Two Sample Hypothesis Tests for Means - Dependent Groups

Data Science for Quality Management: Two Sample Hypothesis Testing

with **Wendy Martin**

Learning objective:

Perform a statistical test for differences in means (dependent groups) with repeated measures and with matched pairs

Repeated Measures t Test for Means

- Used to compare means of repeated measures or paired groups
- Tests the following hypotheses:

$$H_0: \mu_1 = \mu_2 \qquad H_1: \mu_1 \neq \mu_2$$

$$\text{or} \qquad\qquad \text{or}$$

$$H_0: \mu_D = 0 \qquad H_1: \mu_D \neq 0$$

Repeated Measures t Test for Means – Assumptions

- The n pairs of scores are independent of one another
- The population for difference scores is normally distributed, as are the populations for each group

Repeated Measures t Test for Means

- Test Statistic $$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n} - 2r\frac{s_1}{\sqrt{n}}\frac{s_1}{\sqrt{n}}}} = \frac{\bar{D}}{s_d/\sqrt{n}}$$

Where n is the sample size for the pairs of scores

Example 1 – Repeated Measures t Test Problem

- A Black Belt is attempting to solve a noise level problem with a type of wiper motor
- They suspect that a major source of noise problems may be traced to the material from which the bearing is made

Example 1 – Repeated Measures t Test Problem

- To determine whether the average noise level of the motors can be decreased by changing to a new bearing material, they randomly select ten motors from the current production line

Example 1 – Repeated Measures t Test Problem

- A noise level reading is taken on each of the ten motors.
- Next, the bearings are replaced with bearings made from the new material and the motors are retested

Example 1 – Repeated Measures t Test Problem

- The noise level for the each of the modified motors is recorded as was previously done
- All testing is conducted under the same essential conditions as for the first set of measurements. The data were collected and entered in the file **Noise.dat**

Example 1 – Repeated Measures t Test Problem

- Test an appropriate hypothesis to determine whether it is reasonable to assume that the new bearing material will effectively reduce the initial noise level of the motors if implemented across the process. Assume a significance level of 0.05.

Example 1 – Repeated Measures t Test Problem

- In RStudio

```
> t.test.twosample.dependent
> t.test.twosample.dependent.simple.dbar
```

Matched Pairs t Test for Means - Assumptions

- The specimens in the two samples are independent (by nature)
- The population for difference scores is normally distributed, as are the populations for each group

Matched Pairs t Test for Means - Assumptions

- Homogeneity of variance is assumed (Not critical if sample sizes are equal)
- The units or specimens in the two samples are dependent by design.

Matched Pairs t Test for Means

- Test Statistic

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n} - 2r\frac{s_1}{\sqrt{n}}\frac{s_1}{\sqrt{n}}}}$$

Where n is the sample size for the pairs of scores

Example 2 – Matched Pairs t Test Problem

- A production manager wishes to determine whether two secondary blanking presses are producing raw plates with equal average flatness, a smaller number is better
- It is known that hardness of the material is a variable that can affect the flatness of the plates

Example 2 – Matched Pairs t Test Problem

- Therefore, 60 plates are randomly selected from the primary blanking operation and tested for hardness

Example 2 – Matched Pairs t Test Problem

- After rank ordering and pairing the plates according to their tested hardness (i.e., lowest with the next lowest, and so on up the measured hardness scale), they randomly assign the plates from each pair to one of the two groups (i.e., one plate from each pair to the group going to press 1 and one plate from each pair to the group going to press 2)

Example 2 – Matched Pairs t Test Problem

- After the assignments are complete, the manager verifies that there are no significant differences between the two groups in terms of plate hardness on the basis of their variances and means. Why did they do this?

Example 2 – Matched Pairs t Test Problem

- After the assignments are complete, the manager verifies that there are no significant differences between the two groups in terms of plate hardness based on their variances and means.
- Then, the two groups are run concurrently as pairs through the presses, and the resultant flatness data are recorded.

Example 2 – Matched Pairs t Test Problem

- The summary statistics for the two groups of data are as follows.

$\bar{X}_1$	$= 35.24$	$\bar{X}_2$	$= 38.02$
s_1	$= 5.18$	s_2	$= 5.63$
r_{12}	$= 0.60$		

- Test an appropriate hypothesis at an $\alpha = 0.05$

Example 2 – Matched Pairs t Test Problem

- In RStudio

```
> cor.pearson.r.onesample.simple
> t.test.twosample.dependent.simple.meandiff
```

Sources

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