V. Two-tail testing of two sample means from independent populations

A. Populations are independent when a sample selected from one is not related to a sample selected from the other.

B. Examples of independent populations include production time using two different assembly procedures and industrial accidents at two plants.

C. These tests assume the populations are approximately normal with equal variances.

1. These equal variances make a weighted (pooled) point estimate the best estimate of the population σ^2.
2. $S_w^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ $\quad$ S_1^2 is the variance of sample #1 and S_2^2 is the variance of sample #2.

D. Linda Smith wants to compare the time salespeople spend with customers at two of her stores. A sample of 6 salespeople from one store had a mean of 4.5 minutes and variance of 3. A sample of 5 from a second store had a mean of 5.1 minutes and a variance of 3.1. Linda will conduct a .05 level test to determine whether the means are the same for these normally distributed populations.

E. The 5-step approach to hypothesis testing

1. These are the null hypothesis and alternate hypothesis.
   a. $H_0 : \mu_1 = \mu_2$
   b. $H_1 : \mu_1 \neq \mu_2$
2. The level of significance is .05 for this two-tail test.
3. The relevant test statistic is $\bar{x}$. $\quad t = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{s_w^2(\frac{1}{n_1}+\frac{1}{n_2})}}$

$\bar{x}_1$ is 4.5, s_1^2 is 3.0, $\bar{x}_2$ is 5.1, and s_2^2 is 3.1.
n_1 is 6 and n_2 is 5.
s_w^2 is the weighted or pooled estimate of the population variance.
df = items tested - number of samples

4. Reject the null hypothesis when the test statistic is beyond the critical value.
5. Apply the decision rule.

$$s_w^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2} = \frac{(6-1)3.0+(5-1)3.1}{6+5-2} = 3.0$$

$$t = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{s_w^2(\frac{1}{n_1}+\frac{1}{n_2})}} = \frac{4.5-5.1}{\sqrt{3.0(\frac{1}{6}+\frac{1}{5})}} = -.57$$

$df = n_1 + n_2 - 2 = 6 + 5 - 2 = 9 \rightarrow t = \pm 2.262$

Accept H_0 because -.57 is not beyond -2.262. Average customer waiting time at these stores is the same.

VI. Two-tail testing of two sample means from dependent populations using a "paired difference test"

A. Paired sample are used to test a change in environment. Examples include production time before and after training and accidents before and after a safety campaign. A large difference means variables are dependent.

B. Weekly sales at three of Linda's stores, before and after a big promotion, were $1,200, $1,300 and $1,400 and $1,400, $1,500 and $1,500 respectively. Linda will conduct a .10 level test to determine whether the promotion increased sales at these three stores. This is a one-tail test. Any change in sales would be a two-tail test.

1. Paired tests treat data sets as one sample. A large difference results in a negative measure ($\mu_d < 0$).
2. The 5-step approach to hypothesis testing
   a. The null hypothesis and alternate hypothesis are $H_0 : \mu_d \geq 0$ and $H_1 : \mu_d < 0$.
   b. The level of significance is .10.
   c. The relevant statistic is $\bar{d}$.

$\bar{d}$ is the mean difference of paired observations.
s_d is the standard deviation of paired differences.
n is the number of paired observations.
df = n - 1 = 3 - 1 = 2 → t = ±1.886

Store	Sales Dollars Before	Sales Dollars After	Difference d	d^2
1	1,200	1,400	-200	40,000
2	1,300	1,500	-200	40,000
3	1,400	1,500	-100	10,000
Totals			-500	90,000

$$\bar{d} = \frac{\sum d}{n} = \frac{-500}{3} = -\$166.67$$

$$s_d = \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1}} = \sqrt{\frac{90{,}000 - \frac{-500^2}{3}}{3-1}} = 57.7$$

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} = \frac{-166.67}{\frac{57.7}{\sqrt{3}}} = -5.03$$

Reject H_0 because -5.03 is beyond -1.886. Sales increased.

Note: With independent populations, we test the mathematical relationship between different environments. With dependent populations, we test to see if a change in environment affects population parameters.