Chapter 10 (continued)

Hypothesis Test for Difference in Two Population Means using Matched Pairs Samples / Dependent Samples:

Consider the following example:

|  |  |
| --- | --- |
| Population 1 | Population 2 |
| Salaries of al males in Bangladesh | Salaries of all females in Bangladesh |
| Mean salary: $μ\_{1}$ (unknown) | Mean salary: $μ\_{2}$ (unknown) |

We form the hypothesis that the mean salary of males ($μ\_{1}$) is greater than mean salary of females ($μ\_{2}$). To test this hypothesis we draw random samples from each of the two populations.



|  |  |
| --- | --- |
| Sample 1: sample of male salaries | Sample 2: sample of female salaries |
| Size: $n\_{1}$ | Size: $n\_{2}$ |
| Mean: $\overbar{x}\_{1}$ | Mean: $\overbar{x}\_{2}$ |
| Standard deviation: $s\_{1}$ | Standard Deviation: $s\_{2}$ |

In the first part of this chapter, we discussed this same example using Independent Samples. In this section of the chapter, we will not discuss the example using Matched Pairs Samples / Dependent Samples.

Formally, the definition of a Matched Pairs Sample is: Two samples are said to be ***paired*** or ***matched samples*** when for each data value collected from one sample there is a corresponding data value collected from the second sample, and both these data values are collected from the same source.

This means that for every male individual sampled from Population 1, there should be a female individual sampled from Population 2 who has similar relevant characteristics as the male individual. e.g. suppose a male of age 27 with 4 years of job experience is sampled from Population 1. This means we also have to sample a female of age 27 with 4 years of job experience from Population 2.

Another situation where matched paired samples are used is when we are studying the effects before and after a certain intervention. Example: Suppose we want to estimate the difference between the mean weights of all participants before and after a weight loss program. To accomplish this, suppose we take a sample of 40 participants and measure their weights before and after the completion of this program. Note that these two samples include the same 40 participants. This is an example of two ***dependent*** samples.

Terms and symbols:

* $D$ : difference between paired population observations
* $μ\_{D}$ : mean of population differences
* $σ\_{D}$ : standard deviation of population differences
* $d$ : difference between paired sample observations
* $\overbar{d}$ : mean of sample differences
* $s\_{d}$ : standard deviation of sample differences

To demonstrate the way to run matched pairs hypothesis test, refer to the example below:

A researcher wanted to find the effect of a special diet on systolic blood pressure. She selected a sample of seven adults and put them on this dietary plan for 3 months. The following table gives the systolic blood pressures (in mm Hg) of these seven adults before and after the completion of this plan.



Let $μ\_{D}$ be the difference in systolic blood pressure (before – after) due to this special dietary plan. Using the 1% significance level, can you conclude that the mean difference in blood pressure is positive as a result of this dietary course?

Step 1: State the hypothesis: The difference in systolic blood pressure is positive as a result of the special dietary course.

Step 2: $ H\_{0}: μ\_{D}=0 vs. H\_{1}:μ\_{D}>0$

Step 3: Check that population of paired differences is normally distributed or than sample size is greater than equal to 30. The we can use t-distribution to run this test.

Step 4: Calculate the differences in sample observations (before – after) ; also calculate the square of differences.



Calculate sample mean and sample standard deviation of differences. Compute the test statistic:

 where $n$: number of pairs. $d.f. =n-1$

Step 5: Apply rejection region method just as before.

Interval Estimate:

Construct a 95% confidence interval to estimate $μ\_{D}$.

Use the formula: where t is the critical value than has an area of $^{α}/\_{2}$ on its right.

Hypothesis Testing for Differences in Two Population Proportions

Example: A sample study of 1000 individuals revealed that unemployment rate in Country A is 4%, while the same sample study showed that unemployment rate in Country B is 4.25% among a sample size of 850 people. Test to see if the population proportions of unemployment is different in the two countries.

$$H\_{0}: p\_{1}-p\_{2}=0$$

$$H\_{1}: p\_{1}-p\_{2}\ne 0$$

Our sample proportions are $\hat{p}\_{1}=0.04 $ and $\hat{p}\_{2}=0.0425$. Recall that $\hat{q}\_{1}=1-0.04 $ and $\hat{q}\_{2}=1-0.0425$.

In order to use normal distribution to run this hypothesis test, ensure the sample size requirement given below is met:

$$n\_{1}\hat{p}\_{1}>5$$

$$n\_{1}\hat{q}\_{1}>5$$

$$n\_{2}\hat{p}\_{2}>5$$

$$n\_{2}\hat{q}\_{2}>5$$

The z standardized test statistic formula is given below:

where

 and

Once the z test statistic is computed, run a basic rejection region approach on the normal distribution.