11.3: Dependent Sampling – Matched Pairs t-test

The independent models shown above compared samples that were not related. However, it is often advantageous to have related samples that are paired up – two measurements from a single population. The model we will consider here is called the **matched pairs t-test** also known as the paired difference t-test. The advantage of this design is that we can eliminate variability because other factors are not being studied, increasing the power of the design.

In this model we take the difference of each pair and create a new population of differences, so if effect, the hypothesis test is a one population test of mean that we already covered in the prior section.

Matched pairs t-test to compare the means for two dependent populations

Model Assumptions

- Dependent Sampling
- $X_d = X_1 - X_2$
- $\overline{X}_d = \overline{X}_1 - \overline{X}_2$ approximately Normal

Test Statistic

- $t = \dfrac{\overline{X}_d - \mu_d}{s_d/\sqrt{n}}$
- $df = n - 1$

Example: Rental cars

An independent testing agency is comparing the daily rental cost for renting a compact car from Hertz and Avis.

A random sample of 15 cities is obtained and the following rental information obtained.

City	Hertz	Avis
Atlanta	42	40
Baltimore	51	47
Boston	46	42
Chicago	56	52
Cleveland	45	43
Denver	48	48
Dallas	56	54
Honolulu	37	32
Los Angeles	51	48
Kansas City	45	48
Miami	41	39
New York	44	42
San Francisco	48	45
Seattle	46	50
Washington DC	44	43

At the .05 significance level can the testing agency conclude that there is a difference in the rental charged?

Notice in this example that cities are the single population being sampled and that two measurements (Hertz and Avis) are being taken from each city. Using the matched pair design, we can eliminate the variability due to cities being differently priced (Honolulu is cheap because you can't drive very far on Oahu!)

Solution

Design

Research Hypotheses:

$H_o: \mu_1 = \mu_2$ (Hertz and Avis have the same mean price for compact cars.)

$H_a: \mu_1 \neq \mu_2$ (Hertz and Avis do not have the same mean price for compact cars.)

Model will be matched pair t-test and these hypotheses can be restated as: $Ho: \mu_d = 0 \quad Ha: \mu_d \neq 0$

The test will be run at a level of significance (α) of 5%.

Model is two-tailed matched pairs t-test with 14 degrees of freedom. Reject H_o if $t <$-2.145 or $t > 2.145$

Data/Results

City	Hertz	Avis	Difference
Atlanta	42	40	**2**
Baltimore	51	47	**4**
Boston	46	42	**4**
Chicago	56	52	**4**
Cleveland	45	43	**2**
Denver	48	48	**0**
Dallas	56	54	**2**
Honolulu	37	32	**5**
Los Angeles	51	48	**3**
Kansas City	45	48	**-3**
Miami	41	39	**2**
New York	44	42	**2**
San Francisco	48	45	**3**
Seattle	46	50	**-4**
Washington DC	44	43	**1**

We take the difference for each pair and find the sample mean and standard deviation.

$$\overline{X}_d = 1.80$$
$$s_d = 2.513$$
$$n = 15$$
$$t = \frac{1.80 - 0}{2.513/\sqrt{15}} = 2.77$$

Hypothesis Test: Paired Observations

0.000	hypothesized value
46.667	mean Hertz
44.867	mean Avis
1.800	mean difference (Hertz - Avis)
2.513	std. dev.
0.649	std. error
15	n
14	df
2.77	t
.0149	p-value (two-tailed)

Reject H_o under either the critical value or p-value method.

Conclusion

There is a difference in mean price for compact cars between Hertz and Avis. Avis has lower mean prices.

The advantage of the matched pair design is clear in this example. The sample standard deviation for the Hertz prices is $5.23 and for Avis it is $5.62. Much of this variability is due to the cities, and the matched pairs design dramatically reduces the standard deviation to $2.51, meaning the matched pairs t-test has significantly more power in this example.