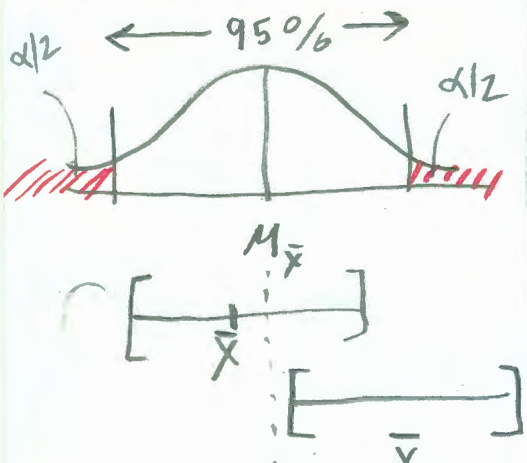
1. Inference about Difference Between 2 Pop. Means
2. Experimental Design →
3. Analysis of Variance →

Chapter 8:

Create Confidence Interval for 1 Pop. Mean



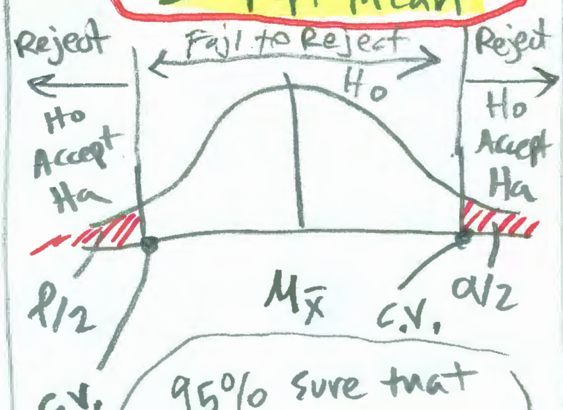
95% sure that Interval contains Pop. Mean

5% chance Interval does not contain Pop. mean.

Z or t Distribution to create Margin of Error

Chapter 9:

Hypothesis Testing for 1 Pop. Mean



95% sure that $\bar{X}$ lie between critical values

5% risk that we Reject Ho & accept Ha, even though Ho was TRUE.

Z or t test Statistic

P-value

Chapter 10:

Hypothesis Testing & confidence Intervals for Difference Between 2 Pop. means

We are interested in whether 2 Pop means are different

Example 1:

Mean Income in Bradford = \$38,010

Mean Income in Kane = \$35,006

difference = \$3,004

Is difference due to chance, or is it a real difference (statistically significant)?

Example 2:

Sample Mean Imported Compact Car MPG = $\bar{X}_1$ = 36.25
Sample Mean Domestic compact car MPG = $\bar{X}_2$ = 33.82

Difference = $\bar{X}_1 - \bar{X}_2$ = 2.69 MPG

Is the difference due to chance or is a statistically significant difference?

1. $\bar{X}_1 - \bar{X}_2$ is point estimate of:

   population Difference = $\mu_1 - \mu_2$

2. what we would like to do is take

$\bar{X}_1 - \bar{X}_2$ and compare against

Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

and then use either:

Z - Distribution
or
t - Distribution

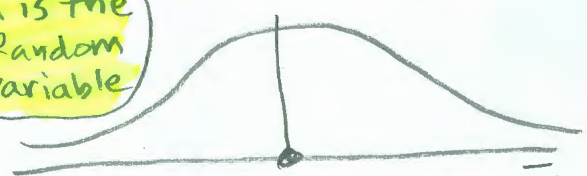
we must remember → chapter 7

chapter 7

Sampling Distribution of $\bar{X}$

Probability Distribution of all possible values of $\bar{X}$

$\bar{X}$ is the Random variable



mean $= \mu_{\bar{x}}$

SD $= \sigma_{\bar{x}}$

Mean of Sampling Distribution of $\bar{X}$

$$\mu = \mu_{\bar{x}} = E(\bar{X}) = \frac{\text{sum of all possible } \bar{X}}{\text{count of all possible } \bar{X}}$$

If we are able to select all possible samples of size **n** from a given population, then the mean of the sampling Distribution of $\bar{X}$ is equal to the pop. mean μ, that is: $\mu = \mu_{\bar{x}}$

and

Standard Deviation of Sampling Distribution of $\bar{X}$, standard Error:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$$

Finite Pop. correction Factor Not Need when $n/N \leq 0.05$

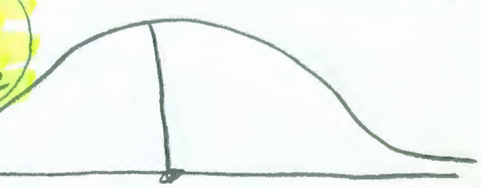
chapter 10

Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

(P.3)

Probability Distribution of all possible values of $\bar{X}_1 - \bar{X}_2$

$\bar{X}_1 - \bar{X}_2$ is the Random variable



mean $= \mu_{\bar{x}_1} - \mu_{\bar{x}_2}$

SD $= \sigma_{\bar{x}_1 - \bar{x}_2}$

mean of Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

$$\mu_1 - \mu_2 = \mu_{\bar{x}_1 - \bar{x}_2} = E(\bar{X}_1 - \bar{X}_2) = \frac{\text{Sum of all possible } \bar{X}_1 - \bar{X}_2}{\text{Count of all possible } \bar{X}_1 - \bar{X}_2}$$

If we are able to select all possible sample differences of size **n** from a given population, then the mean of the sampling Distribution of $\bar{X}_1 - \bar{X}_2$ is equal to the Pop. mean $\mu_1 - \mu_2$, that is: $\mu_1 - \mu_2 = \mu_{\bar{x}_1 - \bar{x}_2}$.

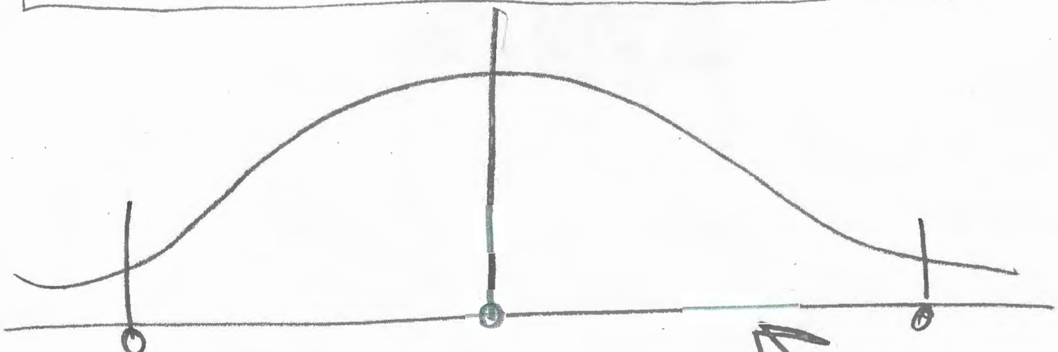
and

Standard Deviation of Sampling Distribution of $\bar{x}_1 - \bar{x}_2$, standard Error:

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} * \sqrt{\frac{N-n}{N-1}}$$

Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

(SD of $\bar{X}_1 - \bar{X}_2$)



$\bar{X}_1 - \bar{X}_2$

$\mu_{\bar{x}_1 - \bar{x}_2}$

$\sigma_{\bar{x}_1 - \bar{x}_2}$

Here?

Here?

We can take the difference between 2 samples and compare to SD of $\bar{X}_1 - \bar{X}_2$

$\bar{X}_1 - \bar{X}_2$

Let's go over to Excel →

3 Conclusions from Excel Example of Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

1

$$\left\{\begin{matrix}\text{Difference}\\ \text{Between}\\ \text{Pop. Means}\end{matrix}\right\} = \mu_1 - \mu_2 = \mu_{\bar{x}_1 - \bar{x}_2} = E(\bar{X}_1 - \bar{X}_2) = \left\{\begin{matrix}\text{Mean of}\\ \text{ALL}\\ \text{sample}\\ \text{Differences}\end{matrix}\right\}$$

2 Sampling Distribution of $\bar{X}_1 - \bar{X}_2$ tends to be symmetrical.

3 Less variation in Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

4

$$\left\{\begin{matrix}\text{Standard}\\ \text{Deviation}\\ \text{of}\\ \text{SD of } \bar{X}_1 - \bar{X}_2\end{matrix}\right\} = \left\{\begin{matrix}\text{Standard}\\ \text{Error}\\ \text{of}\\ \text{SD of } \bar{X}_1 - \bar{X}_2\end{matrix}\right\} = \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

* Finite Population Correction Factor not used in any of textbook problems.

4 Central Limit Theorem

If both populations have a normal distribution, or if the sample size, n, is big enough ($n \geq 30$, skewed $n > 50$) then we can use: ① SD of $\bar{X}_1 - \bar{X}_2$ ② $SE = \sigma_{\bar{x}_1 - \bar{x}_2}$

→ Build Confidence Intervals

→ Do Hypothesis Testing

4

Confidence Interval for $\mu_1 - \mu_2$ when σ_1 and σ_2 are Known

$$\left\{\begin{matrix}\text{point Estimate} \\ \text{of} \\ \mu_1 - \mu_2\end{matrix}\right\} = \bar{X}_1 - \bar{X}_2$$

$$\left\{\begin{matrix}\text{Standard Deviation/} \\ \text{Standard Error of} \\ \text{Sampling Distribution of} \\ \bar{X}_1 - \bar{X}_2\end{matrix}\right\} = \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_2} + \frac{\sigma_2^2}{n_2}}$$

* correct factor Not used in this chapter

$$\left\{\text{Margin of Error}\right\} = \underbrace{Z_{upper}}_{\text{\# of standard Deviations}} * \underbrace{\sigma_{\bar{x}_1 - \bar{x}_2}}_{\text{standard Deviation/Error}}$$

$$\left\{\text{Confidence Interval}\right\} = \bar{X}_1 - \bar{X}_2 \pm Z_{upper} * \sigma_{\bar{x}_1 - \bar{x}_2}$$

Variables:

$\bar{X}_1$ = sample mean from population 1
n_1 = sample size from pop. 1
σ_1 = sigma = standard Deviation from pop. 1
σ_1^2 = variation from pop 1.

$\bar{X}_2$ = sample mean from pop. 2
n_2 = sample size from pop 2
σ_2 = sigma = SD from pop 2
σ_2^2 = variation from pop 2
$\mu_1 - \mu_2$ = pop. Difference between 2 Pop. Means

2 Independent Random Samples (taken separately & independently so no bias.)

Confidence Interval Example σ Know:

- Two cities, Bradford and Kane are separated only by the Conewango River
- The local paper recently reported that the mean household income in Bradford is $38,010 from a sample of 40 households. The population standard deviation (past data) is $6,000.
- The same article reported the mean income in Kane is $35,006 from a sample of 35 households. The population standard deviation (past data) is $7,000.
- From the sample data, create a 95% Confidence Interval to estimate population difference.

Bradford Data = 1	Kane Data = 2
$\bar{X}_1 = \$38{,}010$	$\bar{X}_2 = \$35{,}006$
$n_1 = 40$	$n_2 = 35$
$\sigma_1 = \$6{,}000$	$\sigma_2 = \$7{,}000$
$\sigma_1^2 = 6000^2 = 36{,}000{,}000$	$\sigma_2^2 = 7000^2 = 49{,}000{,}000$

CI = 95%

$\alpha = 5\%$

$\frac{\alpha}{2} = \frac{5\%}{2} = 2.5\%$

$\left\{ \text{Point Estimate of } \mu_1 - \mu_2 \right\} = \bar{X}_1 - \bar{X}_2 = \$38{,}010 - \$35{,}006 = \$3{,}004$

$\left\{ \text{SE of / SD of } \bar{X}_1 - \bar{X}_2 \right\} = \sqrt{\frac{36000000}{40} + \frac{49000000}{35}} = 1516.575089 = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Z_{upper} = NORM.S.INV (1 − 0.025) = 1.959963985

Margin of Error = 1.959963985 * 1516.575089 = 2972.43

$\left\{ \text{Confidence Interval} \right\} = \underbrace{\bar{X}_1 - \bar{X}_2}_{\text{Point Estimate}} \pm \underbrace{Z_{upper} * SE}_{\text{Margin of Error}} = 3004 \pm 2972.43$

Lower = \$31.57

Upper = \$5976.43

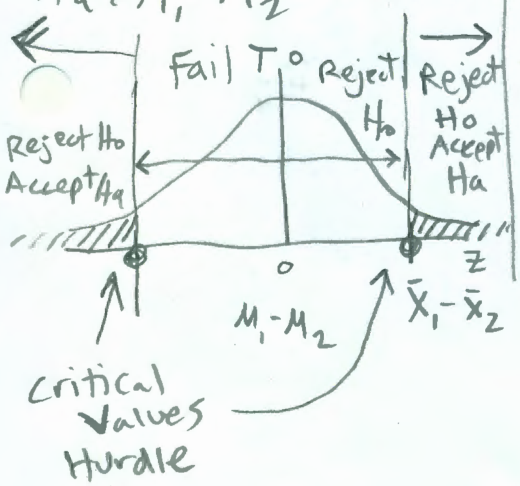
We are 95% sure that population difference between the Bradford mean Income & Kane mean Income is between about \$32 and \$5976.

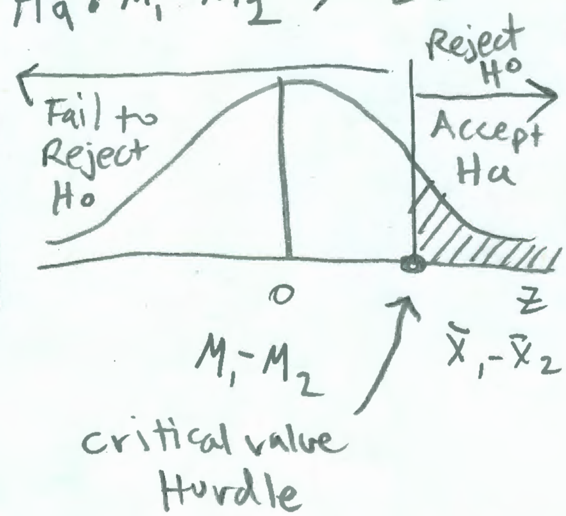
5

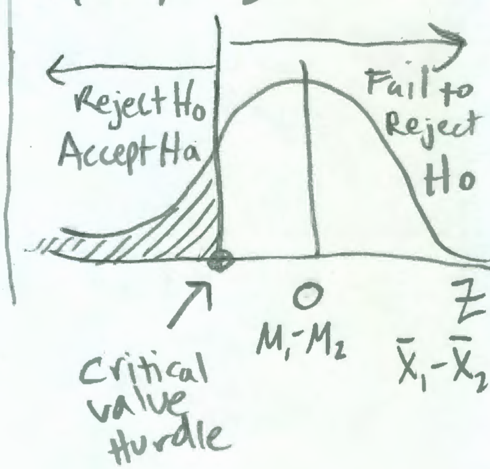
Hypothesis Testing about $\mu_1 - \mu_2$, σ_1 & σ_2 KNOWN

- Same 5 steps as Chapter 9 (P.9)
- $\mu_1 - \mu_2 = D_0$ = Hypothesized Difference
- Possible H_0 & H_a:

2-tail	1 tail upper	1 tail lower
$H_0: \mu_1-\mu_2 = D_0$	$H_0: \mu_1-\mu_2 <= D_0$	$H_0: \mu_1-\mu_2 >= D_0$
$H_a: \mu_1-\mu_2 <> D_0$	$H_a: \mu_1-\mu_2 > D_0$	$H_a: \mu_1-\mu_2 < D_0$







Test statistic sigma KNOWN for Hypothesis Testing about $\mu_1 - \mu_2$

$$= Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Sampling Error

SE of Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

Remember σ_1 & σ_2 are SD
σ_1^2 & σ_2^2 are variance.

Hypothesis Testing continued:

- $\mu_1 - \mu_2$ = difference between 2 pop means.
- σ_1 = pop SD$_1$ σ_2 = Pop SD$_2$

  σ_1^2 = Pop variance$_1$ σ_2^2 = Pop variance$_2$
- n = sample size
- If populations are both normal or n is big enough to envoke Central Limit Theorem, we can use Standard Normal curves.

- To get statistics for Hypothesis Testing we can use:

  Excel → Data Ribbon Tab, Data Analysis,

  Descriptive statistics:

  "Z - Test: Two Samples for Means"

- 2 samples must independent random samples (so no bias)

Example of Hypothesis Testing σ Known: (P.11)

- Two cities, Bradford and Kane are separated only by the Conewango River
- The local paper recently reported that the mean household income in Bradford is $38,010 from a sample of 40 households. The population standard deviation (past data) is $6,000.
- The same article reported the mean income in Kane is $35,006 from a sample of 35 households. The population standard deviation (past data) is $7,000.
- At the .05 significance level can we conclude the mean income in Bradford is more?

More >

Can we conclude the mean income in Bradford (1) is more

Bradford Data = 1	Kate Data = 2
$\bar{X}_1 = \$38,010$	$\bar{X}_2 = \$35,006$
$n_1 = 40$	$n_2 = 35$
$\sigma_1 = \$6000$	$\sigma_2 = \$7000$
$\sigma_1^2 = 6000^2 = 36,000,000$	$\sigma_2^2 = 7000^2 = 49,000,000$

$\bar{X}_1 - \bar{X}_2 = 38,010 - 35,006 = 3,004$

Step 1: List Null and Alternative Hypotheses

$H_0: \mu_1 - \mu_2 \le 0$

$H_a: \mu_1 - \mu_2 > 0$

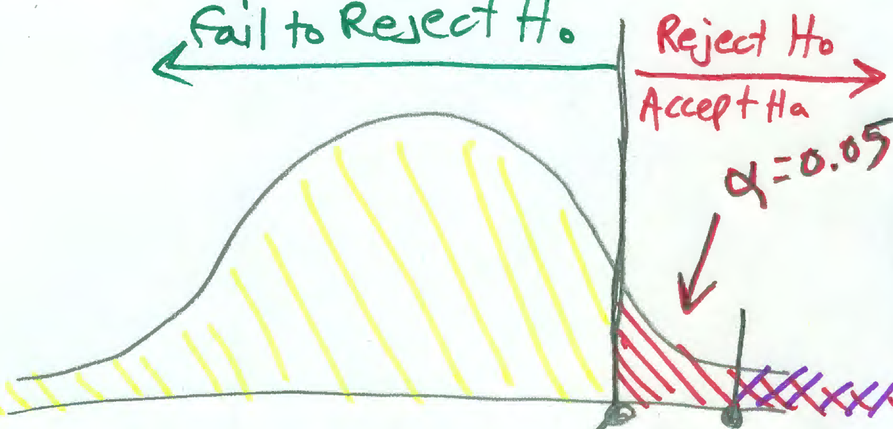
Step 2: Level of Significance = Alpha = Risk of rejecting H0 when it is TRUE

$\alpha = 0.05$

upper test

Step 3: Sample, Calculate, Draw Picture, Calculate Test Statistic

(1) use Z because sigma KNOWN



$$z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$Z = \frac{3004 - 0}{\sqrt{\frac{36M.}{40} + \frac{49M.}{35}}}$$

critical value = 1.6449

p-value = $P(Z \geq 1.98) =$ 0.0238

$Z = 1.9808$ = test statistic

Step 4: If test statistic > 1.649, Reject Ho & Accept Ha, otherwise
If p-value ≤ alpha (0.05), Reject Ho & accept Ha, otherwise... fail to Reject Ho

Step 4: Create Rules for Critical Value and p-value

Critival value upper = 1.6449 = NORM.S.INV(1-.05)

P-value = 1-NORM.S.DIST(1.9808) = 0.0238

Step 5: Conclude

Because our test statistic > critical value & p-value < alpha, We reject Ho & accept Ha. The evidence suggests that the mean income in Bradford is more than in Kane.

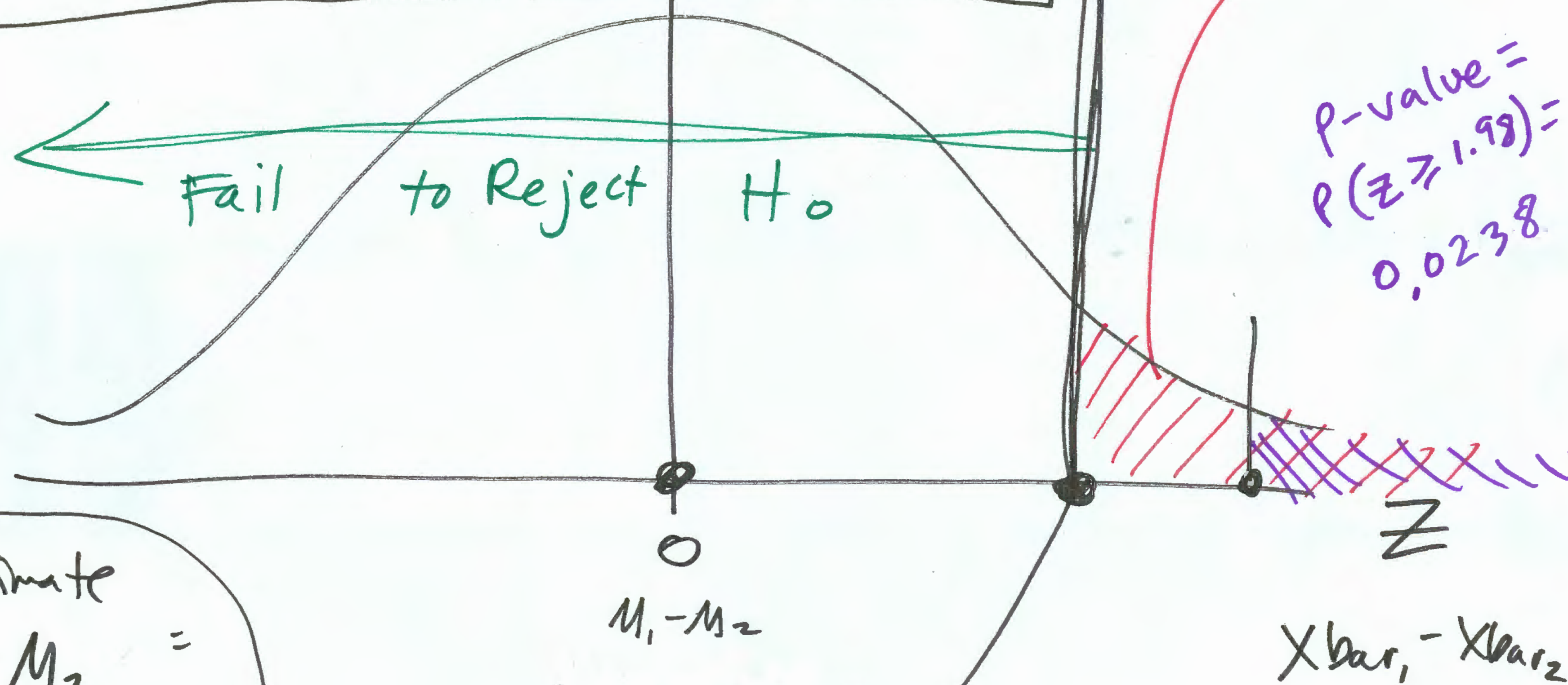
Bradford Mean Annual Income = 1

Kane Mean Annual Income = 2

$H_0: \mu_1 - \mu_2 \le 0$

$H_a: \mu_1 - \mu_2 > 0$

Sampling Distribution of Xbar₁ - Xbar₂



$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Critical value = 1.645

Point Estimate of $\mu_1 - \mu_2$ =

$\bar{X}_1 - \bar{X}$ =

\$38,010 - \$35,006

= \$3,004

6 Confidence Interval For $\mu_1 - \mu_2$, σ_1 & σ_2 UNKNOWN

P.13

- $\mu_1 - \mu_2$ = difference between 2 pop. means
- σ → sigma, pop standard deviation, is not known, so we use sample standard deviation, s, in place of sigma, σ, & use t distribution rather than Z distribution.
- If populations are both normal or sample size, n, is big enough to envoke Central Limit Theorem, we can use t Distributions.

$\bar{x}$ = Xbar = sample mean

s = sample standard Deviation

s^2 = sample variance

n = sample size

α = alpha = risk that pop mean is Not in Interval.

$1-\alpha$ = Confidence Limit/Coefficient

Must use 2 Independent Random Samples

Upper & Lower Limit for Confidence Interval

$$\underbrace{\bar{X}_1 - \bar{X}_2}_{\text{point Estimate for } \mu_1 - \mu_2} \pm \overbrace{t_{upper} * \underbrace{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}_{SE}}^{\text{margin of Error}}$$

t_{upper} Estimates # of standard Deviations

SE = Estimates Standard Deviation of sampling Distribution of Xbar1 - Xbar2 for t Distribution

7 Degrees of Freedom, df, for t Distribution with 2 Independent Random samples:

P.14

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

s_1^2 = variance from sample 1 Data

n_1 = sample size taken from pop. 1

s_2^2 = variance from sample 2 Data

n_2 = sample size taken from pop. 2.

Note 1: Best to round down df to get more conservative Interval. * Excel T functions will automatically truncate df (Rounddown).

Note 2: Excel Hint: Because $\frac{s^2}{n}$ is used so often in formula, use helper cell to calculate $\frac{s^2}{n}$ & then refer to it in formula with cell reference.

Note 3: Does not require that $\sigma_1 = \sigma_2$ (like some other df formulas do). Formula works whether $\sigma_1 = \sigma_2$ or not!!

Formula Not used in this class:

If you can assume $\sigma_1 = \sigma_2$ then:

{pooled variance} = {weighted mean of 2 variances} = $s_p^2 = \frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{n_1 + n_2 - 2}$

t test statistic $= \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2 * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

Confidence Interval Example for σ NOT KNOWN:

- A recent EPA study compared the highway fuel economy of domestic and imported compact cars
- A sample of 20 domestic compact cars revealed a mean of 33.83 MPG with a standard deviation of 2.33 MPG
- A sample of 20 imported compact cars revealed a mean of 36.52 MPG with a standard deviation of 3.43 MPG
- Samples are Independent, Distributions for samples are normal
- At the 95% Confidence Level, construct a Confidence Interval for the Population Difference between Imported Auto MPG and Domestic Auto MPG

P. 15

Imported compact Auto MPG = Data set 1	Domestic Compact Auto MPG = Data set 2
$n_1 = 20$	$n_2 = 20$
$n_1 - 1 = 19$	$n_2 - 1 = 19$
$\bar{X}_1 = 36.5185$ MPG	$\bar{X}_2 = 33.825$ MPG
$s_1 = 3.4329$ MPG	$s_2 = 2.3337$
$s_1^2 = Var_1 = 11.7852$	$s_2^2 = Var_2 = 5.4462$
$s_1^2/n_1 = 11.7852/20 = 0.5893$	$s_2^2/n_2 = 5.4462/20 = 0.2723$

* calcs done in Excel

$$\{\text{Point Estimate of } \mu_1 - \mu_2\} = \bar{X}_1 - \bar{X}_2 = 36.5185 - 33.825 = 2.6935 \text{ MPG}$$

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{0.5893 + 0.2723} = 0.9282 \text{ MPG}$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{[0.5893 + 0.2723]^2}{\frac{(0.5893)^2}{19} + \frac{(0.2723)^2}{19}} = 33.4703 \Rightarrow 33 \text{ df}$$

upper t = T.INV(1−0.95, 33) = 2.0345 = # Standard Deviations

Margin of Error = $t * SE$ = 2.0345 * 0.9282 = 1.8884 MPG

Confidence Interval = $\bar{x}_1 - \bar{x}_2$ +/- upper t * SE = 2.6935 +/- 1.8884 MPG

We are 95% sure that Pop. diff. between MPG for Imported & Domestic Compact Autos is between 0.81 & 4.58 MPG.

8 Hypothesis Testing About $\mu_1 - \mu_2$, σ_1 & σ_2 UNKNOWN

- Same 5 tests for Hypothesis Testing as ch. 9
- $\mu_1 - \mu_2 = D_0$ = Hypothesized Difference
- 2 Independent Random samples.

Test statistic sigma UNKNOWN for Hypothesis Testing about $\mu_1 - \mu_2$

$$t = \frac{\overbrace{(\bar{X}_1 - \bar{X}_2) - D_0}^{\text{sampling Error}}}{\underbrace{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}_{SE}}$$

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

used in Excel T functions

* BIG Excel Hint 2: Data Ribbon, Data Analysis, = t-Test: Two sample Assume unequal variance

Excel

* Big Excel Hint 1: Calculate $\frac{s^2}{n}$ is separate cell & refer to it in formulas with cell References.

General Notes about Hypothesis Testing:

★ Build Alternative Hypothesis, H_a, so that you can make the desired conclusion (Because we control for Type I Error).

For example:

If you want to test whether $\mu_1 > \mu_2$

then, when we accept H_a:

$$H_a: \mu_1 - \mu_2 > 0$$

↓

becomes

$$\mu_1 > \mu_2$$

Notes for sample size for t test:

★ whenever possible, equal sample sizes of $n_1 = n_2$ are recommended. page 448 in text.

★ even small sample sizes of $n_1 + n_2 = 20$, tend to give good results. Page 448 in text

Example Hypothesis Testing for $\mu_1 - \mu_2$, σ NOT KNOWN

- A recent EPA study compared the highway fuel economy of domestic and imported compact cars
- A sample of 20 domestic compact cars revealed a mean of 33.83 MPG with a standard deviation of 2.33 MPG
- A sample of 20 imported compact cars revealed a mean of 36.52 MPG with a standard deviation of 3.43 MPG
- Assume: 1) Samples are random and independent, 2) Distributions for samples are normal.
- At the .05 significance level can the EPA conclude that the MPG is higher on the imported cars?

1 = Domestic Compact Auto MPG	2 = Import Compact Auto MPG
$n_1 = 20$ $\quad n_1 - 1 = 19$	$n_2 = 20$ $\quad n_2 - 1 = 19$
$\bar{x}_1 = 33.83$ MPG	$\bar{x}_2 = 36.52$ MPG
$s_1 = 2.33$ MPG	$s_2 = 3.43$ MPG
$s_1^2 = 2.33^2 = 5.45$	$s_2^2 = 3.43^2 = 11.79$
$s_1^2/n_1 = 5.45/20 = 0.2723$	$s_2^2/n_2 = 11.79/20 = 0.5893$

Step 1: List Null and Alternative Hypotheses

$$H_0: \mu_1 - \mu_2 >= 0$$

$$H_a: \mu_1 - \mu_2 < 0$$

Flip

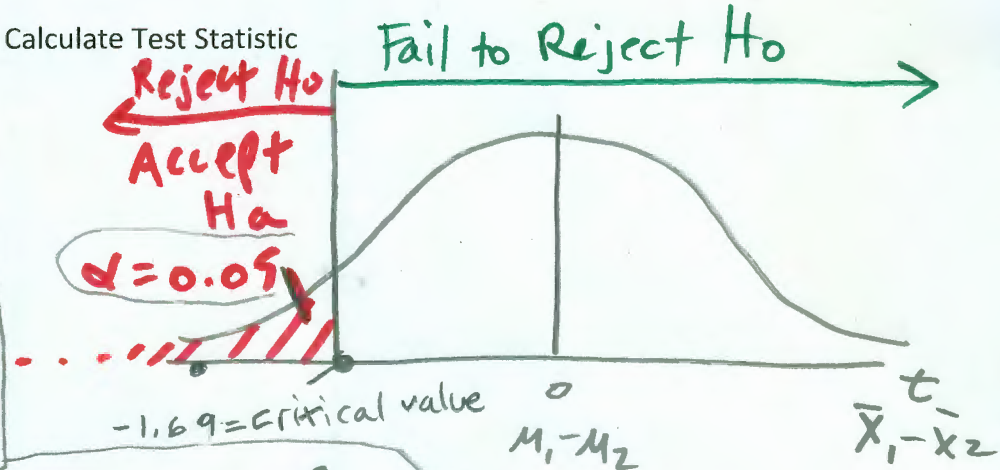
Because "2" should be "higher" $\mu_1 - \mu_2$ would be negative.

Step 2: Pick Level of Significance = Alpha (Risk that H0 is TRUE, but we Reject it)

$$\alpha = 0.05$$

Step 3: Sample, Calculate, Draw Picture, Calculate Test Statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{33.83 - 36.52}{\sqrt{0.2723 + 0.5893}} = -2.9$$



$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{(0.2723 + 0.5893)^2}{\frac{(0.2723)^2}{19} + \frac{(0.5893)^2}{19}} = 33$$

critical value →

p-value →

Step 4: Create Rules for Critical Value and p-value

$$\text{critical value} = T.INV(0.05, 33) = -1.69$$

$$\text{p-value} = T.DIST(-2.9, 33, 1) = 0.0032$$

Critical value Rule:

If our calculated test statistic is greater than our critical value, we Reject Ho & accept Ha, otherwise we fail to Reject Ho

p-value Rule:

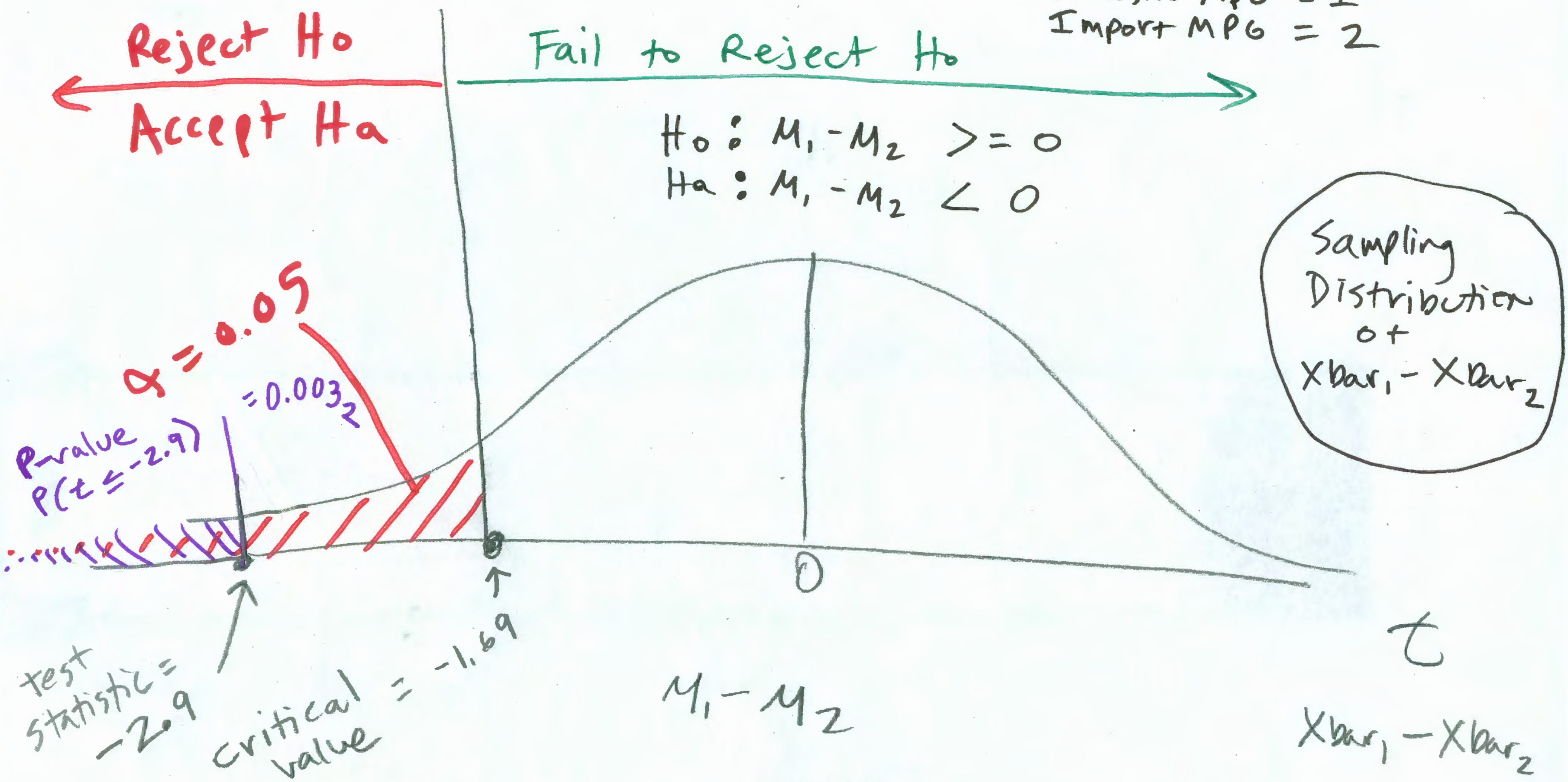
If the p-value is less than our alpha, we Reject Ho & accept Ha, otherwise we fail to Reject Ho.

Step 5: Conclusion

Because our calculated test statistic of -2.9 is less than our critical value of -1.69, we reject Ho & accept Ha. The sample evidence strongly suggests that the MPG is higher for Imported compact cars than it is for Domestic compact cars.

Domestic MPG = 1
Import MPG = 2

$H_0: \mu_1 - \mu_2 >= 0$
$H_a: \mu_1 - \mu_2 < 0$



$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

9 Inference (Hypothesis Testing or Confidence Interval) about

$\mu_1 - \mu_2$ Matched Samples

Matched Samples:

1. Each sampled element provides a pair of Data values.

Examples:

1. 1 worker tries 2 different production methods & provides 2 different times to complete production
2. 1 person gives a rating for a product (0 to 10) before watching an ad and after watching an ad.
3. 2 different (but similar) workers provide times for production method 1 & production method 2.

Because they are same person for each data value, variability's & sampling Error between two values is reduced.

2. The pair of matched sample values are then subtracted to get the difference.

3 Compare & contrast Matched & Independent Samples:

Independent Sample Design:

Samples selected from two populations in such a way that the elements making up one sample are chosen independently of the elements making up the other sample.

Matched Sample Design:

One simple random sample of elements is selected & 2 data values are obtained for each element (or similar elements).

★ Matched Sample Design often leads to smaller Sampling Error than the Independent Design because variation between sampled items (like difference in production workers) is significantly reduced as a source of sample error.

Experimental Design

4. Key to analysis of matched sample design: we consider only column of differences.

5. Example of Matched Samples: (p.22)

① Established method for making Quad Boomerangs  = Sample 1 = Method 1

New method for making Quad Boomerangs  = Sample 2 = Method 2

② simple Random Sample to pick 10 employess to try both methods

③ We time both methods for each worker:

worker	method 1 time (Minutes)	Method 2 Time (Minutes)	difference (d_i)
1	12	11	12−11 = −1.0
2	11	11.3	11−11.3 = −0.3
3	13	10.5	13−10.5 = 2.5
4	12.2	12	12.2−12 = 0.2
5	12	12	12−12 = 0.0
6	13	13.1	13−13.1 = −0.1
7	11.2	10.4	11.2−10.4 = 0.8
8	10.9	10.8	10.9−10.8 = 0.1
9	12.5	12	12.5−12 = 0.5
10	12.4	11.8	12.4−11.8 = 0.6
			Total =

(6)

Formulas

(4) {Confidence Interval} $\bar{d} \pm \underbrace{\underset{\#\text{S.D.}}{t_{\alpha/2}} * \underset{\text{SE}}{\frac{s_d}{\sqrt{n}}}}_{\text{margin of Error}}$

(P.23)

(1) {mean of matched sample differences} $= \bar{d} = \sum \frac{d_i}{n}$

Excel: AVERAGE

(2) {standard Deviation of Matched sample differences} $= s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}}$

Excel: STDEV.S

(3) {test statistic t} $= \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$

Same formulas as we did in chapter 3 for single variables or samples.

$\bar{d}$ = mean of Matched Sample Diff.

d_i = particular difference

n = sample size

Excel

(5) Excel Data Ribbon Tab, Data Analysis, "t-test: Paired 2 sample for means"

s_d = Standard Deviation of Matched Sample Differences

μ_d = Pop mean of differences

t = test statistic

$t_{\alpha/2}$ = upper t for Confidence Interval

(6) Same Rules for Confidence Interval & Hypothesis Testing as earlier in class.

7 Hypothesis Test Matched sample Example:

Question: Is there a difference between 2 methods?

Step 1

$H_0 : \mu_d = 0$

$H_a : \mu_d <> 0$

Step 2

Alpha = 0.05

Alpha/2 = 0.025

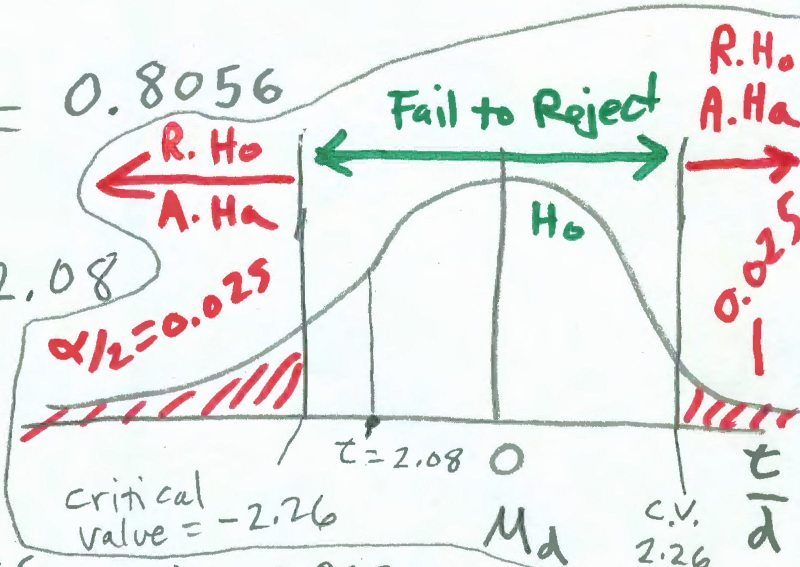
Step 3

$\bar{d} = \Sigma \frac{d_i}{n} = 0.53$

$S_d = \sqrt{\frac{\Sigma (d_i - \bar{d})^2}{n-1}} = 0.8056$

$t = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}} = 2.08$

df = 10 − 1 = 9



Step 4

P-value = T.DIST.2T(2.08, 9) = 0.0672

upper t = T.INV(1 − 0.025, 9) = 2.26

Rules: If our calculated test statistic is greater than 2.26 or less than −2.26, we reject Ho & accept Ha, otherwise we fail to Reject Ho.

If our p-value is smaller than our alpha, we Reject Ho and Accept Ha, otherwise we fail to Reject Ho.

Step 5

Because our calculated Test statistic is between −2.26 & 2.26 AND because our p-value is bigger than our alpha, we Fail to Reject Ho. → more steps continued:

The sample evidence does not support the conclussion that the new method is faster (different from) than the old method. we do run a 5% risk of Failing to Reject Ho, even though Ha was TRUE (New method was different from old method).

8 Confidence Interval for Matched Samples:

$$\begin{Bmatrix}\text{confidence}\\ \text{Interval}\end{Bmatrix} = \bar{d} +/- \underbrace{t_{\alpha/2} * \frac{S_d}{\sqrt{n}}}_{\text{Margin of Error}}$$

$t_{\alpha/2}$ ↑ # standard Deviations

$\frac{S_d}{\sqrt{n}}$ ← Standard Error

$$= 0.53 +/- 2.26 * \frac{0.8056}{\sqrt{10}}$$

$$0.53 +/- 0.5763 \Rightarrow -0.0463 \;\&\; 1.1063$$

We are 95% sure that the population difference lies between -0.0463 & 1.1063 minutes difference.

Final Note about Matched Samples:

when comparing 2 population means, the matched sample procedure generally provides better precision than the independent sample approach, and therefore it is the recommended design when possible.

Experimental Design

BUT what do we do if we are comparing more than 2 population means?

Completely Randomized Design $_{CRD}$ & Analysis of Variance ANOVA

*Before we define these ↑, Let's look at an example.

1 Example of CRD & ANOVA

1) Boomerang manufacturer is considering 4 different methods for manufacturing Quads
2) A random sample of 20 workers is selected.

3) The workers are randomly assigned 1 of 4 manufacturing methods
4) Each worker is timed making the Quad with the assigned manufacturing method.
5) Time is recorded in minutes rounded to 2 decimals
6) The mean time for each method is recorded
7) We want to determine if there is a significant difference between the 4 mean times — or "Are all means equal?"
8) collected data (times & means):

	method 1 (mins)	method 2 (mins)	method 3 (mins)	method 4 (mins)
	12.56	11.87	15.57	12.94
	15.92	12.94	10.64	12.01
	10.42	11.23	14.05	10.21
	14.41	15.97	13.08	10.18
	10.32	17.14	16.71	11.21
Xbar	12.73	13.83	14.01	11.31
S^2	6.04	6.73	5.49	1.41

sample mean

sample variation

From this sample data we want to ask the question:

$$\mu_1 = \mu_2 = \mu_3 = \mu_4$$

We have:

1 = Method 1 mean time = 12.73 mins
2 = Method 2 mean time = 13.83 mins
3 = Method 3 mean time = 14.01 mins
4 = Method 4 mean time = 11.31 mins

If we were to compare each mean, one at a time, we would have to run six tests:

(1 vs. 2), (1 vs. 3), (1 vs. 4), (2 vs. 3), (2 vs. 4), (3 vs. 4)

(*) If we ran six tests we would build up a lot of Type 1 Error (Risk of Rejecting H_0, when it is TRUE).

<u>Example:</u> @ alpha = 0.05

P (All correct) = 0.95 * 0.95 * 0.95 * 0.95 * 0.95 * 0.95 =
= 0.735

P(at least 1 NOT correct) = 1 − 0.735 = 0.265

(*) we would also have to do many calculations if we tested all six sets of means.

So, we have to learn a different statistical method to test:

$$\mu_1 = \mu_2 = \mu_3 = \mu_4$$

2 Analysis of Variance ANOVA

- A statistical method to check whether K population means are equal $\mu_1 = \mu_2 = \cdots = \mu_k$

- Define Terms used in ANOVA:

Factor = independent variable of interest = Manufacturing Method for Quad Boom.

Treatment = different Populations
= Different Levels of Factor } Synonyms

4 pops.
4 Treatments

	Treatment 1	Treatment 2	Treatment 3	Treatment 4
	Method 1 (mins)	Method 2 (mins)	Method 3 (mins)	Method 4 (mins)
	12.56	11.87	15.57	12.94
	15.92	12.94	10.64	12.01
	10.42	11.23	14.05	10.21
	14.41	15.97	13.08	10.18
	10.32	17.14	16.71	11.21
			14.01	11.31
sample mean xbar	12.73	13.83	5.49	1.41
sample variance s^2	6.04	6.73		

Time to Make Quad in Mins

Response variable = Dependent variable

Single Factor Experiment = 1 Factor

Experimental unit = object of interest in an experiment → our example = people making Quads.

But, we must be careful to make sure we have a random experiment so we will use:

3) Completely Randomized Design (CRD)

- An experimental Design in which the treatments are randomly assigned to experimental units
- For us CRD means randomly assign manufacturing methods to Workers.

Employees Making Boomerangs:
Annette Caldwell
Sophia Wise
Florence Ferguson
Jorge Reid
Belinda Ruiz
Everett Marsh
Guy Frazier
Angel Clayton
Johnnie Zimmerman
Jenna Rios
Alvin Mccoy
Carole Sullivan
Winifred Ortega
Lucas Ortiz
Sharon Richardson
Marsha Gutierrez
Jodi Houston
Ora Kim
Loren Freeman
Danny Mullins
Jared Young
Sylvester Hodges
Daniel Watson
Nathan Beck
Lamar Hammond
Tommie Turner
Garrett King
Bradford Boone
Julian Quinn
Phyllis Waters
Carlton Matthews
Kimberly Palmer
Arlene Baker
Jose Goodman
Floyd Erickson
Faith Salazar
Kathy Dean
Ernestine Weaver
Judy Mckenzie

From List Of 100, randomly select 20 Workers ==>

60 more in list below

v v v v

Workers
Sophia Wise
Alexandra Moran
Herbert Garrett
Sharon Richardson
Robert Valdez
Winifred Ortega
Carole Sullivan
Willard Payne
Irvin Owens
Constance Henderson
Jodi Houston
Danny Mullins
Jaime Gibson
Arlene Baker
Guy Frazier
Van Simmons
Floyd Erickson
Denise Farmer
Reginald Montgomery
Edwin Watkins

Randomly Assign Manufacturing Method ==>

Workers	Method
Sophia Wise	Method 4
Alexandra Moran	Method 4
Herbert Garrett	Method 3
Sharon Richardson	Method 4
Robert Valdez	Method 2
Winifred Ortega	Method 3
Carole Sullivan	Method 3
Willard Payne	Method 1
Irvin Owens	Method 2
Constance Henderson	Method 2
Jodi Houston	Method 1
Danny Mullins	Method 4
Jaime Gibson	Method 2
Arlene Baker	Method 3
Guy Frazier	Method 1
Van Simmons	Method 1
Floyd Erickson	Method 1
Denise Farmer	Method 4
Reginald Montgomery	Method 3
Edwin Watkins	Method 2

Time and record data ==>>

Pop 1	Pop 2	Pop 3	Pop 4
Treatment 1	Treatment 2	Treatment 3	Treatment 4
Method 1	Method 2	Method 3	Method 4
12.56	11.87	15.57	12.94
15.92	12.94	10.64	12.01
10.42	11.23	14.05	10.21
14.41	15.97	13.08	10.18
10.32	17.14	16.71	11.21

experimental units = workers who make Quads

treatments = populations = methods to make Quad 1, 2, 3, 4.

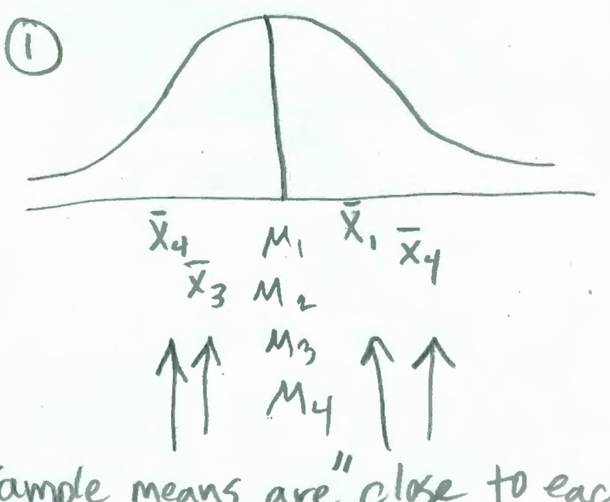
Before we use ANOVA to test, we must check what assumptions are neccessary:

4) Assumptions for ANOVA (Analysis of Variance):

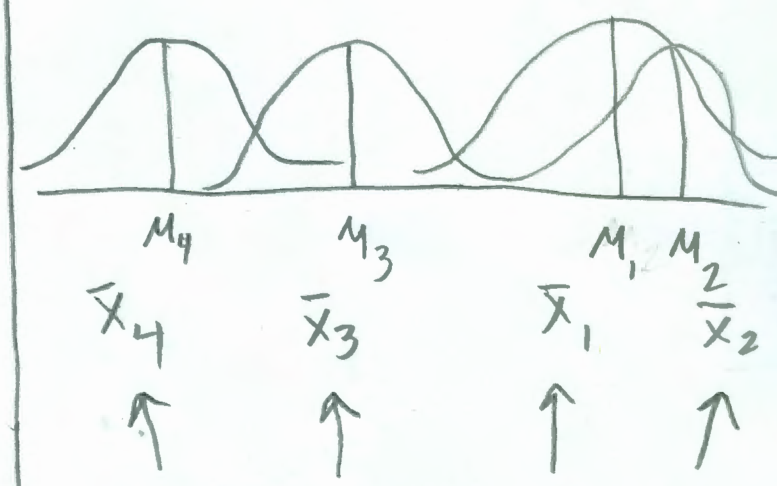
1. The population for the response variables (time to make Quad) follow Normal Distribution.
2. The population variance of response variables (σ^2) are the same for all populations.
3. observations must be independent (Not related or linked or influenced by other observations).

5) Conceptual overview of ANOVA:

1.



Sample means are "close to each other" because there is only one sampling Distribution when H_0 is TRUE



Sampling means come from different sampling Distributions and are not close together when H_0 is FALSE

5 conceptual overview of ANOVA continued:

2 with ANOVA we will develop 2 independent estimates of the common population variance & then compare them with division P.32

1 MSTR = estimate of variability **Between** the samples

MSE = estimate of variability **Within** Each sample

2 $\frac{MSTR}{MSE}$ too big we don't think means are equal.

$\frac{MSTR}{MSE}$ close to 1 then reasonable to assume means equal. (F critical value will tell us how close to 1 we need to be.)

6 Hypothesis Testing with ANOVA

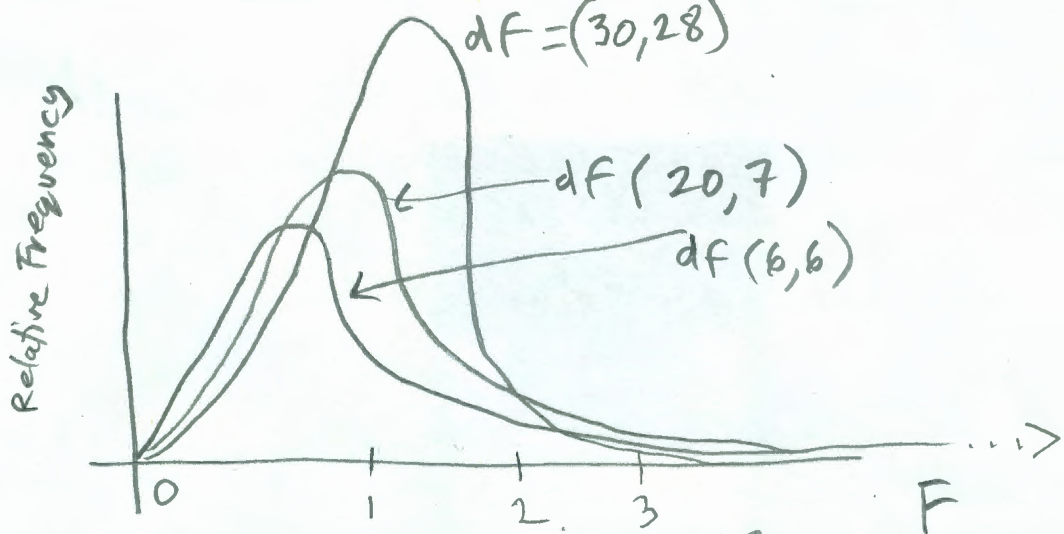
$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

H_a: "Not all population means are equal"

→ If H_0 Rejected, we cannot conclude that all means are different, only that "at least 2 pop. means are different"

→ ANOVA is procedure to determine if we reject H_0.

7 F Distribution used in ANOVA table



① Family of F Distributions (many F Distributions)

② df in numerator = 1st estimate of Variance
df in denominator = 2nd estimate of variance

③ F Distribution is continuous from 0 to ∞

④ F Distribution cannot be negative (0 1st number)

⑤ Positive skew (long tail to right, as both df increase, approach Normal)

⑥ Used when you want to test whether two samples are from populations having equal variances.

⑦ Used when calculating ANOVA table

⑧ For F Critical value in Excel use:
F.INV.RT (alpha, Degrees of Freedom for MSTR, Degrees of Freedom for MSE)

⑨ P-value for F Hypothesis Test in Excel use:
F.DIST.RT ($\frac{MSTR}{MSR}$, Degrees of Freedom for MSTR, Degrees of Freedom for MSE)

8 How to do ANOVA in Excel

Next 3 pages →

	A	B	C	D	E
1	$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$				
2	Ha = Not all population means are the same (at least two are different)				
3	Decision Rule: p-value < alpha, Reject H0 & Accept Ha or Test Statistic F > Critical F,				
4	Reject H0 & Accept Ha, otherwise we fail to Reject H0				
5		Population 1	Population 2	Population 3	Population 4
6		Treatment 1	Treatment 2	Treatment 3	Treatment 4
7		Method 1	Method 2	Method 3	Method 4
8		12.56	11.87	15.57	12.94
9		15.92	12.94	10.64	12.01
10		10.42	11.23	14.05	10.21
11		14.41	15.97	13.08	10.18
12		10.32	17.14	16.71	11.21
13	Sample Size = n =	5	5	5	5
14	Sample Size = Xbar =	12.726	13.83	14.01	11.31
15	Sample Variance = s_2 =	6.04288	6.73235	5.48675	1.41095
16					
17	Grand Over All Count = n_T =	20			
18	Grand Over All Sample Mean = XbarG =	12.969	**Check:**	12.969	
19	# Treatments = k =	4			
20	Calculations for **Total** Variance for All Observations =				
21	Sum of Squares Total = SST =	101.87338	**Check:**	101.87338	101.87338
22	df_T =	19			
23	Calculations for Variance **Between** Treatments =				
24	Sum of Squares of Treatments = SSTR =	23.18166			
25	df_{TR} =	3			
26	Estimate of Variance **Between** Treatments = MSTR = SSTR/df_{TR}	7.72722			
27	Calculations for Variance **Within** Treatments =				
28	Sum of Squares Error = SSE =	78.69172	**Check:**	78.69172	
29	df_E =	16			
30	Estimate of Variance **Within** Treatments = MSE = SSE/df_E	4.9182325			
31	F Test Statistic = MSTR/MSE =	1.571137599			
32	Alpha =	0.05			
33	p-value =	0.235295845			
34	F Critical	3.238871517			
35					
36	Because the F Calculated Test Statistic is not past our F Critical, We Fail To Reject H0.				
37	It seems reasonable to assume that the pop means are the same.				

1 Quad Example Final Results

	A	B	C	D	E
1					
2					
3					
4					
5		Population 1	Population 2	Population 3	Population 4
6		Treatment 1	Treatment 2	Treatment 3	Treatment 4
7		Method 1	Method 2	Method 3	Method 4
8		12.56	11.87	15.57	12.94
9		15.92	12.94	10.64	12.01
10		10.42	11.23	14.05	10.21
11		14.41	15.97	13.08	10.18
12		10.32	17.14	16.71	11.21
13	Sample Size = n =	=COUNT(B8:B12)	=COUNT(C8:C12)	=COUNT(D8:D12)	=COUNT(E8:E12)
14	Sample Size = Xbar =	=AVERAGE(B8:B12)	=AVERAGE(C8:C12)	=AVERAGE(D8:D12)	=AVERAGE(E8:E12)
15	Sample Variance = s_2 =	=VAR.S(B8:B12)	=VAR.S(C8:C12)	=VAR.S(D8:D12)	=VAR.S(E8:E12)
16					
17	Grand Over All Count = n_T =	=COUNT(B8:E12)			
18	Grand Over All Sample Mean = XbarG =	=AVERAGE(B8:E12)	**Check:**	=SUMPRODUCT(B13:E13,B14:E14)/SUM(B13:E13)	
19	# Treatments = k =	=COUNTA(B7:E7)			
20	Calculations for **Total** Variance for All Observations =				
21	Sum of Squares Total = SST =	=SUMPRODUCT((B8:E12-B18)^2)	**Check:**	=B24+B28	=VAR.P(B8:E12)*B17
22	df_T =	=B17-1			
23	Calculations for Variance **Between** Treatments =				
24	Sum of Squares of Treatments = SSTR =	=SUMPRODUCT((B14:E14-B18)^2,B13:E13)			
25	df_{TR} =	=B19-1			
26	Estimate of Variance **Between** Treatments = MSTR = SSTR/df_{TR}	=B24/B25			
27	Calculations for Variance **Within** Treatments =				
28	Sum of Squares Error = SSE =	=SUMPRODUCT(B15:E15,B13:E13-1)	**Check:**	=SUMPRODUCT((B8:E12-B14:E14)^2)	
29	df_E =	=B17-B19			
30	Estimate of Variance **Within** Treatments = MSE = SSE/df_E	=B28/B29			
31	F Test Statistic = MSTR/MSE =	=B26/B30			
32	Alpha =	0.05			
33	p-value =	=F.DIST.RT(B31,B25,B29)			
34	F Critical	=F.INV.RT(B32,B25,B29)			

P.36

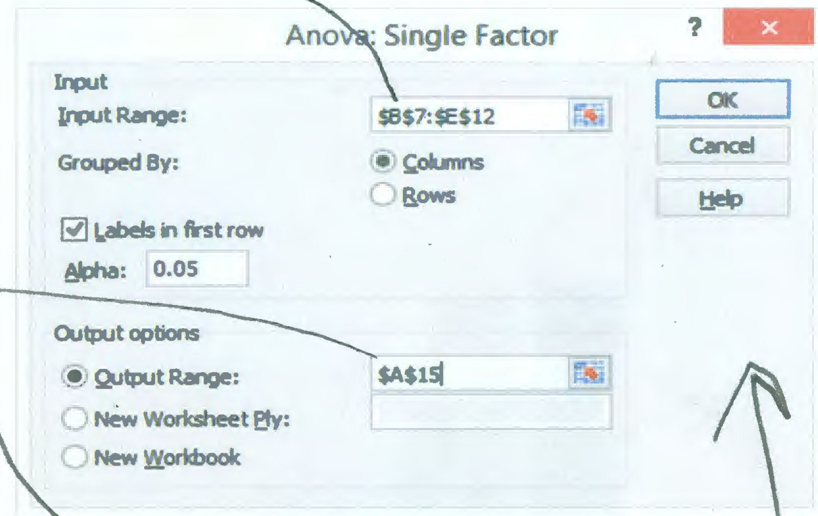
(2) Excel Formulas

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

Ha = Not all population means are the same (at least two are different)

Decision Rule: p-value < alpha, Reject H0 & Accept Ha or Test Statistic F > Critical F, Reject H0 & Accept Ha, otherwise we fail to Reject H0

Population 1	Population 2	Population 3	Population 4
Treatment 1	Treatment 2	Treatment 3	Treatment 4
Method 1	Method 2	Method 3	Method 4
12.56	11.87	15.57	12.94
15.92	12.94	10.64	12.01
10.42	11.23	14.05	10.21
14.41	15.97	13.08	10.18
10.32	17.14	16.71	11.21



Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Method 1	5	63.63	12.726	6.04288
Method 2	5	69.15	13.83	6.73235
Method 3	5	70.05	14.01	5.48675
Method 4	5	56.55	11.31	1.41095

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	23.18166	3	7.72722	1.571137599	0.235296	3.238871517
Within Groups	78.69172	16	4.9182325			
Total	101.87338	19				

Because our p-value is larger than , We Fail To Reject H0. It seems reasonable to assume that the pop means are the same.

OUTPUT

3 Excel, Data Ribbon Tab, Data Analysis button, "Anova: Single Factor"

Excel

11 Analysis of variance to test equality of K population means for completely Randomized Design / Book Formulas:

① $H_0: \mu_1 = \mu_2 = \dots = \mu_K$

H_a: Not all population means are equal

② Variables:

j = particular treatment "jth Treatment"

μ_j = Pop. mean of j^{th} population

K = # of treatments / populations

σ^2 = {common population variance}

n_j = Random sample size for jth pop = # observations

X_{ij} = Value of obervational i for treatment j

$\bar{X}_j$ = sample mean for treatment j

s_j^2 = sample Variance for treatment j

s_j = sample standard Deviation for treatment j

$\bar{\bar{X}}$ = overall sample mean

③ Formulas:

$n_T = n_1 + n_2 + \dots + n_K$

$\boxed{1}$ $$\bar{X}_j = \frac{\sum_{i=1}^{n_j} X_{ij}}{n_j}$$

$\boxed{2}$ $$s_j^2 = \frac{\sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{n_j - 1}$$

AVERAGE STDEV.S Functions.

$\boxed{3}$ $$\bar{\bar{X}} = \frac{\sum_{j=1}^{K} \sum_{i=1}^{n_j} x_{ij}}{n_T}$$

or if all n_j (sample size) are same → $$\frac{\sum_{j=1}^{K} \bar{X}_j}{K}$$

* $\bar{\bar{X}}$ → if sample means provided & sample size not same, we can do a weighted mean.

$$\bar{\bar{x}} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + \dots + n_K\bar{x}_K}{n_T}$$

③ Formulas contined:

Between - Treatments Estimate of population Variance (σ^2) (sample sizes equal)

$$\left\{\begin{matrix}\text{variability} \\ \text{Between /} \\ \text{Among} \\ \text{samples}\end{matrix}\right\} = \left\{\begin{matrix}\text{Mean Square} \\ \text{due to} \\ \text{Treatments}\end{matrix}\right\} = MSTR = \frac{\sum_{j=1}^{K} n_j (\bar{x}_j - \bar{\bar{x}})^2}{K-1}$$

$\boxed{4}$ $$SSTR = \left\{\begin{matrix}\text{Sum of Squares} \\ \text{due to} \\ \text{Treatments}\end{matrix}\right\} = \sum_{J=1}^{K} n_j (\bar{x}_j - \bar{\bar{x}})^2$$

treatment j mean

overall mean

$\boxed{5}$ $$\left\{\begin{matrix}\text{Degrees of} \\ \text{Freedom} \\ \text{Associated} \\ \text{w/ SSTR}\end{matrix}\right\} = K - 1$$

$\boxed{6}$ $$\left\{\begin{matrix}\text{Mean Square} \\ \text{due to} \\ \text{Treatments}\end{matrix}\right\} = MSTR = \frac{SSTR}{K-1}$$

If H_0 is TRUE ($\mu_1 = \mu_2 = \cdots = \mu_K$) → MSTR is unbiased estimate of σ^2

If H_a is TRUE (Not all pop. means are equal) → ① MSTR is NOT unbiased estimate of σ^2

② MSTR will overestimate σ^2

*, if each sample size is same:

$$MSTR = \frac{\sum_{j=1}^{K} n_j (\bar{x}_j - \bar{\bar{x}})^2}{K-1} = n \left[\frac{\sum_{j=1}^{K} (\bar{x}_j - \bar{\bar{x}})^2}{K-1}\right] = n S_{\bar{x}}^2$$

n * variation of sample Means

③ Formulas Continued:

Within Treatments Estimate of Pop. variance (σ^2) (sample sizes equal)

$$\left\{\begin{matrix}\text{variability} \\ \text{within each} \\ \text{sample}\end{matrix}\right\} = \left\{\begin{matrix}\text{mean square} \\ \text{due to} \\ \text{Error}\end{matrix}\right\} = MSE = \frac{\sum_{j=1}^{K} (n_j - 1) s_j^2}{n_T - K}$$

$\boxed{7}$ $$\left\{\begin{matrix}\text{Sum of Squares} \\ \text{due to} \\ \text{Error}\end{matrix}\right\} = SSE = \sum_{j=1}^{K} (n_j - 1) s_j^2$$

OR Alternative:

$$SSE = \sum_{j=1}^{K} \sum_{i=1}^{n_j} (x_{ij} - \bar{X}_j)^2$$

$\boxed{8}$ $$\left\{\begin{matrix}\text{Degrees of} \\ \text{Freedom} \\ \text{Associated} \\ \text{w/ SSE}\end{matrix}\right\} = n_T - K$$

$\boxed{9}$ $$\left\{\begin{matrix}\text{mean square} \\ \text{due to Error}\end{matrix}\right\} = MSE = \frac{SSE}{n_T - K}$$

Note1: MSE is based on the variation within each of the treatments

Note2: It is not influenced by whether H_0 is TRUE

Note3: Therefore: MSE always provides an unbiased estimate of common pop. variance σ^2

* if sample size is same: then: $n_T = Kn$ & $n_T - K = K(n-1)$

thus:

$$MSE = \frac{\sum_{j=1}^{K} (n-1) S_j^2}{K(n-1)} = \frac{(n-1)\sum_{j=1}^{K} S_j^2}{K(n-1)} = \frac{\sum_{j=1}^{K} S_j^2}{K} = \left\{\begin{matrix}\text{Average} \\ \text{of K} \\ \text{sample} \\ \text{variances}\end{matrix}\right\}$$

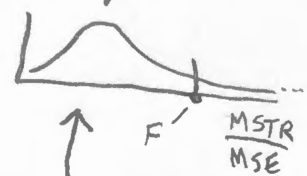
③ Formulas continued:

Comparing common Pop. Variance σ^2 Estimates MSTR & MSE with F Test:

① If H_0 TRUE ($\mu_1 = \mu_2 = \cdots = \mu_K$) MSTR & MSE provide 2 independent, unbiased estimates of σ^2

② If three assumptions for ANOVA (each Response variable has normal Distribution, σ^2 for response variable is same for all pops., observations, x_{ij} are independent) are valid:

1) sampling Distribution of MSTR/MSE is an F Distribution



Sampling Distribution of $\frac{MSTR}{MSE}$

2) Numerator (MSTR) $df = K - 1$

3) Denominator (MSE) $df = n_T - K$

4) If H_0 TRUE, $\frac{MSTR}{MSE}$ should appear to have been selected from F distribution

4) If H_0 FALSE, MSTR overestimates σ^2 & $\frac{MSTR}{MSE}$ will be large

$\boxed{10}$

$$\left\{ \begin{array}{l} \text{Test statistic} \\ \text{for equality of} \\ \text{K population Means} \end{array} \right\} = F = \frac{MSTR}{MSE}$$

3 Formulas continued:

11 $\left\{\begin{matrix}\text{Total sum} \\ \text{of squares}\end{matrix}\right\} = SST = SSTR + SSE =$

$$= \sum_{j=1}^{K} \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2 =$$

If we took $\dfrac{SST}{n_T - 1}$ we would get overall sample variance for all observations

12 ANOVA Table for completely Randomized Design:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	P-value
Treatments	SSTR	$K-1$	$MSTR = \dfrac{SSTR}{K-1}$	$\dfrac{MSTR}{MSE}$	F critical
Error	SSE	$n_T - K$	$MSE = \dfrac{SSE}{n_T - k}$		
Total	SST	$n_T - 1$			

13 $P\text{-value} = P\left(F \geq \dfrac{MSTR}{MSE}\right)$ Reject H_0, Accept H_a if p-value $\leq \alpha$

14 critical value $= F_\alpha$ Reject H_0, Accept H_a if $F \geq F_\alpha$

Testing for Equality of K Pop. Means for an observational Study

Same procedures for experimental studies, but we don't have as much control.

HW #6

$H_0: \mu_1 - \mu_2 >= 0$
$H_a: \mu_1 - \mu_2 < 0$

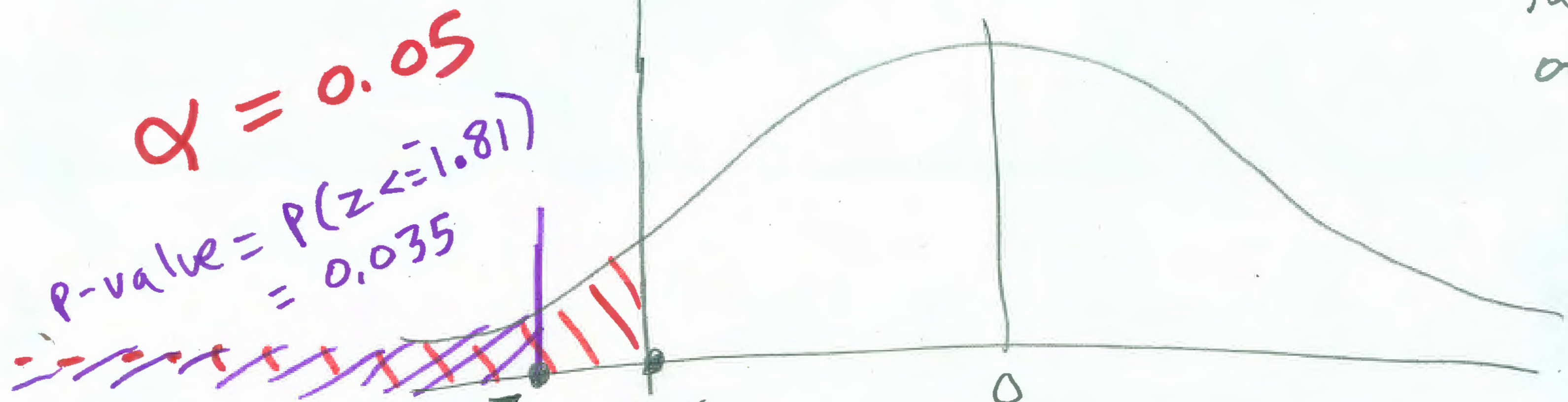
Reject Ho, Accept Ha

Fail to Reject Ho

Sampling Distribution of Xbar₁ − Xbar₂

$\alpha = 0.05$

p-value = $P(z <= -1.81)$
= 0.035



test statistic = −1.81

critical value = −1.645

$\mu_1 - \mu_2$

Z

Xbar₁ − Xbar₂

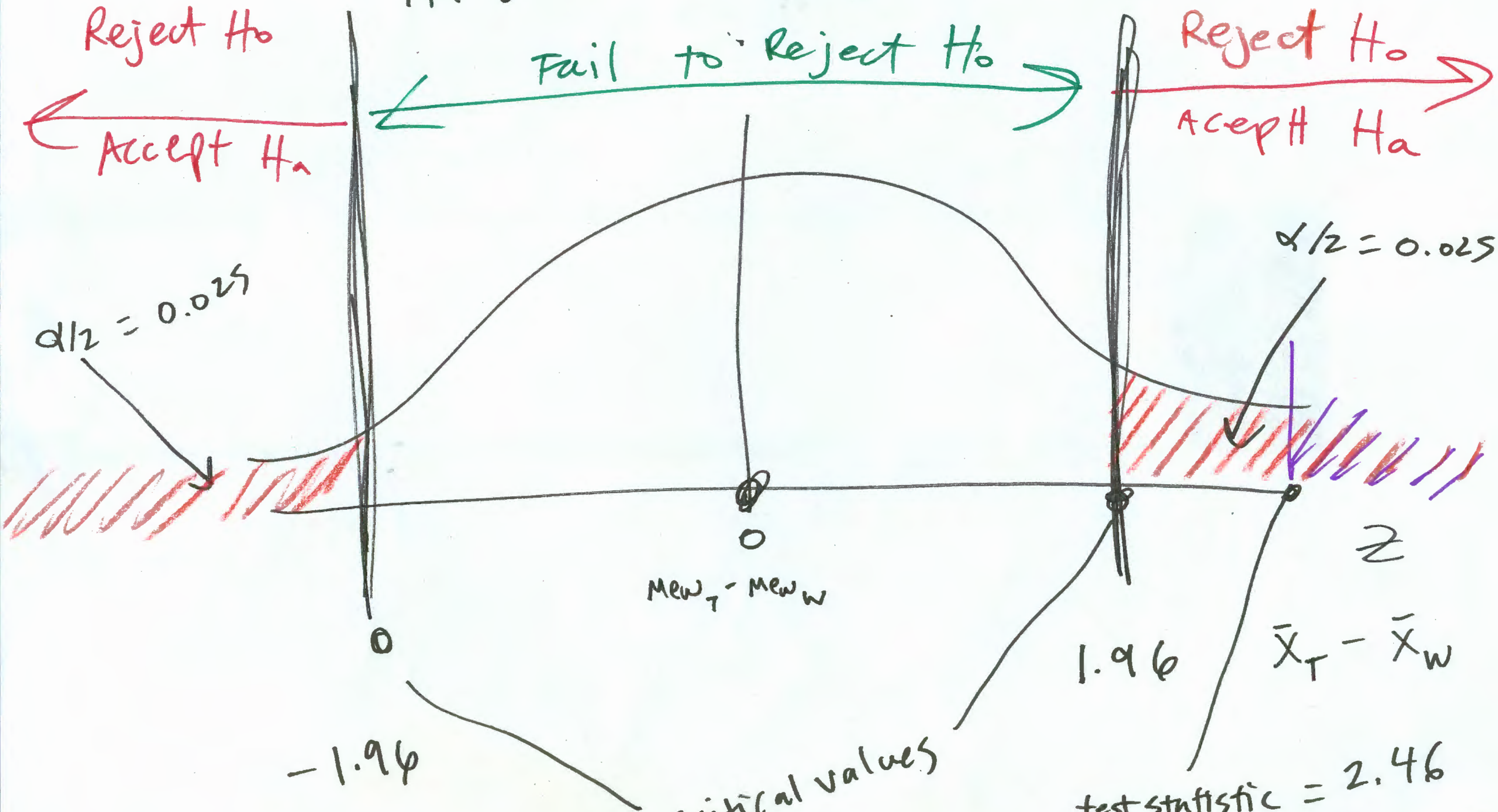
$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$ = Standard Deviation of Sampling Distribution of Xbar₁ − Xbar₂

Sampling Distribution of $Xbar_T - Xbar_W$

HW 7

H_0 : $Mew_T - Mew_H = 0$

H_a : $Mew_T - Mew_H <> 0$

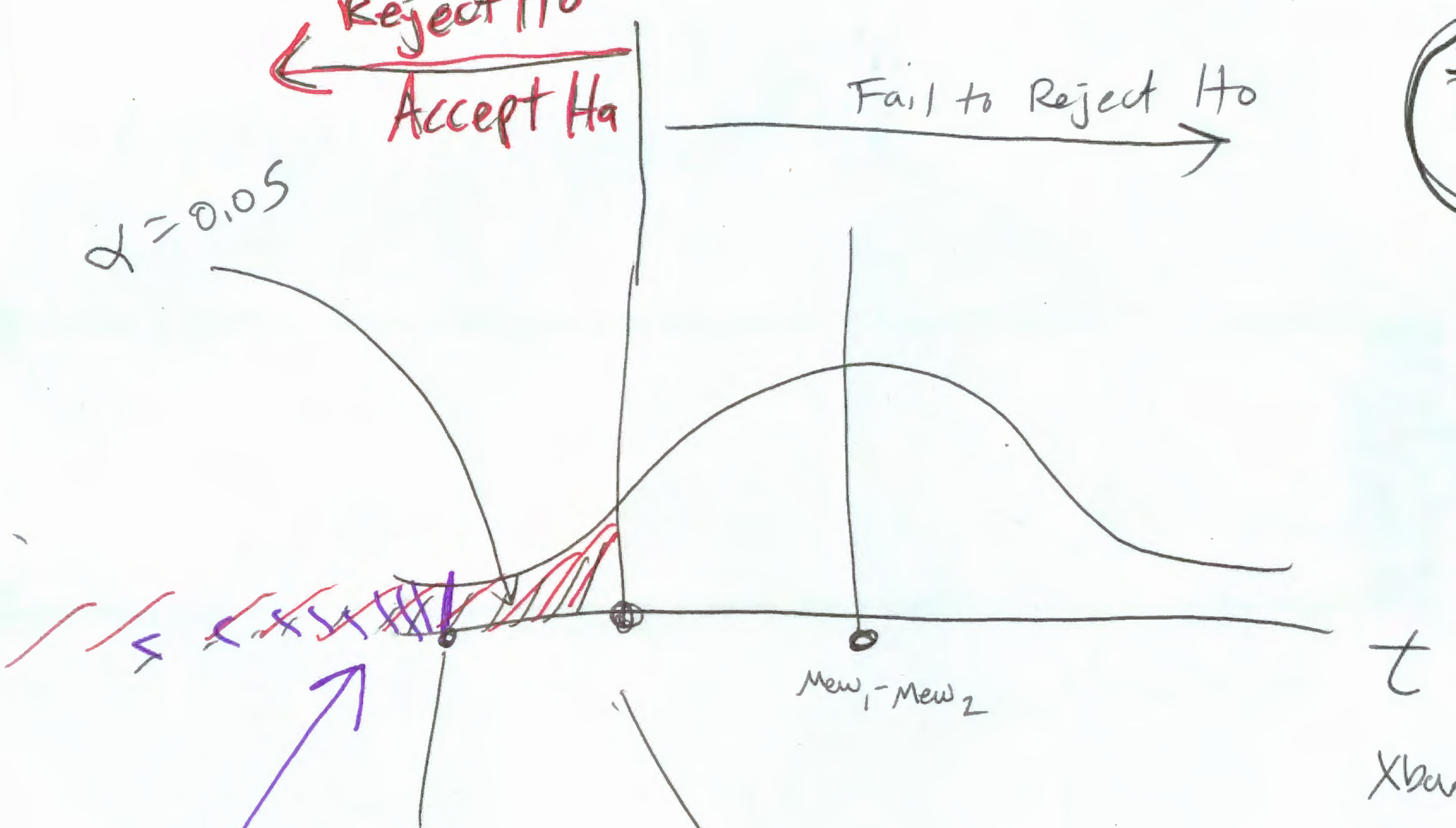


p-value = $P(Xbar_T - Xbar_W \geq 2.46$ OR $Xbar_T - Xbar_W \leq -2.46) = 0.0138$

Sampling Distribution of Differences

t distribution

HW 14



test statistic $= -2.4$

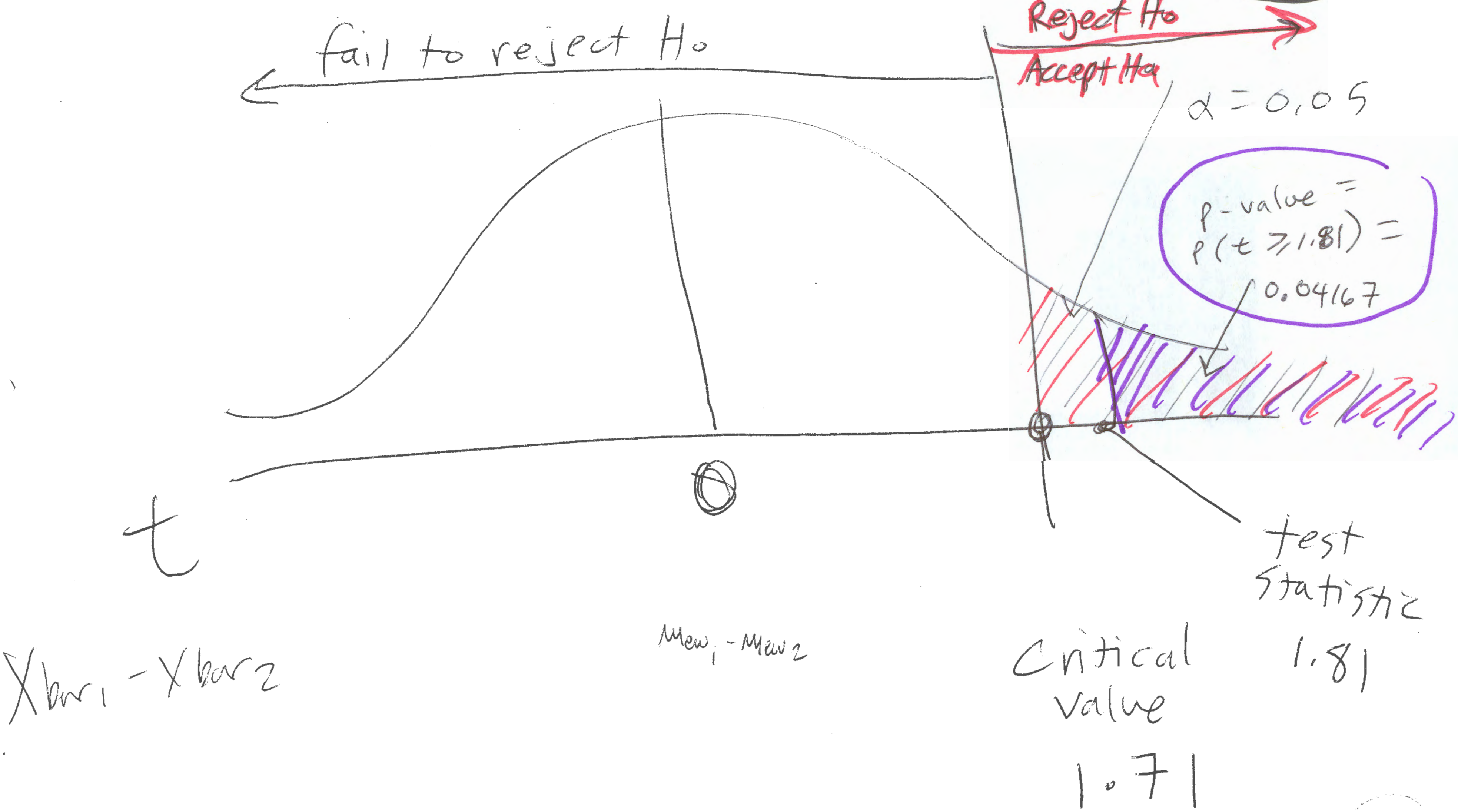
Critical Value -1.66

$p\text{-value} = P(t \le -2.4) = 0.0091$

Sampling Distribution of Differences $\bar{X}_1 - \bar{X}_2$

t Distribution

HW 16



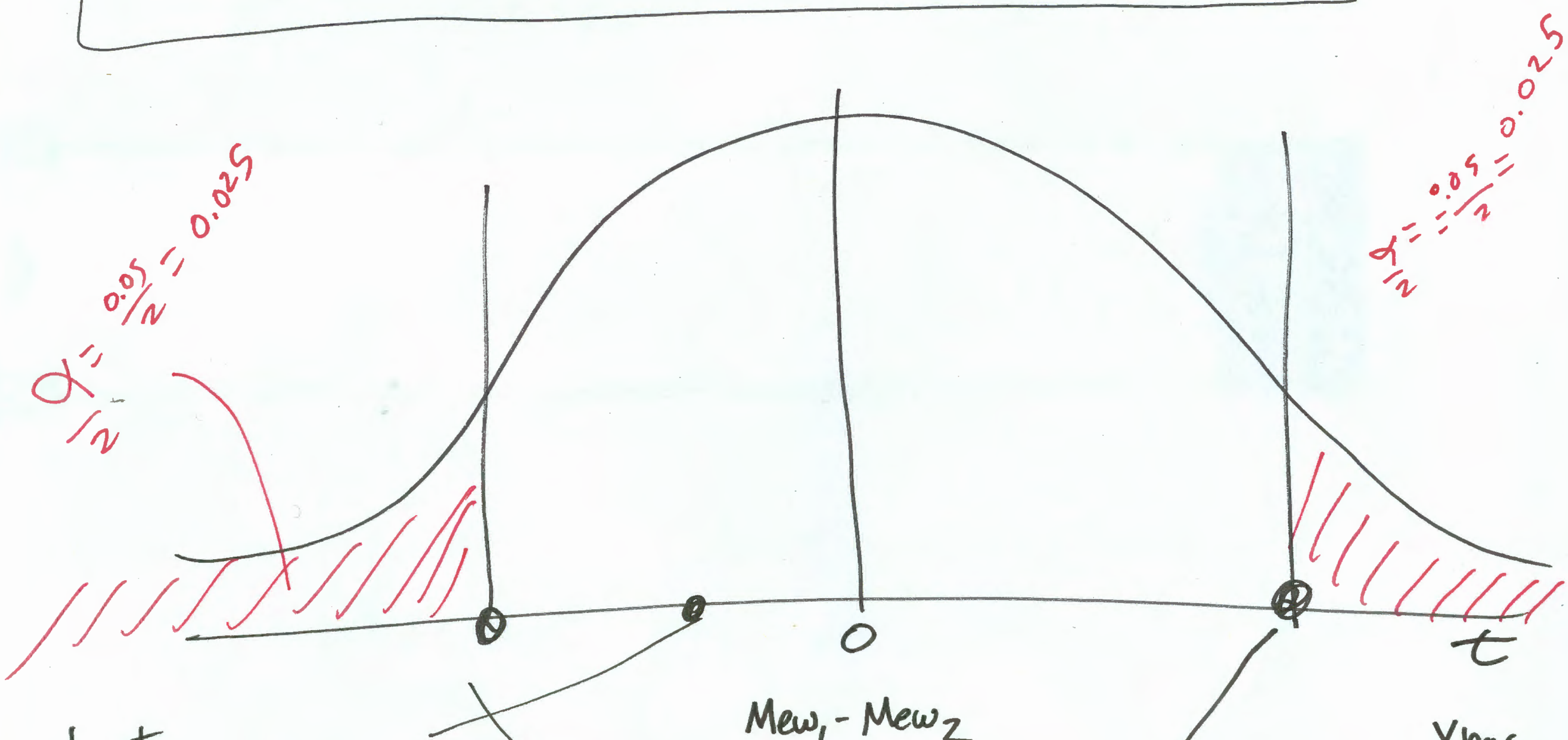
P. 47

HW 18

$H_0 : M_1 - M_2 = 0$

$H_a : M_1 - M_2 <> 0$

Sampling Distribution of Xbar$_1$ − Xbar$_2$



test statistic = -0.32

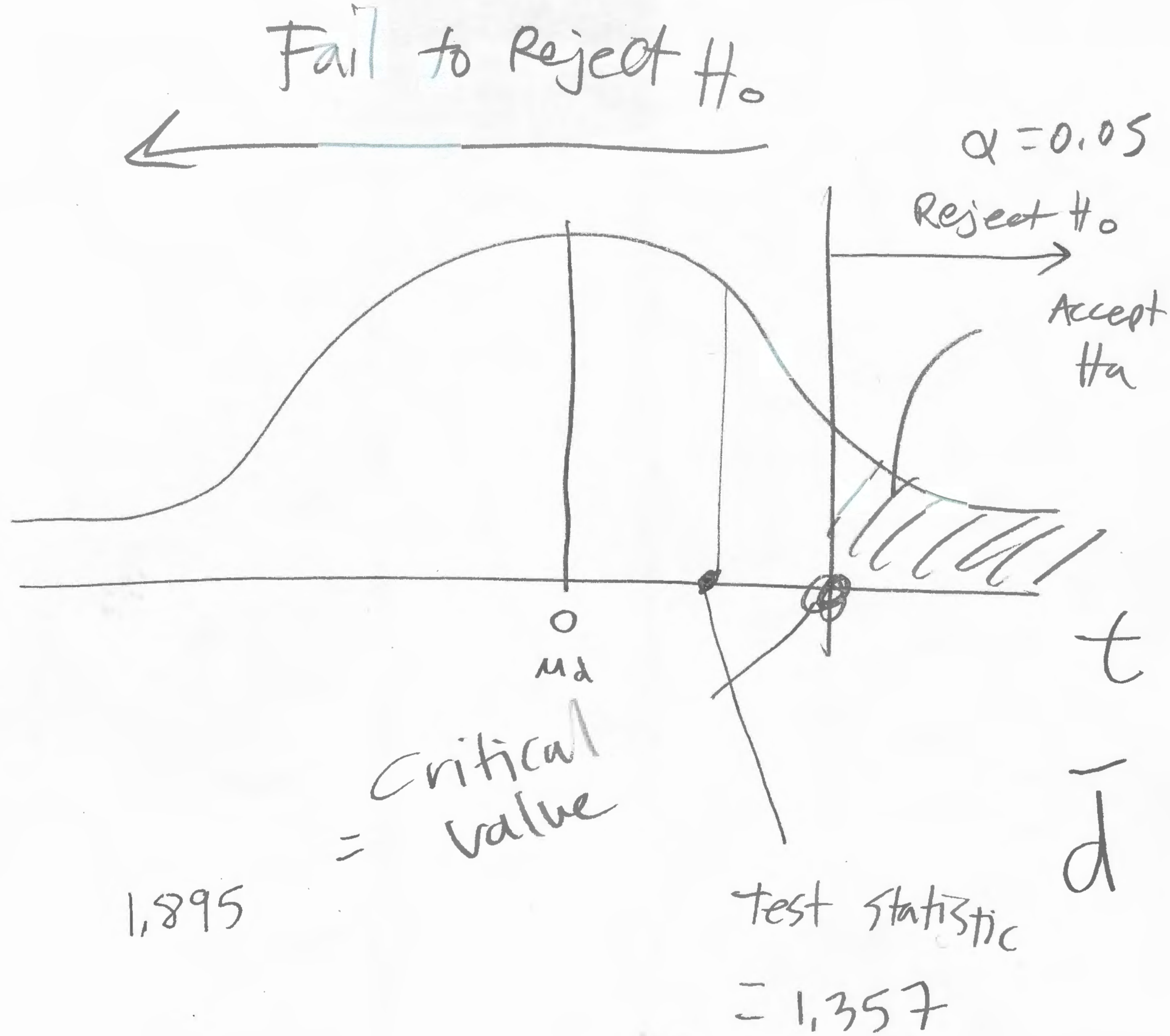
critical values +/- 2.02

p-value = P (t ≤ -0.32 OR t ≥ 0.32) = 0.7535

P.48

P.49

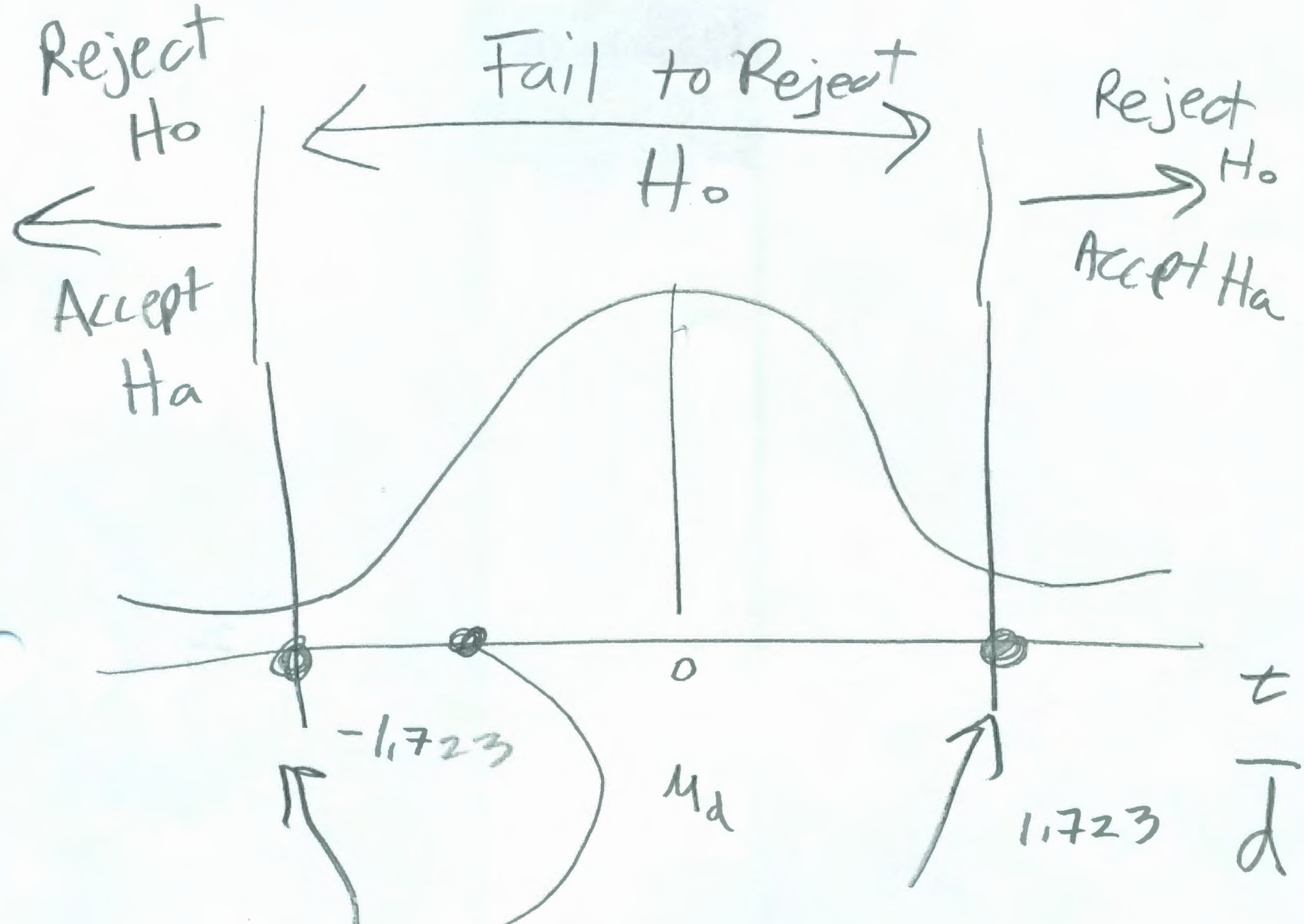
HW
#21



$p\text{-value} = P(t \geq 1.357) = 0.1084$

HW (p. 50)

26



Critical values

test statistic = -1.42

p-value = $p(-1.7 \leq t \geq 1.73) = 0.173$