m^2 \le 1$ vs. $H_a: \sigma_f^2 / \sigma_m^2 > 1$

Step 2 $F^* = s_f^2 / s_m^2$ $\quad df_m = 5$, $df_f = 20$, $\alpha = .01$

Step 3

Step 4

Step 5



10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Is variance of female heights greater than that of males? $\alpha = .01$ 27 values

Step 1

$$H_0: \sigma_f^2 \le \sigma_m^2 \text{ vs. } H_a: \sigma_f^2 > \sigma_m^2$$

$$H_0: \sigma_f^2 / \sigma_m^2 \le 1 \text{ vs. } H_a: \sigma_f^2 / \sigma_m^2 > 1$$

Step 2

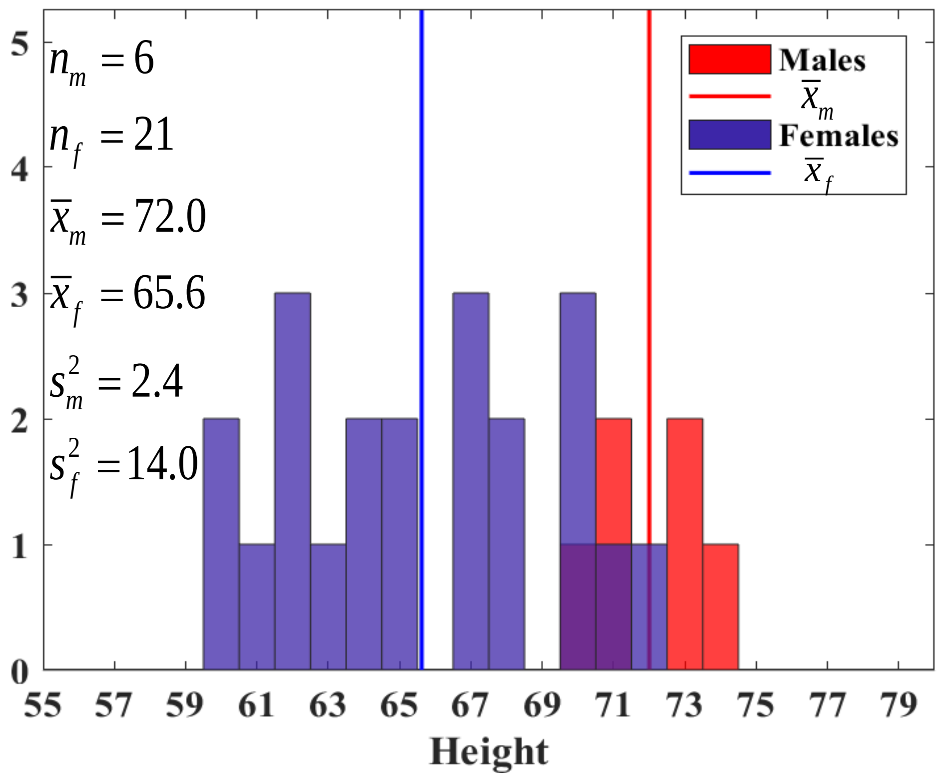
$$F^* = s_f^2 \Big/ s_m^2 \qquad df_m = 5 \quad df_f = 20 \quad \alpha = .01$$

Step 3

$$F^* = 14.0 / 2.4 = 5.83$$

Step 4

Step 5



10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Is variance of female heights greater than that of males? $\alpha = .01$ 27 values

Step 1

$H_0: \sigma_f^2 \le \sigma_m^2$ vs. $H_a: \sigma_f^2 > \sigma_m^2$

$H_0: \sigma_f^2 / \sigma_m^2 \le 1$ vs. $H_a: \sigma_f^2 / \sigma_m^2 > 1$

Step 2

$F^* = s_f^2 \big/ s_m^2$

$df_m = 5$
$df_f = 20$
$\alpha = .01$

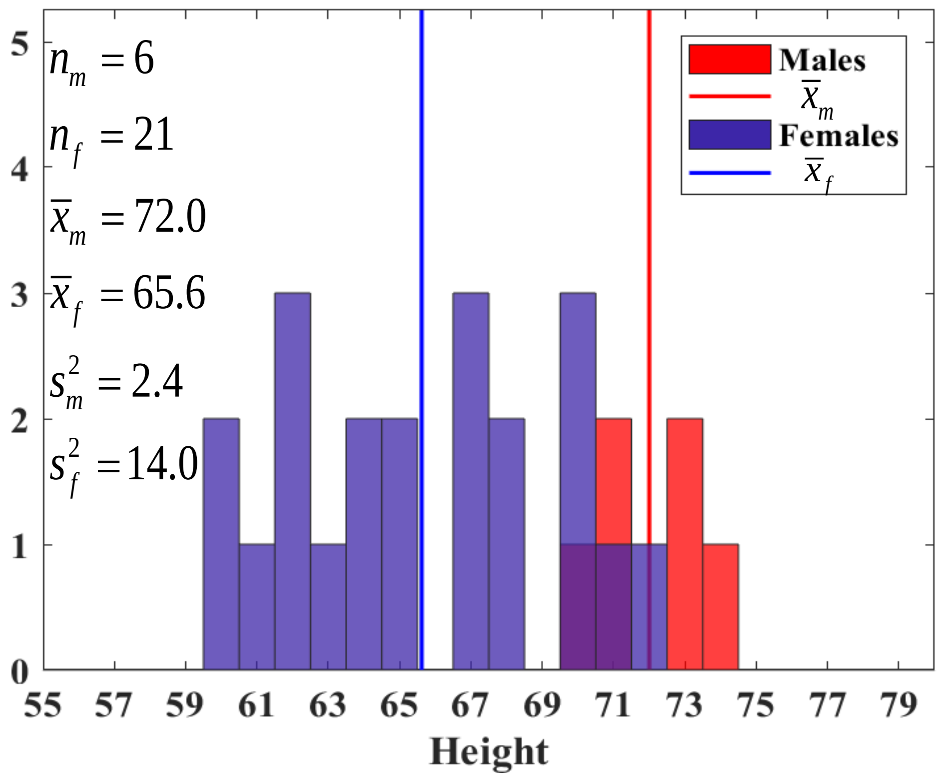
Step 3

$F^* = 14.0 / 2.4 = 5.83$

Step 4

$F(20, 5, .01) = 9.55$

Step 5



10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Is variance of female heights greater than that of males? $\alpha = .01$ 27 values

Step 1

$H_0: \sigma_f^2 \leq \sigma_m^2$ vs. $H_a: \sigma_f^2 > \sigma_m^2$

$H_0: \sigma_f^2 / \sigma_m^2 \leq 1$ vs. $H_a: \sigma_f^2 / \sigma_m^2 > 1$

Step 2

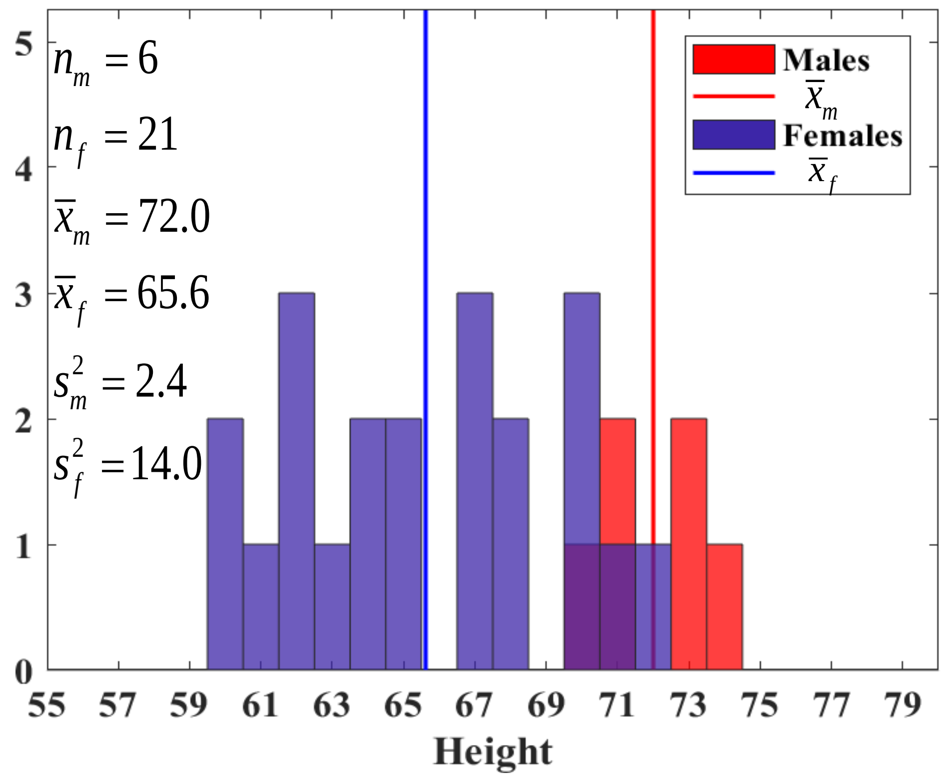
$$F^* = s_f^2 \big/ s_m^2 \qquad df_m = 5 \quad df_f = 20 \quad \alpha = .01$$

Step 3

$$F^* = 14.0 / 2.4 = 5.83$$

Step 4

$$F(20, 5, .01) = 9.55$$



Step 5 Do not reject H_0 since $5.83 < 9.55$ and conclude σ_f^2 not $> \sigma_m^2$.

Chapter 10: Inferences Involving Two Populations

Questions?

Homework: Read Chapter 10.4-10.5

WebAssign

Chapter 10# 83, 85, 91, 98, 99, 101, 111, 113, 115, 117, 119, 125, 133