Two Sample Hypothesis Tests for Means - Independent Groups

Data Science for Quality Management: Two Sample Hypothesis Testing

with **Wendy Martin**

Learning objective:

Perform a statistical test for differences in means (independent groups) when variances are equal and unequal

Two Independent Sample Tests for Means

- Unknown σ_1^2 and σ_2^2 (t test)
  - σ_1^2 and σ_2^2 presumed equal
  - σ_1^2 and σ_2^2 presumed unequal

Two Sample Equal Variance t Test for Means – Assumptions

- The samples are randomly selected from two independent populations or processes
- The underlying processes are normally distributed

Two Sample Equal Variance t Test for Means – Assumptions

- Homogeneity of variance is assumed
- $\sigma_1^2 = \sigma_2^2 = \sigma^2$
- Population variances are unknown

Two Sample Equal Variance t Test for Means

- Hypotheses

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Two Sample Equal Variance t Test for Means

- Test Statistic

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}}}$$

- Has df = $n_1 + n_2 - 2$, where

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

Example 1 - t Test Problem

- A design engineer finds that for two different designs of the same motor, the brush box wear when tested appeared as follows:

Example 1 - t Test Problem

$\bar{X}_1 = 0.0060$ $\bar{X}_2 = 0.0090$

$s_1 = 0.0015$ $s_2 = 0.0013$

$n_1 = 25$ $n_2 = 30$

Example 1 - t Test Problem

- Based upon the two random samples, is it reasonable to assume that the average amount of wear is equal for the two design populations? Assume α = 0.05.

Example 1- t Test Problem

- In RStudio

```
> t.test.twosample.independent
> t.test.twosample.independent.simple
```

Example 2 - t Test Problem

- The characteristic “cap pull force” refers to the amount of effort required (measured in pounds) to remove a lipstick cap from an assembly base

- A cap that pulls off too easily results in an assembly that may fall apart

Example 2 - t Test Problem

- A cap that is difficult to remove would also tend to frustrate the end user
- The assembly components are made from injection molded plastic
- There are two molds that make the cap

Example 2 - t Test Problem

- The plant manager wanted to know if the average cap pull force was equal for caps produced on the two molds

- Two groups of random samples were drawn, one from each of the two production lines, with each batch representing a different cap mold

Example 2 - t Test Problem

- Appropriate procedures were followed in measuring cap pull force for the two batches
- The resultant data are recorded below and are stored in the data file **CapPull2.dat**
- Assume $\alpha = 0.05$

Two Sample Unequal Variance t Test for Means – Assumptions

- The samples are randomly selected from two independent populations or processes.
- The underlying processes are normally distributed.
- The population or process variances are not equal.

Two Sample Unequal Variance t Test for Means

- Hypotheses

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Two Sample Unequal Variance t Test for Means

- Test Statistic

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- Has df = df*

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

Example 3 - t Test Problem

- A production manager wants to compare two ultrasonic welders for average resistance to destruction for the parts made on each.

- History has shown that very different amounts of variation can occur between the two machines, and the parts are very expensive to test.

Example 3 - t Test Problem

- The randomly selected samples show

$\bar{X}_1$	= 75	$\bar{X}_2$	= 82
s_1	= 20	s_2	= 9
n_1	= 12	n_2	= 12

- Test an appropriate hypothesis at an $\alpha = 0.10$.

Sources

- Luftig, J. An Introduction to Statistical Process Control & Capability. Luftig & Associates, Inc. Farmington Hills, MI, 1982