# Two-Sample t-test
# Non-Homogenous Variances

Course: Statistics 1

Lecturer: Dr. Courtney Pindling

# Example

- Case:
  - Independent Samples: different variances
  - Sample: Compare means of two samples

| **Group 1** | **Group 2** |
|---|---|
| $N = 8$ | $N = 8$ |
| Mean, $M_1 = 3$ | Mean, $M_2 = 4.5$ |
| Std Dev, $S_1 = 1.51$ | Std Dev, $S_2 = 3.21$ |
| Variance, $S_1^2 = 2.29$ | Variance, $S_2^2 = 10.29$ |
| $\sum X = 24$ | $\sum X = 36$ |
| $\sum X^2 = 88$ | $\sum X^2 = 234$ |

# *F* Critical Value

- $F_{cv}$ = 3.79 (*F*-Dist. Table, $df_1 = 7$, $df_2 = 7$, a = 0.05)

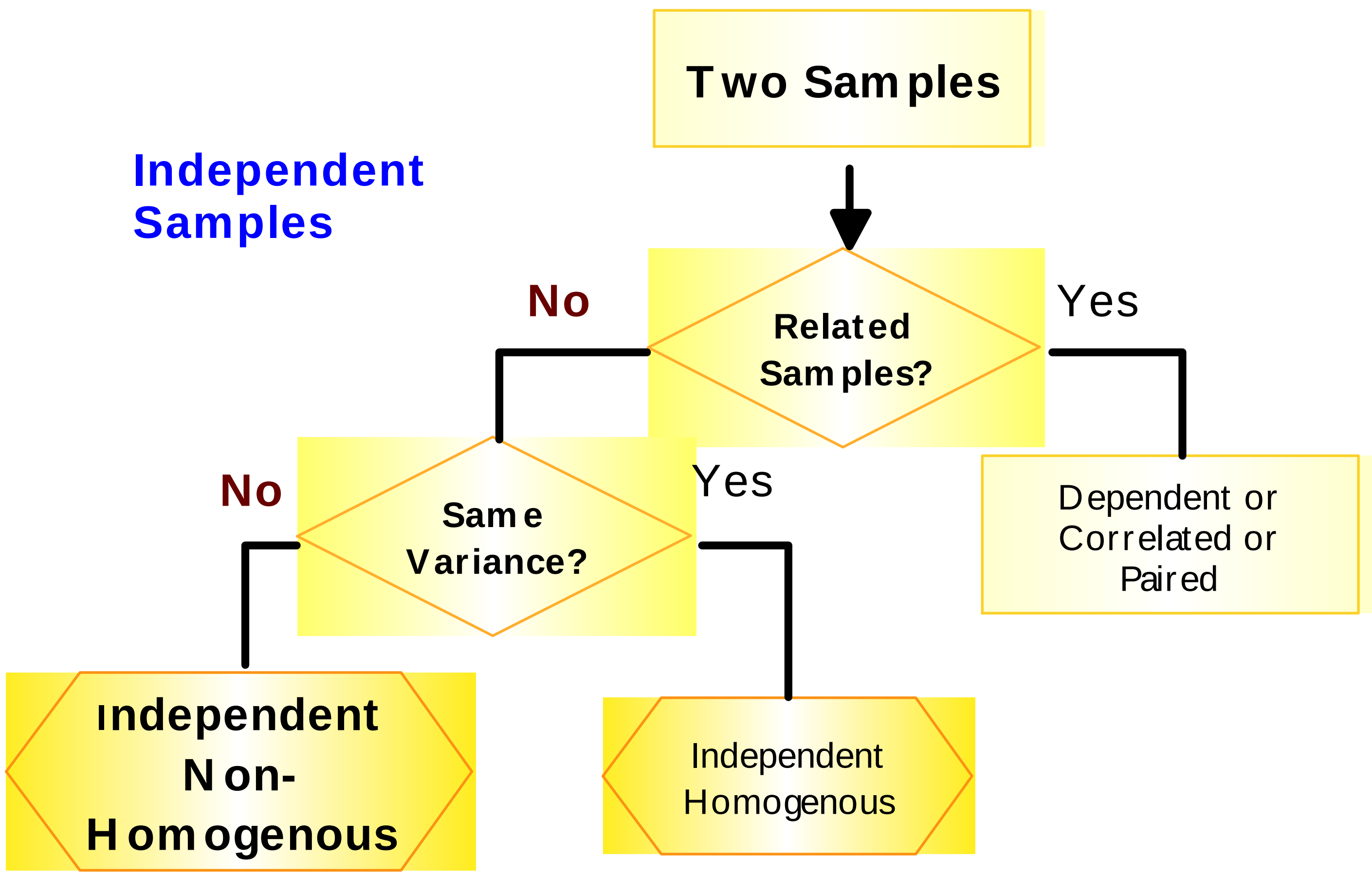
| *df 2* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | *df 2* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.4462 | 199.4995 | 215.7067 | 224.5833 | 230.1604 | 233.9875 | 236.7669 | 238.8842 | 240.5432 | 241.8819 | 1 |
| 2 | 18.51276 | 19.00003 | 19.16419 | 19.24673 | 19.29629 | 19.32949 | 19.35314 | 19.37087 | 19.38474 | 19.39588 | 2 |
| 3 | 10.12796 | 9.552082 | 9.276619 | 9.117173 | 9.013434 | 8.940674 | 8.88673 | 8.845234 | 8.812322 | 8.785491 | 3 |
| 4 | 7.70865 | 6.944276 | 6.591392 | 6.388234 | 6.256073 | 6.163134 | 6.094211 | 6.041034 | 5.9988 | 5.964353 | 4 |
| 5 | 6.607877 | 5.786148 | 5.409447 | 5.192163 | 5.050339 | 4.950294 | 4.875858 | 4.818332 | 4.77246 | 4.735057 | 5 |
| 6 | 5.987374 | 5.143249 | 4.757055 | 4.533689 | 4.387374 | 4.283862 | 4.206669 | 4.146813 | 4.099007 | 4.059956 | 6 |
| 7 | 5.59146 | 4.737416 | 4.34683 | 4.120309 | 3.971522 | 3.865978 | **3.79** | 3.725717 | 3.676675 | 3.636529 | 7 |
| 8 | 5.317645 | 4.458968 | 4.06618 | 3.837854 | 3.687504 | 3.580581 | 3.50046 | 3.438103 | 3.388124 | 3.347168 | 8 |
| 9 | 5.117357 | 4.256492 | 3.862539 | 3.63309 | 3.481659 | 3.373756 | 3.29274 | 3.229587 | 3.18 | 3.137274 | 9 |

# Test for Homogeneity of Variance

- $H_0$: $s_1^2 = s_2^2$ So, $H_a$: $s_1^2 \neq s_2^2$
- $F_{cv} = 3.79$ ($df_1 = 7$, $df_2 = 7$ and a = 0.05)
- $F_{stat} = 4.5$ (10.29/2.29 = 4.5)
- Decision: Reject $H_0$ that variances are same
  - Since $F_{stat} > F_{cv}$

**Conclusion**: Variances are **not** homogenous

# Independent Samples with Non-Homogenous Variance



# Degree of Freedom, *df*

- Underlying *t*-distribution
- Degree of freedom, *df*, defined by formula
- *df* = 10

**Given: $n_1 = 8$, $s_1^2 = 2.29$** **Given: $n_2 = 8$, $s_2^2 = 10.29$**

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left(\frac{2.29}{8} + \frac{10.29}{8}\right)^2}{\frac{\left(\frac{2.29}{8}\right)^2}{7} + \frac{\left(\frac{10.29}{8}\right)^2}{7}} = 9.96 \simeq 10$$

## Step 1: Hypotheses

- Null, $H_0$ (***no difference in means***)

  $\mu_1 - \mu_2 = 0$

- Alternative, $H_a$ (Non-Directional)

  $\mu_1 - \mu_2 \neq 0$

# Step 2: Set Rejection Criterion

- Significance Level: a = 0.05
- Critical value: *t*-distribution, *df* = 10

  (from Formula)

  - Two-tailed (non-directional)
  - $t_{cv} = 2.228$
  - Reject $H_0$ if $t_{stat} >= 2.228$

# Step 3: Compute Test Statistics

Given: $n_1 = 8, M_1 = 3, s_1 = 1.51, s_1^2 = 2.29, \Sigma X = 24, \Sigma X^2 = 88$

Given: $n_2 = 8, M_2 = 4.5, s_1 = 3.21, s_1^2 = 10.29, \Sigma X = 36, \Sigma X^2 = 234$

Std Error: $s_{(M_1 - M_2)} = \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)} = \sqrt{\left(\frac{2.29}{8} + \frac{10.29}{8}\right)} = 1.25$

Note: Standard Error Formula

Test statistics, $t = \frac{M_1 - M_2}{s_{(M_1 - M_2)}} = \frac{3 - 4.5}{1.25} = -1.2$

Absolute value of $t_{stat}$ = **1.2**

# Step 4: Confidence Interval

- CI = Statistics +/- Critical Value (Standard Error)
- Mean Difference, $M_D$ = -1.5, Since 3 – 4.5
- $t_{cv}$ = 2.228 (two-tailed, *df* = 10 and a = 0.05)

- $CI_{95}$ = -1.5 +/- 2.228(1.25) = -4.285 **to** 1.285

# Step 5: Effect Size

- ES = Mean Difference / Standard Error = $(M_1 - M_2)/s$
- Calculated $s^2$ (pooled estimate) = 6.29
- So ***s*** = Sqrt(20) = 2.50

- ES = (3 – 4.5)/2.5 = -1.5/2.5 = -0.60; **Use absolute value**, *d* = **0.60**

- **Conclusion**: Medium effect (about 0.6)

$$s^2 = \frac{[\sum X_1^2 - \frac{(\sum X_1)^2}{n_1}] + [\sum X_2^2 - \frac{(\sum X_2)^2}{n_2}]}{n_1 + n_2 - 2} = \frac{[88 - \frac{576}{8}] + [234 - \frac{1296}{8}]}{8 - 8 - 2} = 6.29$$

# Step 6: Decision

- Homogeneity of Variance assumption **<u>not</u>** met

- **<u>Do not</u>** Reject $H_0$:
  - 1. $t_{stat} < t_{cv}$ or 1.2 < *2.228*
  - *2. Hypothesized population difference of 0 is within* $CI_{95}$
    - $CI_{95}$: *-4.285 to 1.285*
  - 3. ES = 0.60 >= 0.6, is Medium effect

- **Conclusion**: The group 1 mean is <u>not</u> significantly different from group 2 mean

# SPSS Outputs

**Group Statistics**

| | Group2 | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Words | 1.00 | 8 | 3.0000 | 1.51186 | .53452 |
| | 2.00 | 8 | 4.5000 | 3.20713 | 1.13389 |

**Independent Samples Test**

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | 95% Confidence Interval of the Difference | |
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | Lower | Upper |
| Words | Equal variances assumed | 8.400 | .012 | -1.197 | 14 | .251 | -1.50000 | 1.25357 | -4.18863 | 1.18863 |
| | Equal variances not assumed | | | -1.197 | 9.965 | .259 | -1.50000 | 1.25357 | -4.29446 | 1.29446 |

***F*-Test:** Reject null hypothesis - **not** same variance, since $F_{sig}$ = **0.012** < **0.05**

***t*-Test:** Don't reject null – means are same, since:

**1.** $t_{test}$ = **1.197** < $t_{cv}$ = **2.228** (two-tailed, *df* = 10, a = 0.05)

**2.** *p*-value = **0.259** > **0.05** and **3.** $CI_{95}$ contains **0**