IE 340/440

PROCESS IMPROVEMENT THROUGH PLANNED EXPERIMENTATION

Two Sample Tests with Numerical Data

Dr. Xueping Li

University of Tennessee

Chapter Topics

- Comparing Two Independent Samples
  - Independent samples Z test for the difference in two means
  - Pooled-variance t test for the difference in two means
- F Test for the Difference in Two Variances
- Comparing Two Related Samples
  - Paired-sample Z test for the mean difference
  - Paired-sample t test for the mean difference

Chapter Topics

(continued)

- Wilcoxon Rank Sum Test
  - Difference in two medians
- Wilcoxon Signed Ranks Test
  - Median difference

Comparing Two Independent Samples

- Different Data Sources
  - Unrelated
  - Independent
    - Sample selected from one population has no effect or bearing on the sample selected from the other population
- Use the Difference between 2 Sample Means
- Use Z Test or Pooled-Variance t Test

Independent Sample Z Test (Variances Known)

- Assumptions
  - Samples are randomly and independently drawn from normal distributions
  - Population variances are known
- Test Statistic
  - $$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2}}}$$

Independent Sample (Two Sample) *Z* Test in Excel

- Independent Sample *Z* Test with Variances Known
  - Tools | Data Analysis | *Z* test: Two Sample for Means

Pooled-Variance *t* Test (Variances Unknown)

- Assumptions
  - Both populations are normally distributed
  - Samples are randomly and independently drawn
  - Population variances are unknown but assumed equal
  - If both populations are not normal, need large sample sizes

Developing the Pooled-Variance t Test

- Setting Up the Hypotheses

$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	**OR**	$H_0: \mu_1 - \mu_2 = 0$ $H_1: \mu_1 - \mu_2 \neq 0$	**Two Tail**
$H_0: \mu_1 \leq \mu_2$ $H_1: \mu_1 > \mu_2$	**OR**	$H_0: \mu_1 - \mu_2 \leq 0$ $H_1: \mu_1 - \mu_2 > 0$	**Right Tail**
$H_0: \mu_1 \geq \mu_2$ $H_1: \mu_1 < \mu_2$	**OR**	$H_0: \mu_1 - \mu_2 \geq 0$ $H_1: \mu_1 - \mu_2 < 0$	**Left Tail**

Developing the Pooled-Variance t Test

(continued)

- Calculate the Pooled Sample Variance as an Estimate of the Common Population Variance

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

S_p^2 : Pooled sample variance

n_1 : Size of sample 1

S_1^2 : Variance of sample 1

n_2 : Size of sample 2

S_2^2 : Variance of sample 2

Developing the Pooled-Variance *t* Test

(continued)

- Compute the Sample Statistic

$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Hypothesized difference

$df = n_1 + n_2 - 2$

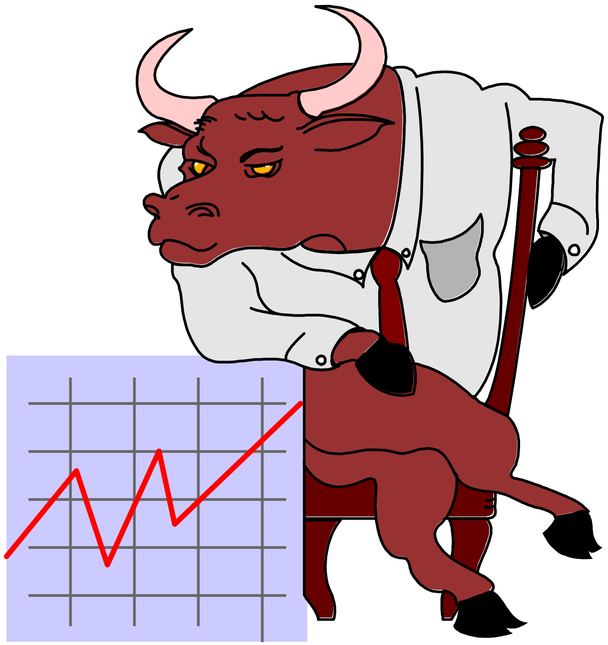
$$S_p^2 = \frac{\left(n_1 - 1\right)S_1^2 + \left(n_2 - 1\right)S_2^2}{\left(n_1 - 1\right) + \left(n_2 - 1\right)}$$

Pooled-Variance t Test: Example

You're a financial analyst for Charles Schwab. Is there a difference in average dividend yield between stocks listed on the NYSE & NASDAQ? You collect the following data:

	NYSE	NASDAQ
Number	21	25
Sample Mean	3.27	2.53
Sample Std Dev	1.30	1.16

Assuming equal variances, is there a difference in average yield (α = **0.05**)?



Calculating the Test Statistic

$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\left(3.27 - 2.53\right) - 0}{\sqrt{1.510\left(\frac{1}{21} + \frac{1}{25}\right)}} = 2.03$$

$$S_p^2 = \frac{\left(n_1 - 1\right)S_1^2 + \left(n_2 - 1\right)S_2^2}{\left(n_1 - 1\right) + \left(n_2 - 1\right)}$$

$$= \frac{\left(21 - 1\right)1.30^2 + \left(25 - 1\right)1.16^2}{\left(21 - 1\right) + \left(25 - 1\right)} = 1.502$$

Solution

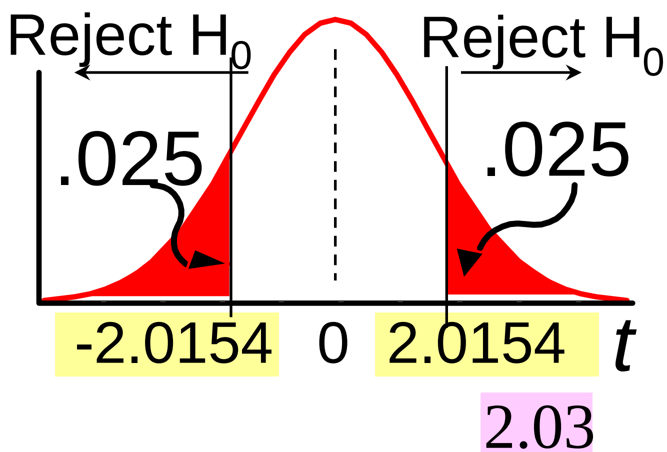
H_0: $\mu_1 - \mu_2 = 0$ i.e. ($\mu_1 = \mu_2$)

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. ($\mu_1 \neq \mu_2$)

$\alpha = 0.05$

df = 21 + 25 - 2 = 44

Critical Value(s):



Test Statistic:

$$t = \frac{3.27 - 2.53}{\sqrt{1.502\left(\frac{1}{21} + \frac{1}{25}\right)}} = 2.03$$

Decision:

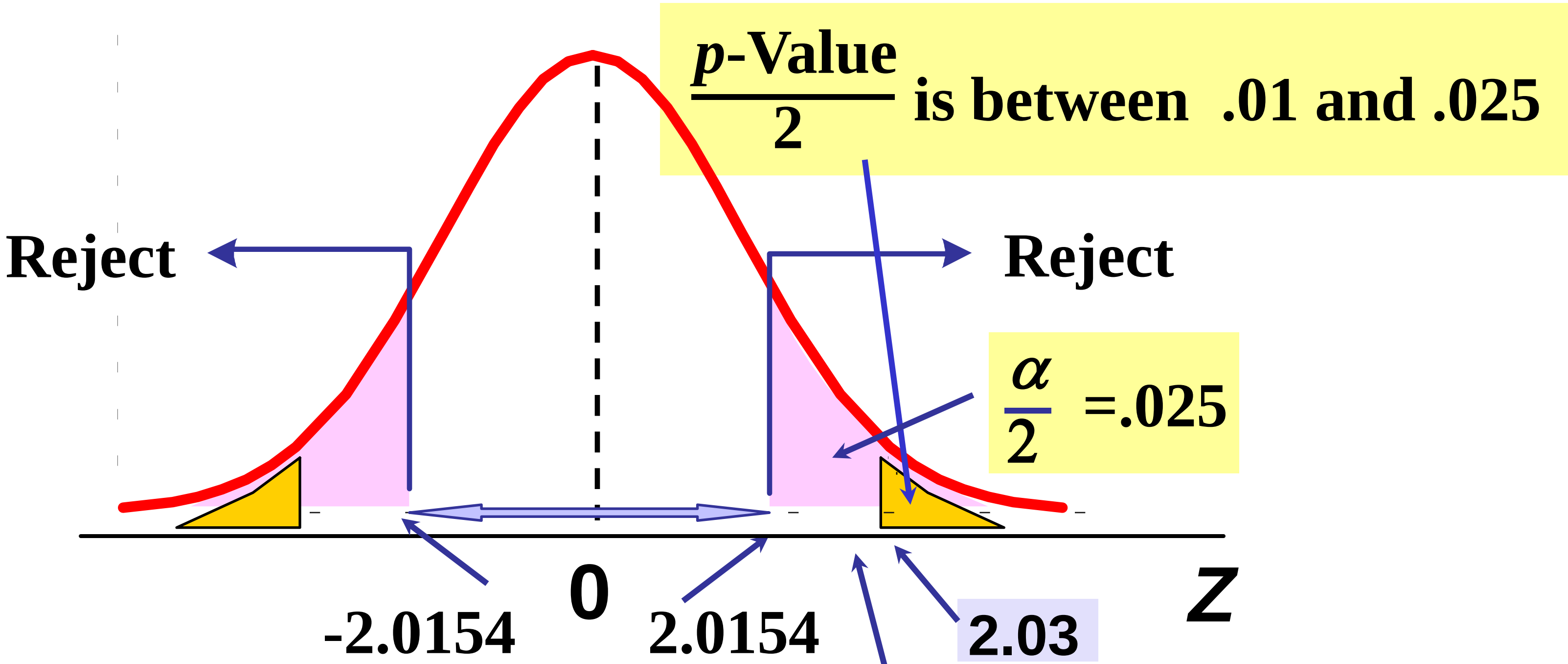
Reject at $\alpha = 0.05$.

Conclusion:

There is evidence of a difference in means.

p-Value Solution

(p-Value is between .02 and .05) < (α = 0.05) Reject.



Test Statistic 2.03 is in the Reject Region

Pooled-Variance *t* Test in PHStat and Excel

- If the Raw Data are Available:
  - Tools | Data Analysis | *t* Test: Two-Sample Assuming Equal Variances
- If Only Summary Statistics are Available:
  - PHStat | Two-Sample Tests | *t* Test for Differences in Two Means...

Solution in Excel

- Excel Workbook that Performs the Pooled-Variance *t* Test

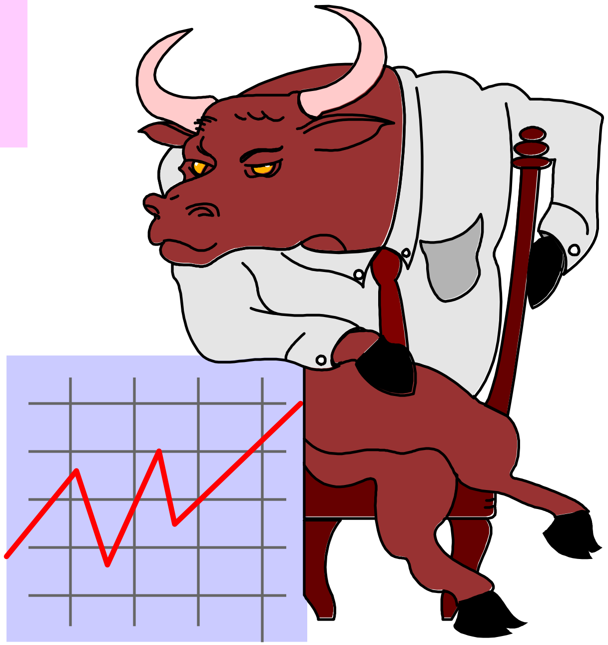
Microsoft Excel Worksheet

Example

You're a financial analyst for Charles Schwab. You collect the following data:

	NYSE	NASDAQ
Number	**21**	**25**
Sample Mean	**3.27**	**2.53**
Sample Std Dev	**1.30**	**1.16**

You want to construct a 95% confidence interval for the difference in population average yields of the stocks listed on NYSE and NASDAQ.



Example: Solution

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

$$= \frac{(21-1)1.30^2 + (25-1)1.16^2}{(21-1) + (25-1)} = 1.502$$

$$\left(\bar{X}_1 - \bar{X}_2\right) \pm t_{\alpha/2, n_1+n_2-2} \sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$(3.27 - 2.53) \pm 2.0154 \sqrt{1.502 \left(\frac{1}{21} + \frac{1}{25}\right)}$$

$$0.0088 \le \mu_1 - \mu_2 \le 1.4712$$

Solution in Excel

- An Excel Spreadsheet with the Solution:

Microsoft
xcel Workshe

F Test for Difference in Two Population Variances

- Test for the Difference in 2 Independent Populations
- Parametric Test Procedure
- Assumptions
  - Both populations are normally distributed
    - Test is not robust to this violation
  - Samples are randomly and independently drawn

The *F* Test Statistic

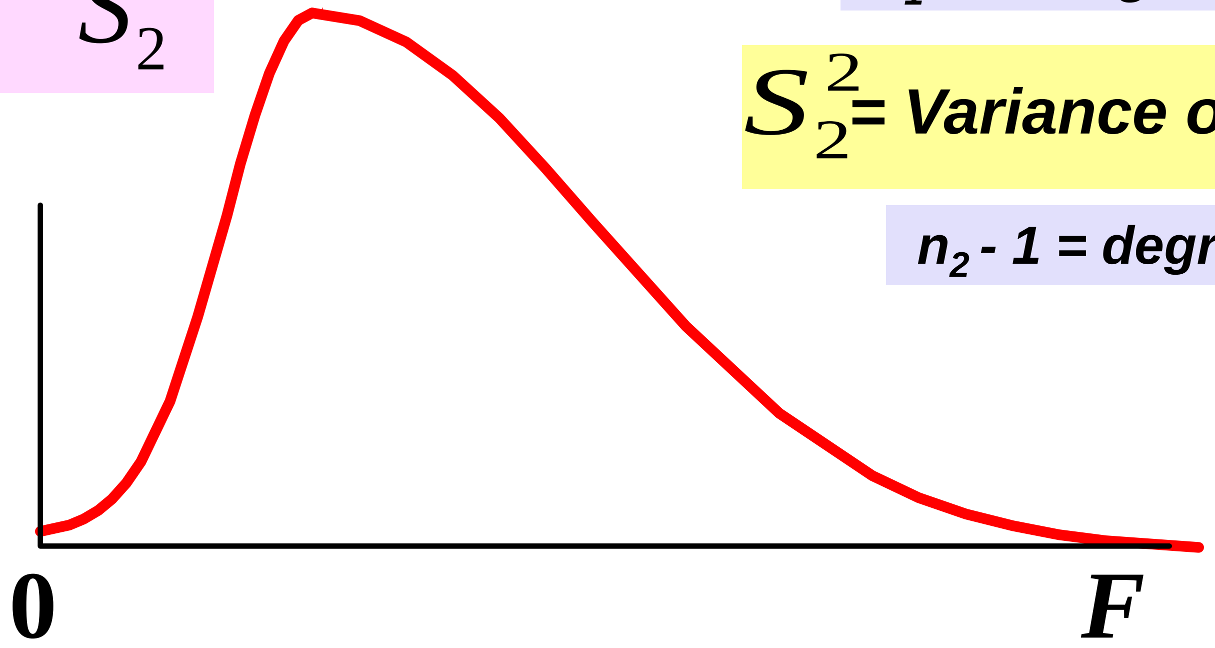
$$F = \frac{S_1^2}{S_2^2}$$

S_1^2 = ***Variance of Sample 1***

n_1 - 1 = degrees of freedom

S_2^2 = ***Variance of Sample 2***

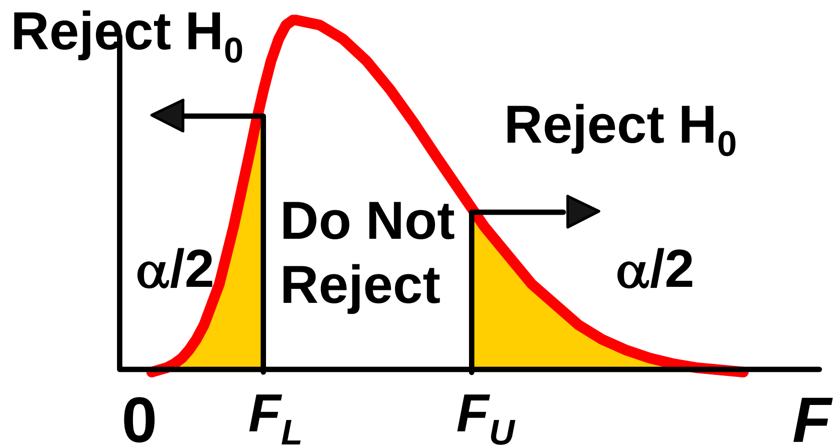
n_2 - 1 = degrees of freedom



Developing the F Test

- **Hypotheses**
  - H_0: $\sigma_1^2 = \sigma_2^2$
  - H_1: $\sigma_1^2 \neq \sigma_2^2$
- **Test Statistic**
  - $F = S_1^2 / S_2^2$
  - **Two Sets of Degrees of Freedom**
    - $\text{df}_1 = n_1 - 1$; $\text{df}_2 = n_2 - 1$
  - **Critical Values:** $F_{L(n_1-1,\, n_2-1)}$ and $F_{U(n_1-1,\, n_2-1)}$

$F_L = 1/F_U^*$ (*degrees of freedom switched)



F Test: An Example

Assume you are a financial analyst for Charles Schwab. You want to compare dividend yields between stocks listed on the NYSE & NASDAQ. You collect the following data:

	NYSE	NASDAQ
Number	**21**	**25**
Mean	**3.27**	**2.53**
Std Dev	**1.30**	**1.16**

Is there a difference in the variances between the NYSE & NASDAQ at the $\alpha = 0.05$ level?

F Test: Example Solution

- Finding the Critical Values for α = .05
  - $df_1 = n_1 - 1 = 21 - 1 = 20$

    $df_2 = n_2 - 1 = 25 - 1 = 24$
  - $F_{L(20,24)} = 1/F_{U(24,20)} = 1/2.41 = .415$

    $F_{U(20,24)} = 2.33$

F Test: Example Solution

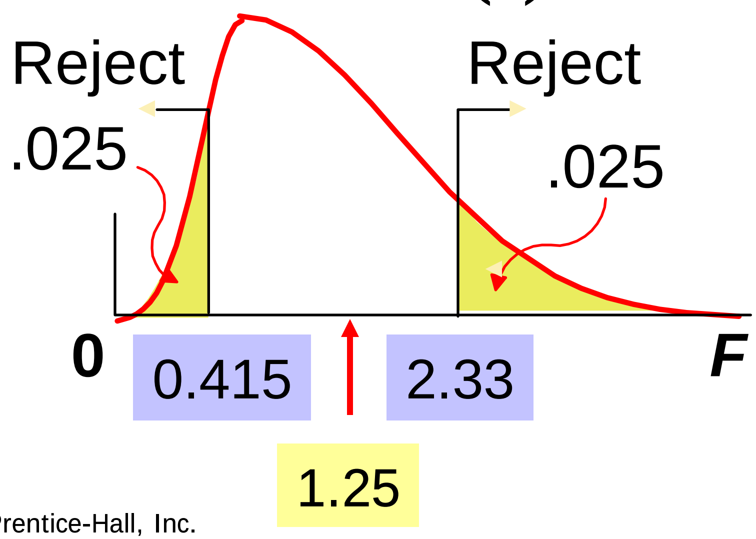
H_0: $\sigma_1^2 = \sigma_2^2$

H_1: $\sigma_1^2 \neq \sigma_2^2$

$\alpha = .05$

df$_1$ = 20 **df$_2$** = 24

Critical Value(s):



Test Statistic:

$$F = \frac{S_1^2}{S_2^2} = \frac{1.30^2}{1.16^2} = 1.25$$

Decision:

Do not reject at α = 0.05.

Conclusion:

There is insufficient evidence to prove a difference in variances.

F Test in PHStat

- PHStat | Two-Sample Tests | *F* Test for Differences in Two Variances
- Example in Excel Spreadsheet

Microsoft Excel
Worksheet

F Test: One-Tail

$H_0: \sigma_1^2 \geq \sigma_2^2$

$H_1: \sigma_1^2 < \sigma_2^2$

or

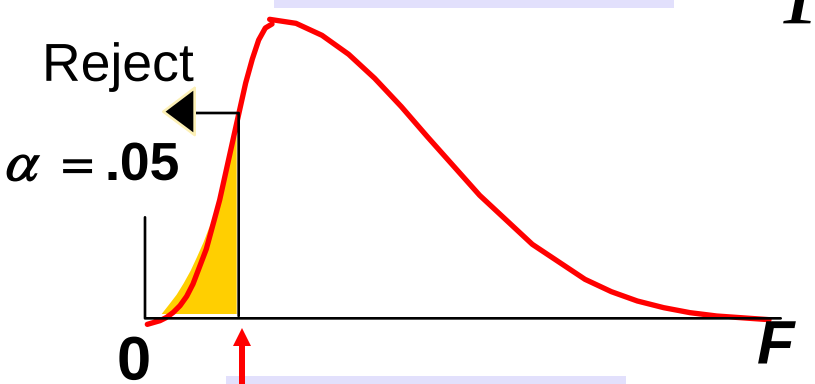
$H_0: \sigma_1^2 \leq \sigma_2^2$

$H_1: \sigma_1^2 > \sigma_2^2$

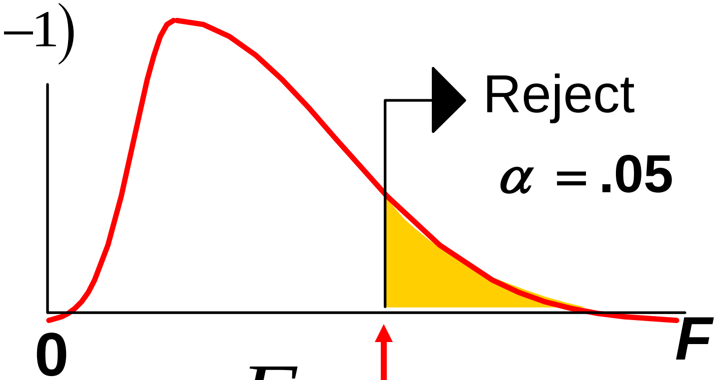
$\alpha = .05$

$$F_{L(n_1-1,n_2-1)} = \frac{1}{F_{U(n_2-1,n_1-1)}}$$

Degrees of freedom switched



$F_{L(n_1-1,n_2-1)}$



$F_{U(n_1-1,n_2-1)}$

Comparing Two Related Samples

- Test the Means of Two Related Samples
  - Paired or matched
  - Repeated measures (before and after)
  - Use difference between pairs

$$D_i = X_{1i} - X_{2i}$$

- Eliminates Variation between Subjects

Z Test for Mean Difference (Variance Known)

- Assumptions
  - Both populations are normally distributed
  - Observations are paired or matched
  - Variance known
- Test Statistic
  - $$Z = \frac{\bar{D} - \mu_D}{\frac{\sigma_D}{\sqrt{n}}} \qquad \bar{D} = \frac{\sum_{i=1}^{n} D_i}{n}$$

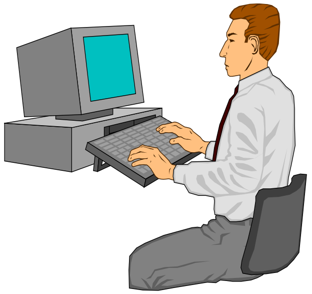
t Test for Mean Difference (Variance Unknown)

- Assumptions
  - Both populations are normally distributed
  - Observations are matched or paired
  - Variance unknown
  - If population not normal, need large samples
- Test Statistic

$$t = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \qquad \bar{D} = \frac{\sum_{i=1}^{n} D_i}{n} \qquad S_D = \sqrt{\frac{\sum_{i=1}^{n} (D_i - \bar{D})^2}{n-1}}$$

Paired-Sample *t* Test: Example

Assume you work in the finance department. Is the new financial package faster (α=0.05 level)? You collect the following processing times:



User	Existing System (1)	New Software (2)	Difference D_i
C.B.	9.98 Seconds	9.88 Seconds	.10
T.F.	9.88	9.86	.02
M.H.	9.84	9.75	.09
R.K.	9.99	9.80	.19
M.O.	9.94	9.87	.07
D.S.	9.84	9.84	.00
S.S.	9.86	9.87	- .01
C.T.	10.12	9.98	.14
K.T.	9.90	9.83	.07
S.Z.	9.91	9.86	.05

$$\bar{D} = \frac{\sum D_i}{n} = .072$$

$$S_D = \sqrt{\frac{\sum \left(D_i - \bar{D}\right)^2}{n-1}}$$

$$= .06215$$

Paired-Sample t Test: Example Solution

Is the new financial package faster (0.05 level)?

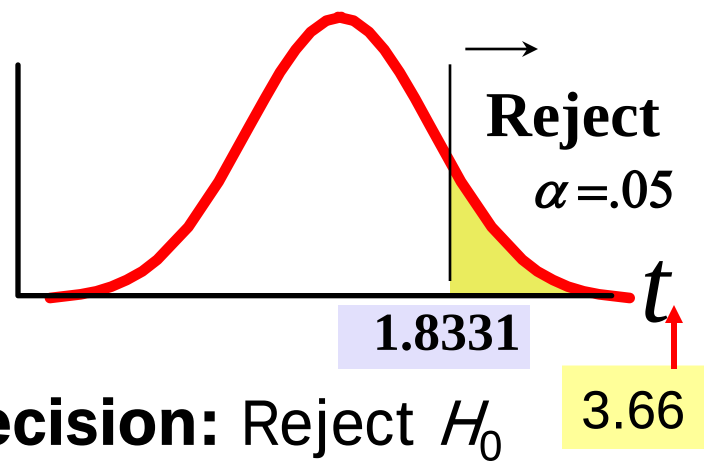
H_0: $\mu_D \leq 0$
H_1: $\mu_D > 0$

$\alpha = .05$ $\bar{D} = .072$

Critical Value=1.8331
df = n - 1 = 9

Test Statistic

$$t = \frac{\bar{D} - \mu_D}{S_D / \sqrt{n}} = \frac{.072 - 0}{.06215 / \sqrt{10}} = 3.66$$



Decision: Reject H_0

t Stat. in the rejection zone.

Conclusion: The new software package is faster.

Paired-Sample *t* Test in Excel

- Tools | Data Analysis... | *t* test: Paired Two Sample for Means
- Example in Excel Spreadsheet

Confidence Interval Estimate for μ_D of Two Related Samples

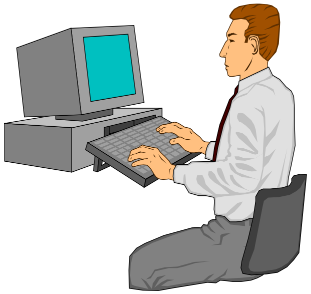
- Assumptions
  - Both populations are normally distributed
  - Observations are matched or paired
  - Variance is unknown
- $100(1-\alpha)\%$ Confidence Interval Estimate:

$$\bar{D} \pm t_{\alpha/2,n-1} \frac{S_D}{\sqrt{n}}$$

Example

Assume you work in the finance department. You want to construct a 95% confidence interval for the mean difference in data entry time. You collect the following processing times:

User	Existing System (1)	New Software (2)	Difference D_i
C.B.	9.98 Seconds	9.88 Seconds	.10
T.F.	9.88	9.86	.02
M.H.	9.84	9.75	.09
R.K.	9.99	9.80	.19
M.O.	9.94	9.87	.07
D.S.	9.84	9.84	.00
S.S.	9.86	9.87	- .01
C.T.	10.12	9.98	.14
K.T.	9.90	9.83	.07
S.Z.	9.91	9.86	.05



Solution:

$$\bar{D} = \frac{\sum D_i}{n} = .072 \qquad S_D = \sqrt{\frac{\sum \left(D_i - \bar{D}\right)^2}{n-1}} = .06215$$

$$t_{\alpha/2,n-1} = t_{0.025,9} = 2.2622$$

$$\bar{D} \pm t_{\alpha/2,n-1} \frac{S_D}{\sqrt{n}}$$

$$.072 \pm 2.2622 \left(\frac{.06215}{\sqrt{10}} \right)$$

$$0.0275 < \mu_{\mathrm{D}} < 0.1165$$

Wilcoxon Rank Sum Test for Differences in 2 Medians

- Test Two Independent Population Medians
- Populations Need Not Be Normally Distributed
- Distribution Free Procedure
- Used When Only Rank Data is Available
- Can Use Normal Approximation if $n_j > 10$ for at least One j

Wilcoxon Rank Sum Test: Procedure

- Assign Ranks, R_i, to the $n_1 + n_2$ Sample Observations
  - If unequal sample sizes, let n_1 refer to smaller-sized sample
  - Smallest value $R_i = 1$
  - Assign average rank for ties
- Sum the Ranks, T_j, for Each Sample
- Obtain Test Statistic, T_1 (Smallest Sample)

Wilcoxon Rank Sum Test: Setting of Hypothesis

Two-Tail Test	Left-Tail Test	Right-Tail Test
H_0: $M_1 = M_2$ H_1: $M_1 \neq M_2$	H_0: $M_1 \geq M_2$ H_1: $M_1 < M_2$	H_0: $M_1 \leq M_2$ H_1: $M_1 > M_2$
Reject | Do Not Reject | Reject	Reject | Do Not Reject	Do Not Reject | Reject
T_{1L} T_{1U}	T_{1L}	T_{1U}

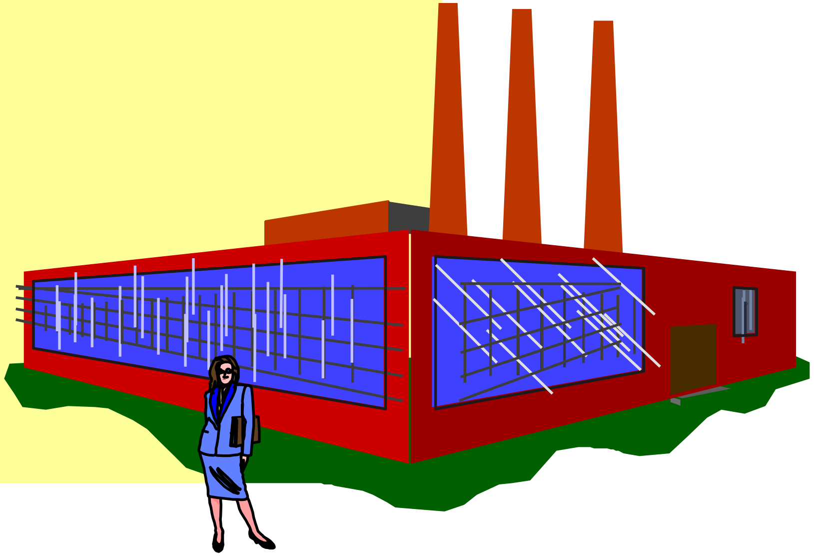
M_1 = median of population 1

M_2 = median of population 2

Wilcoxon Rank Sum Test: Example

Assume you're a production planner. You want to see if the median operating rates for the 2 factories is the same. For factory 1, the rates (% of capacity) are 71, 82, 77, 92, 88. For factory 2, the rates are 85, 82, 94 & 97.

Do the factories have the same median rates at the 0.10 significance level?



Wilcoxon Rank Sum Test: Computation Table

Factory 1		Factory 2	
Rate	**Rank**	**Rate**	**Rank**
71	1	85	5
82	Tie ~~3~~ 3.5	82	Tie ~~4~~ 3.5
77	2	94	8
92	7	97	9
88	6	...	...
Rank Sum	T_2=**19.5**		T_1=**25.5**

Lower and Upper Critical Values T_1 of Wilcoxon Rank Sum Test

n_2	α		n_1	
	One-Tailed	Two-Tailed	4	5
4				
5	.05	.10	12, 28	19, 36
	.025	.05	11, 29	17, 38
	.01	.02	10, 30	16, 39
	.005	.01	--, --	15, 40
6				

Wilcoxon Rank Sum Test: Solution

- H_0: $M_1 = M_2$
- H_1: $M_1 \neq M_2$
- $\alpha = .10$
- $n_1 = 4 \quad n_2 = 5$
- **Critical Value(s):**

Reject	Do Not Reject	Reject

12 **28**

Test Statistic:

$T_1 = 5 + 3.5 + 8 + 9 = 25.5$
(Smallest Sample)

Decision:

Do not reject at $\alpha = 0.10$.

Conclusion:

There is no evidence medians are not equal.

Wilcoxon Rank Sum Test (Large Sample)

- For Large Sample, the Test Statistic T_1 is Approximately Normal with Mean μ_{T_1} and Standard Deviation σ_{T_1}
  - $\mu_{T_1} = \dfrac{n_1(n+1)}{2}$ $\qquad \sigma_{T_1} = \sqrt{\dfrac{n_1 n_2 (n+1)}{12}}$
  - $n_1 \le n_2 \qquad n = n_1 + n_2$
- Z Test Statistic
  - $Z = \dfrac{T_1 - \mu_{T_1}}{\sigma_{T_1}}$

Wilcoxon Signed Ranks Test

- Used for Testing Median Difference for Matched Items or Repeated Measurements When the *t* Test for the Mean Difference is NOT Appropriate
- Assumptions:
  - Paired observations or repeated measurements of the same item
  - Variable of interest is continuous
  - Data are measured at interval or ratio level
  - The distribution of the population of difference scores is approximately symmetric

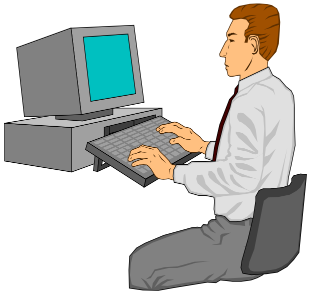
Wilcoxon Signed Ranks Test: Procedure

1. Obtain a difference score D_i between two measurements for each of the n items
2. Obtain the n absolute difference $|D_i|$
3. Drop any absolute difference score of zero to yield a set of $n' \leq n$
4. Assign ranks R_i from 1 to n' to each of the non-zero $|D_i|$; assign average rank for ties
5. Obtain the signed ranks $R_i^{(+)}$ or $R_i^{(-)}$ depending on the sign of D_i
6. Compute the test statistic: $W = \sum_{i=1}^{n'} R_i^{(+)}$
7. If $n' \leq 20$, use Table E.9 to obtain the critical value(s)
8. If $n' > 20$, W is approximated by a normal distribution with

$$\mu_W = \frac{n'(n'+1)}{4} \qquad \sigma_W = \sqrt{\frac{n'(n'+1)(2n'+1)}{24}}$$

Wilcoxon Signed Ranks Test: Example

Assume you work in the finance department. Is the new financial package faster (α=0.05 level)? You collect the following processing times:



User	Existing System (1)	New Software (2)	Difference D_i
C.B.	9.98 Seconds	9.88 Seconds	.10
T.F.	9.88	9.86	.02
M.H.	9.84	9.75	.09
R.K.	9.99	9.80	.19
M.O.	9.94	9.87	.07
D.S.	9.84	9.84	.00
S.S.	9.86	9.87	- .01
C.T.	10.12	9.98	.14
K.T.	9.90	9.83	.07
S.Z.	9.91	9.86	.05

Wilcoxon Signed Ranks Test: Computation Table

User	Existing	New	Di	|Di|	Ri	Ri(+)
C.B.	9.98	9.88	0.1	0.1	7	7
T.F.	9.88	9.86	0.02	0.02	2	2
M.H.	9.84	9.75	0.09	0.09	6	6
R.K.	9.99	9.8	0.19	0.19	9	9
M.O.	9.94	9.87	0.07	0.07	4.5	4.5
D.S.	9.84	9.84	0	0		
S.S.	9.86	9.87	-0.01	0.01	1	
C.T.	10.12	9.98	0.14	0.14	8	8
K.T.	9.9	9.83	0.07	0.07	4.5	4.5
S.Z.	9.91	9.86	0.05	0.05	3	3

$$W = \sum_{i=1}^{n'} R_i^{(+)} = 44$$

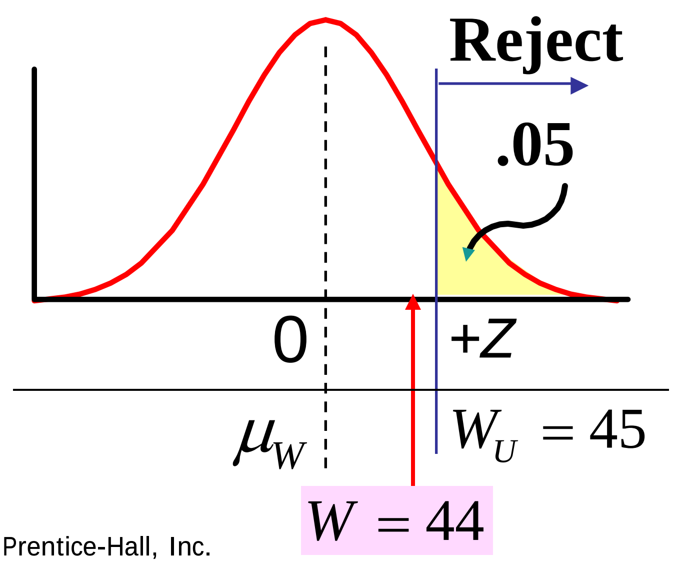
Upper Critical Value of Wilcoxon Signed Ranks Test

n	*ONE-TAIL* $\alpha = 0.05$ *TWO-TAIL* $\alpha = 0.10$	$\alpha = 0.025$ $\alpha = 0.05$
	(*Lower, Upper*)	
9	8,37	5,40
10	10,45	8,47
11	13,53	10,56

Upper critical value ($W_U = 45$)

Wilcoxon Signed Ranks Test: Solution

- H_0: $M_1 \leq M_2$
- H_1: $M_1 > M_2$
- $\alpha = .05$
- $n = 10$
- Critical Value:



Test Statistic:

$W = 44$

Decision:

Do not reject at $\alpha = 0.05$.

Conclusion:

There is no evidence the median difference is greater than 0. There is not enough evidence to conclude that the new system is faster.

Chapter Summary

- Compared Two Independent Samples
  - Performed Z test for the differences in two means
  - Performed t test for the differences in two means
- Addressed F Test for Difference in Two Variances
- Compared Two Related Samples
  - Performed paired sample Z tests for the mean difference
  - Performed paired sample t tests for the mean difference

Chapter Summary

(continued)

- Addressed Wilcoxon Rank Sum Test
  - Performed tests on differences in two medians for small samples
  - Performed tests on differences in two medians for large samples
- Illustrated Wilcoxon Signed Ranks Test
  - Performed tests for median differences for paired observations or repeated measurements