**Hypothesis Testing**

**Tutorial Questions**

Data for these questions is provided in the Excel File **Hypothesis Testing Tutorial Questions (2021-01).xlsx**

**Question 1**

Water quality experts working for the forestry service are concerned about the water quality in a nearby lake. In particular they would like to test whether the average phosphorous content has fallen below 15 parts per billion (ppb).

A random sample taken at different points in the water supply was collected over 12 days with the following readings:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 12.3 | 17.7 | 6.6 | 16.4 | 8.2 | 18.0 | 15.5 | 9.1 | 17.4 | 10.2 | 16.3 | 8.3 |

Conduct the test on phosphorous content assuming the level of significance is 0.05 assuming phosphorous content is normally distributed.

**Hypotheses**

This is a one tail, lower tail, since we want to test if phosphorous has fallen, on the MEAN phosphorous level.

H0: µ ≥ 15

H1: µ < 15

**Decision Rule**

Reject Ho if p – value < α

α = 0.05 given in the question

**Calculating the p – value**

Probability distribution is the t-distribution as the Standard deviation is not known and must be calculated from the sample.

Validity check – t distribution is appropriate because the distribution of phosphorous is normal as given in the question.

From the data we can use Data Analysis in Excel to find:

|  |  |
| --- | --- |
| *Phosphorous Content (ppb)* | |
|  |  |
| Mean | 13 |
| Standard Error | 1.245537 |
| Median | 13.9 |
| Mode | #N/A |
| Standard Deviation | 4.314668 |
| Sample Variance | 18.61636 |
| Kurtosis | -1.86516 |
| Skewness | -0.2082 |
| Range | 11.4 |
| Minimum | 6.6 |
| Maximum | 18 |
| Sum | 156 |
| Count | 12 |

As a lower one tail test of means:

p-value = P(

Using t distribution in excel

p-value = P( = t.dist(tcalc, v, cumulative)

v = n – 1 = 11

p-value = t.dist(-1.606, 11, true) = 0.0683

**Make Decision:**

p-value = 0.0683 is not less than α = 0.05 and hence we can not reject the null hypothesis. There is not sufficient evidence to conclude that the average phosphorous level has fallen below 15ppb.

Alternative Approach – My Preference

This test can be done using the Data / Data Analysis / t-Test: Two Sample Assuming Unequal Variances function

**Hypotheses**

This is a one tail, lower tail, since we want to test if phosphorous has fallen, on the MEAN phosphorous level.

H0: µ ≥ 15

H1: µ < 15

**Decision Rule**

Reject Ho if p – value < α (0.05)

**Calculating the p – value**

Set up the data so that there is a second “Dummy” column with two zeros

|  |  |
| --- | --- |
| Phosphorous Content (ppb) | Dummy |
| 12.3 | 0 |
| 17.7 | 0 |
| 6.6 |  |
| 16.4 |  |
| 8.2 |  |
| 18 |  |
| 15.5 |  |
| 9.1 |  |
| 17.4 |  |
| 10.2 |  |
| 16.3 |  |
| 8.3 |  |

Select the Data Analysis / t-Test: Two-sample assuming unequal variances function

Graphical user interface, text, application

Description automatically generated

Enter the original data as Array 1 and the Dummy variable as Array 2

The hypothesised mean difference is our hypothesised value of µ = 15

Select α

Graphical user interface, application

Description automatically generated

You will get the resulting output



As this is a one tail test we get the p-value = 0.0683

Our p-value of 0.0683 is not less than α and hence, as above, we do not reject the null hypothesis and we do not have sufficient evidence to conclude that the phosphorous level has fallen below 15ppb.

Question 2

The University administration is concerned about the proportion of full time students who also appear to be working in a full time job. Five years ago, an extensive census of students revealed that 3 percent of full time students worked in full time jobs. A recent survey of 1000 full time students revealed that 45 of them worked full time. Test, using a 5% significance level, whether the proportion of full time students in full time work has increased over the past five years. Do the test using each of the test statistic and p-value methods.

Hypotheses

Here we wish to test if the PROPORTION has increased above 3% and hence it is a one tail, upper tail test on p

H0: p ≤ 0.03

H1: p > 0.03

**Decision Rule**

Reject Ho if p – value < α

α = 0.05 given in the question

Calculating p - value

As a test on proportions we use the z distribution provided we have a large enough sample size to ensure the sampling distribution is normal.

np = 1000(0.03) = 30, n(1-p) = 1000(0.97) = 970

Both are > 10, and hence it is valid to use the normal distribution

P – value = P() = 1 – normdist(

P – value = P() = 1 – normdist( = 0.00271

Make Decision

Since p – value = 0.00271 is less than α = 0.05 we can reject the null hypothesis. There is sufficient evidence to conclude that the proportion of students doing full time work while studying full time has increased.

Question 3

The p-value represents the probability that the null hypothesis is false. True or False? Why?

False. The p-value is the chance of getting a value of our sample statistic that is at least as extreme as the value that is observed. As such it is the probability of getting that sample result if the null hypothesis were true. Turning this around we can say this is the probability the null hypothesis is true given this is what we found in the sample. Although not technically exact, the p-value is generally taken to be the probability the null hypothesis is true given our sample data.

**Question 4**

A drink vending machine is set to dispense 200mls of drink into each plastic cup with a standard deviation of 5ml. The operator checks a random sample of 10 cups with the following results:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 205 | 198 | 197 | 201 | 207 | 199 | 202 | 198 | 204 | 203 |

1. Conduct a test at the 5% level of significance to determine if the process is dispensing the correct amount of fluid.
2. What assumption must you make in order to carry out this test?

Hypotheses

This is a two tail test of means. Since we want to see if the process is filling correctly it could be broken if it underfills or overfills implying two tails.

H0: µ = 200

H1: µ ≠ 200

**Decision Rule**

Reject Ho if p – value < α

α = 0.05 given in the question

Calculating p – value

Assuming the population of drink dispensed is normally distributed we will use a z distribution for this problem as a value for the population standard deviation, σ, is given – known from population

Test is a two tail test

From the data:

which is greater than µ = 200

P – value = 2 x (1- norm.dist(, true))

σ = 5 which is the process (known standard deviation)

P – value = 2 x (1 – norm.dist(201.4, 200, 1.581, true) = 2 x (1 – 0.8121) = 0.3758

Make Decision

Because p – value 0 0.3758 is greater than α = 0.05 we do not sufficient evidence to reject the null hypothesis. As such we continue to work on the belief that the process is filling cups correctly.

Question 5

In 2016 the Reserve Bank of Australia reported that 43% of individuals still used cash on a regular basis to make payments. The 2019 Survey of Consumer Payments showed that in a sample of 1500 respondents, 480 indicated they use cash on a regular basis. Does this data indicate that there has been a decrease in the proportion of consumers using cash regularly at a 1% level of significance?

Hypotheses

One tail lower tail test on proportion

H0: p ≥ 0.43

H1: p < 0.43

Decision Rule

Reject H0 if p – value < α (0.01)

Calculating the p - value

Testing we can use normal:

np = 1500(0.43) = 645, n(1-p) = 1500(0.57) = 855

Both are > 10 and hence we can use the normal

P – value = norm.dist(, true)

P – value = norm.dist(0.32, 0.43, 0.01278, true) = 3.8E-18 = 0

Make Decision

A p- value of 0 implies there is 0% chance the null hypothesis is true and so we can reject the null hypothesis and conclude there is sufficient evidence that the proportion of customers using cash on a regular basis has decreased.

**Question 6**

A manufacturer of steel wire wishes to test whether or not the wire being produced has an average diameter of 4mm. From past experience it is known that the standard deviation of the process is 0.04mm. A random sample of 25 lengths of wire produced an average diameter of 4.02mm.

Assuming the population distribution of wire diameters is normal, conduct a test using each of the test statistic and p-value methods to see if the process is not working according to specifications. Use a 2% level of significance.

Hypotheses

This is a test to see if the MEAN diameter is not equal to 4mm, hence it is a two tail test on µ

H0: µ = 4

H1: µ ≠ 4

Decision Rule

Reject H0 if p – value < α (0.02)

Calculate p -value

Given that the population standard deviation is known, σ = 0.04 we can use the standard normal, z, because we are told the population distribution for the diameter is normal

From the information in the question we are told =4.02, σ = 0.04 and n = 25

Since =4.02 is greater than µ = 4 the p-value is given by:

p-value = 2 x (1- norm.dist(, true))

p-value = 2 x (1- norm.dist(4.02, 4, 0.008, true)) = 2(1 – 0.9938) = 0.0124

Make Decision

Since p – value = 0.0124 is less than α = 0.02 we can reject the null hypothesis and conclude that the process is not working according to specification and the average diameter of wire is not equal to 4.0mm

**Question 7**

### Following the research report assessment a statistics lecturer is concerned that the class this year is performing more poorly than his previous class. The average score on the report last year was 17 marks. A random sample of students from this class had the following results:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 29 | 6 | 15 | 19 | 25 | 12 | 16 | 10 | 15 |

### State the null and alternative hypothesis you would use to test whether or not the students were performing more poorly this year.

### What is the probability distribution that should be used to conduct this test. Why? What assumptions will you need to make.

### Conduct the hypothesis test. What conclusion do you come to?

### What is the p – value for this particular test?

H0: µ ≥ 17

H1: µ < 17

Standard deviation is not known and so we have to calculate it from the sample data provided. As such we need to use the t distribution. This can be done only if we can assume that the population of student results follows a normal distribution.

Decision Rule:

Reject H0 if p – value < α

I will assume α = 0.05 (common value)

Calculating p – value

Because we are using the t distribution

p-value = P(tv < tcalc) = t.dist(tcalc , d.o.f(v), true)

From the data we can use Data Analysis in Excel to find:

|  |  |
| --- | --- |
| *Mark* | |
|  |
| Mean | 16.33333 |  |
| Standard Error | 2.392117 |  |
| Median | 15 |  |
| Mode | 15 |  |
| Standard Deviation | 7.17635 |  |
| Sample Variance | 51.5 |  |
| Kurtosis | -0.04656 |  |
| Skewness | 0.547336 |  |
| Range | 23 |  |
| Minimum | 6 |  |
| Maximum | 29 |  |
| Sum | 147 |  |
| Count | 9 |  |

p-value = P(tv < tcalc) = t.dist(tcalc , d.o.f(v), true) = t.dist(-0.28, 9, true) = 0.393

Graphical user interface, text, application, email

Description automatically generated

p-value = 0.393 is not less than α = 0.05 and hence we do not reject Ho. There is not sufficient evidence to conclude that the average results of students has decreased this year.

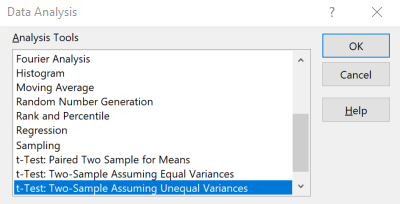
Alternative Dummy Variable Approach

Note this test can also be done using the Data Analysis function t-test: Two-sample using unequal variances

Set up the data so that there is a second “Dummy” column with two zeros

|  |  |
| --- | --- |
| Mark | Dummy |
| 29 | 0 |
| 6 | 0 |
| 15 |  |
| 19 |  |
| 25 |  |
| 12 |  |
| 16 |  |
| 10 |  |
| 15 |  |

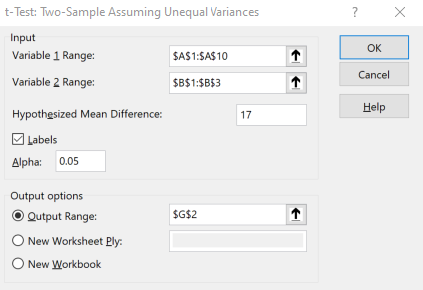
Select the Data Analysis / t-Test: Two-sample assuming unequal variances function



Enter the original data as Array 1 and the Dummy variable as Array 2

The hypothesised mean difference is our hypothesised value of µ = 17

Select α



You will get the resulting output



As this is a one tail test we get the p-value = 0.393 as before and also the tcalc given here as t Stat = -0.2787 the same as we calculated above

**Question 8**

The manager of Bowl-a-Rama wants to test to see if a new promotion using vouchers where a customer gets 3 games for the price of two has increased the number of customers. Prior to the promotion the business was averaging 450 customers per day. In the 14 days following the promotion the following customer numbers were recorded:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 476 | 521 | 654 | 635 | 245 | 389 | 454 |
| 488 | 559 | 641 | 659 | 268 | 377 | 448 |

Test to see if the new promotion has been successful using a 5% level of significance. What assumption is necessary to conduct this test?

Hypotheses

A successful campaign would increase the average number of customers, hence it is a one tail upper tail test

H0: µ ≤ 450

H1: µ > 450

Decision Rule

Reject H0 if p – value < α (0.05)

Calculating p – value

Standard deviation is not known, and hence needs to be calculated using sample data - hence we use t distribution.

Simplest approach is to use dummy variable method and t-test: Two Sample Assuming Unequal Variances

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances |  |  |
|  |  |  |
|  | *Customers* | *Dummy* |
| Mean | 486.714 | 0 |
| Variance | 18491.758 | 0 |
| Observations | 14 | 2 |
| Hypothesized Mean Difference | 450 |  |
| df | 13 |  |
| t Stat | 1.0102 |  |
| P(T<=t) one-tail | 0.1654 |  |
| t Critical one-tail | 1.7709 |  |
| P(T<=t) two-tail | 0.3308 |  |
| t Critical two-tail | 2.1604 |  |

p-value = 0.1654

Make Decision

p-value = 0.1654 is greater than α = 005 and hence we do not reject Ho. There is not sufficient evidence to conclude that the average number of customers has increased.

Question 9

Producers of Shaker chocolate milk have in the past controlled 35% of the flavoured milk market, but they are hoping that a new advertising campaign will boost this share. If a sample of 200 customers indicates that 82 of them drink Shaker milk would you recommend that the company keep going with the campaign? Use a 2% level of significance. Calculate the p-value. Would you recommend they keep going at the 5% level of significance?

Hypotheses

We are talking about the percentage of people – hence it is a test of proportions. To keep going with a campaign we would want to know if the proportion has increased – hence it is an upper one tail test

H0: p ≤ 0.35

H1: p > 0.35

Decision Rule

Reject H0 if p – value < α

Calculating the p - value

Testing we can use normal:

np = 200(0.35) = 70, n(1-p) = 200(0.65) = 130

Both are > 10 and hence we can use the normal

P – value = 1 - norm.dist(, true)

P – value = 1 - norm.dist(0.41, 0.35, 0.03373, true) = 1 – 0.96238 = 0.03762

Make Decision

A p- value of 003762 implies there is 3.76% chance the null hypothesis is true.

Balancing this probability against the levels of significance given in the question we could reject the null hypothesis at a 5% level, but not at a 2% level of significance.

In the end it will come down to whether or not the company is happy that there was only a 3.76% chance of error to continue to spend on this campaign.

**Question 10**

Previous studies have indicated travelling times from Wynnum to the city during peak hour are normally distributed with a mean of 45 minutes and a standard deviation of 2 minutes. Following a number of changes to the road system, the Department of Transport wants to test to see if the travelling time to the city has been altered. A random sample of 20 motorists provided a mean travel time of 43.9 minutes.

1. State the null and alternative hypotheses to be tested.
2. Construct an appropriate decision rule and conduct the test. Let α=0.01
3. Calculate the p-value for this test.
4. Redo parts (a), (b) and (c) if the transport department is only interested in testing to see if the mean travel time has been reduced. Explain the difference between the two tests.

Hypotheses

Here we are testing the MEAN travel time in minutes and wish to see if travel time has been altered. Hence this is a two tail test of µ

H0: µ = 45

H1: µ ≠ 45

Decision Rule

Reject H0 if p – value < α (0.01)

Calculating the p-value

Standard deviation is given as a previously known one, so we can assume this is σ and hence use z distribution.

This can be done as a z test as we are also told the distribution of travel times is normal.

σ = 2

p – value for a two tail test where the value of (43.9) is less than µ is:

p-value = 2 x P( < 43.9) = 2 x norm.dist(43.9, 45, 0.447, true) = 2 x 0.00693 = 0.01386

p – value = 0.01386 is not less than α = 0.01 and hence we do not reject Ho. There is not sufficient evidence to conclude that the average travel time has changed from 45 minutes at a 1% level of significance.

If we were to do this test as a one tail test if the mean travel time has reduced, the differences are:

H0: µ ≥ 45

HA: µ < 45

Decision Rule

Reject H0 if p – value < α (0.01)

Calculating the p-value

p – value for a lower one tail test

p-value = P( < 43.9) = norm.dist(43.9, 45, 0.447, true) = 0.00693

p – value = 0.00693 is less than α = 0.01 and hence we do not reject Ho. There is not sufficient evidence to conclude that the average travel time has changed from 45 minutes at a 1% level of significance.

The purpose of this question is to show that doing a one tail test is actually a more powerful test. By that we mean it has a better chance of minimizing the chance of a Type II error and is better at recognising if there has been a shift in the true value of the parameter of interest (in this case µ)

**Question 11**

1. An inspector examines output from a machine which if properly adjusted produces nails with an average length of 5cm and a standard deviation of 0.175cm. A sample of 35 nails indicated an average of 5.0625 cm. Is there any indication at a 5% level of significance that the process is out of adjustment.

This is a test to see if the MEAN diameter is not equal to 5mm, hence it is a two tail test on µ

H0: µ = 5

H1: µ ≠ 5

Given that the population standard deviation is known, σ = 0.175 we can use the standard normal, z. In this instance since the sample size n > 30 (n=35) we can assume the sampling distribution of the sample mean is normal

Reject H0 if:

P – value > α (0.05)

For a two tail test of means from a z distribution where

P – value = 2 x (1 – norm.dist(, µ, , true))

From the data we are told =5.0625

P – value = 2 x (1 – norm.dist(, µ, , true)) = 2 x (1 – norm.dist(5.0625, 5, 0.0296, true) = 2 x (1 – 0.9826) = 0.0348

Since p – value = 0.0348 is less than α = 0.05 we can reject the null hypothesis and conclude that the process is not working according to specification and the average length of nails is not equal to 5.0cm

Question 12

A consumer magazine wanted to test the fuel consumption of two vehicles competing in the same market. Samples of each of the makes were trialled in order to determine the average fuel consumption measuring fuel consumption in Kms per litre.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model 1: | 15.4 | 17.1 | 14.9 | 15.8 | 16.4 | 16.0 | 16.9 | 15.5 |  |  |
| Model 2: | 15.7 | 18.2 | 15.4 | 16.2 | 15.8 | 17.1 | 16.9 | 16.8 | 17.2 | 16.5 |

Conduct a test using the test statistic method to see if there is a difference in performance of the two vehicles on average using a 5% level of significance

Hypotheses

This is a two population question where we are looking at differences in MEANS between the two models of car.

It is a two tail test because we are only looking to see if there is a difference, not if one is better than the other

H0: µ1 = µ2 or rather µ1 - µ2 = 0

H1: µ1 ≠ µ2 or rather µ1 - µ2 ≠ 0

Decision Rule:

Reject H0 if p – value < α (0.05)

Calculating p – value

This question requires us to calculate the sample standard deviation for both models, and hence it is a t test.

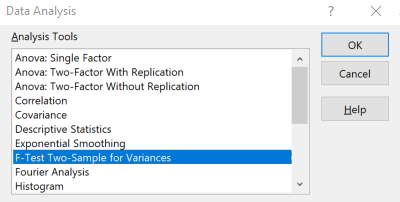
To conduct the t – test we first need to determine if the variances are equal or not. This requires an initial test of equality of variances which is an F test

H0:

H1:

Reject Ho if p-value < α (0.05)

Using the Excel Data / Data Analysis F Test Two Sample for Variances function



Graphical user interface, application

Description automatically generated

Which provides the following output:

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances | | |
|  |  |  |
|  | *Model 1:* | *Model 2:* |
| Mean | 16 | 16.58 |
| Variance | 0.577143 | 0.706222 |
| Observations | 8 | 10 |
| df | 7 | 9 |
| F | 0.817226 |  |
| P(F<=f) one-tail | 0.404023 |  |
| F Critical one-tail | 0.271985 |  |

Here the p-value for a one tail test is given as 0.404, and hence the p-value for our two tail test is:

p-value = 0.404 x 2 = 0.808

As p-value is not less than α = 0.05 we do not reject the null hypothesis and conclude that the variances are equal.

We now use excel:

T – test: Two Sample Assuming Equal Variances

From Excel we get:

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Equal Variances | | |
|  |  |  |
|  | *Model A:* | *Model B:* |
| Mean | 16 | 16.58 |
| Variance | 0.577143 | 0.706222222 |
| Observations | 8 | 10 |
| Pooled Variance | 0.64975 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 16 |  |
| t Stat | -1.51692 |  |
| P(T<=t) one-tail | 0.074397 |  |
| t Critical one-tail | 1.745884 |  |
| P(T<=t) two-tail | 0.148794 |  |
| t Critical two-tail | 2.119905 |  |

You can see from this table that the p-value for a two tail test is 0.14.

As the p-value of 0.14 is not less than the level of significance of α = 0.05. We do not reject the Null Hypothesis which in this case means we do not have sufficient evidence to say that the fuel consumption of the two vehicles is different. We conclude there is no difference.

Question 13

Countries are always accusing each other of engaging in unfair trading tactics that affect the ability of their own manufacturers or farmers to compete. This claim was tested by an Economist in the United States by comparing the sales price of the same Japanese manufactured automobile in Japan and in the US. She obtained information on a random sample of 50 US sales and 30 Japanese sales.

|  |  |  |
| --- | --- | --- |
|  | **US Sales** | **Japanese Sales (Converted to $US)** |
| **Sample Size** | 50 | 30 |
| **Sample Mean** | $24,545 | $25,243 |
| **Sample Standard Deviation** | $2,147.58 | $1,410.30 |

1. Construct a set of hypotheses that would allow you to test if Japanese car manufacturers are setting prices in the US that might give them an unfair advantage in the market.
2. Conduct the test assuming α = 0.05. Now use α = 0.1. Does the answer change? What conclusion do you draw from the data?

Hypotheses

This is a one tail test of averages. IF Japan is setting unfair prices to support manufacturers then average price in US should be less than average price in Japan.

Let 1 = US

Let 2 = Japan

H0: µ1 ≥ µ2 or rather µ1 - µ2 ≥ 0

H1: µ1 < µ2 or rather µ1 - µ2 < 0

Decision Rule

Reject H0 if p – value < α

Calculate p - value

This question provides us only with sample standard deviations so it is a t test. T- test is valid because both sample sizes are > 30.

H0:

H1:

Reject Ho if p-value < α (0.05)

Using the Excel Data / Data Analysis F Test Two Sample for Variances function

From Excel we get:

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances |  |  |
|  |  |  |
|  | *US* | *Japan* |
| Mean | 24545 | 25243 |
| Variance | 4612088 | 1988948 |
| Observations | 50 | 30 |
| df | 49 | 29 |
| F | 2.318859 |  |
| P(F<=f) one-tail | 0.008728 |  |
| F Critical one-tail | 1.777387 |  |

p-value for a one tail test is p – value = 0.0087 and hence p – value for this test is 2 x 0.0087 = 0.0174

p-value is less than α = 0.05 and hence we can reject the null hypothesis and conclude the variances are NOT equal

As such we use the t-test: Two Sample Assuming Unequal Variances

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances |  |  |
|  |  |  |
|  | *US* | *Japan* |
| Mean | 24545 | 25243 |
| Variance | 4612088 | 1988948 |
| Observations | 50 | 30 |
| Hypothesized Mean Difference | 0 |  |
| df | 77 |  |
| t Stat | -1.75302 |  |
| P(T<=t) one-tail | 0.04179 |  |
| t Critical one-tail | 1.664885 |  |
| P(T<=t) two-tail | 0.083579 |  |
| t Critical two-tail | 1.991254 |  |

You can see from this table that the p-value for a one tail test is 0.042.

P – value = 0.042 is less than α of both 0.05 and 0.1 and hence we can reject the null hypothesis at both of those levels and conclude that the Japanese car prices are lower on average in Japan than the US and hence they may be creating an unfair advantage.

Note that if we used α = 0.01 we would have drawn a different conclusion. This highlights the importance of either setting the correct α up front, or consider the p-value and what it means and draw our conclusion from that.

Question 14

The use of gifts or special promotions has long been used by many firms to attract customers away from competitors. The problem with the use of such tactics is that they attract customers who are likely to be less loyal. Hence results may only be a short term phenomenon, not cost effective. A study of one such program for a major bank looked at the loyalty of customers who came to the bank via the promotion as compared to those that opened accounts for other reasons. A random sample of 200 customers who opened accounts for non-promotion reasons showed that 178 of these customers still had an active account one year later. Of 300 customers who opened accounts because of the promotion, 216 had accounts still active 12 months later.

Conduct a hypothesis test at the 10% level of significance to determine whether or not there is greater loyalty in the group of customers who opened accounts for non promotion reasons.

Hypotheses

This is a comparison of two populations where we want to look at the PROPORTION of still active customers among those who came for promotion on non special promotion reasons.

Let 1 – Customers who came for non promotion purposes

Let 2 – Customers who came because of a special promotion.

H0: P1 ≤ P2 or rather P1 – P2 ≤ 0

H1: P1 > P2 or rather P1 – P2 > 0

Decision Rule:

Reject H0 if p – value < α (0.1)

Calculate the p – value

As a test of proportions from two samples, the sampling distribution of the difference in sample proportions is normal provided:

x1 , (n1 – x1), x2 and (n2 – x2) are all > 10

178, (200 – 178), 216, (300 – 216) are all > 10 and hence we can use the normal distribution – sampling distribution of the difference in sample proportions is normal

As an upper one tail test using z, the p – value is calculated using:

P – value = 1 – norm.dist((, (), , true))

P1 – P2 = 0 (from hypotheses)

P – value = 1 – norm.dist((, (), , true)) = 1 – norm.dist( 0.17, 0, 0.0373, true) = 1 – 0.99999 = 0.00001

Hence the p – value is less than α = 0.1 and we can reject the null hypothesis and conclude that the loyalty among non-promotion customers is greater than for those who came for a special promotion as the proportion retained was higher.

Given that p – value is essentially 0 we could conclude that this is true at any level, which gives us great confidence that promotions do not attract loyal customers.

Question 15

It has been argued that the Labour party in Australia places a greater emphasis on so called “Social Issues” including those to do with race and gender. This may affect the way the party is seen by non Australian born voters. A random sample of 500 Australian born voters included 210 who indicated a preference for the Labour party. In a sample of 200 non Australian born voters, 105 indicated a preference for the Labour party. Does the data indicate that there is a difference in support for the Labour party between Australian born and Non Australian born voters? Use a 1% level of significance.

Hypotheses:

This question asks for us to test if there is a difference in support for the Labour party among the two voter groups – Australian born and non-Australian born. Hence this is a two tail test of the DIFFERENCE IN POPULATION PROPORTIONS

Let 1 be Australian born

Let 2 be non – Australian born

H0: P1 = P2 or rather P1 – P2 = 0

H1: P1 ≠ P2 or rather P1 – P2 ≠ 0

Decision Rule:

Reject Ho if p – value < α (0.01)

Calculate the p – value

As a test of proportions from two samples, the sampling distribution of the difference in sample proportions is normal provided:

x1 , (n1 – x1), x2 and (n2 – x2) are all > 10

210, (500 – 210), 105, (200 – 105) are all > 10 and hence we can use the normal distribution – sampling distribution of the difference in sample proportions is normal

Calculation of the p – value depends upon if is less than or greater than 0.

If : p – value = 2 x norm.dist((, (), , true))

If : p – value = 2 x (1 - norm.dist((, (), , true)))

p – value = 2 x norm.dist((, (), , true))

P1 – P2 = 0 (from hypotheses)

p – value = 2 x norm.dist((, (), , true)) = 2 x norm.dist(-0.105, 0, 0.0416, true) = 2 x 00582 = 0.01164

Hence the p – value = 0.01164 is not less than α = 0.01 and we can not reject the null hypothesis. There is insufficient evidence to conclude there is a difference in the preference for the Labour party among Australian born and non-Australian born voters.

That said, for any level of significance higher than 0.02 (e.g. 2% or 5% etc) we would reject H0

Question 16

Behavioural researchers have developed an index which measures managerial success based on length of time in the organisation, level of attainment, colleague acceptance etc. A question of interest is whether or not managers who have significant external contacts (clients, suppliers, government etc) achieve a better level of success than those who spend the majority of their time dealing only with people within the work unit. Independent random samples of managers from within each of these two groups produced the following results with regard to the success index.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Managerial Success Index (Out of 100) | | | | | | | | | | | | |
| External Contact | | | | | |  | Internal Operations | | | | | |
| 65 | 58 | 78 | 60 | 68 | 69 |  | 62 | 53 | 36 | 34 | 56 | 50 |
| 66 | 70 | 53 | 71 | 63 | 63 |  | 42 | 57 | 46 | 68 | 48 | 42 |
|  |  |  |  |  |  |  | 52 | 53 | 43 |  |  |  |

Conduct a test at the 5% level of significance to if the average success rating of managers with significant external contacts is higher than that of those with primarily internal relationships.

Hypotheses

Let 1 – Managers with significant external contacts

Let 2 – Managers who concentrate on internal operations

H0: µ1 ≤ µ2 or rather µ1 - µ2 ≤ 0

H1: µ1 > µ2 or rather µ1 - µ2 > 0

This is an upper one tail test using a level of significance of α = 0.05.

Decision Rule:

Reject H0 if p – value < α (0.05)

Calculate p -value

This is a t test as the only standard deviations we can get come from the sample. This can only be done if we assume the index scores are normally distributed. This is necessary as the sample sizes are individually too small. If this assumption is not true then the test can only be done using other methods (not covered in this course). For the purpose of this question we will make the assumption.

In order to conduct the t test we use Excel

First step is to determine whether or not the variances are equal

To do this we need to conduct an initial test of equality of variances which is an F test

H0:

H1:

Reject Ho if p-value < α (0.05)

Using the Excel Data / Data Analysis F Test Two Sample for Variances function

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances | | |
|  |  |  |
|  | *External Contact* | *Internal Operations* |
| Mean | 65.333333 | 49.4666667 |
| Variance | 43.69697 | 87.1238095 |
| Observations | 12 | 15 |
| df | 11 | 14 |
| F | 0.5015503 |  |
| P(F<=f) one-tail | 0.1276907 |  |
| F Critical one-tail | 0.3651436 |  |

Here the p – value for the one tail test is 0.128

Hence two tail p -value = 0.128 x 2 = 0.256

We can not reject the null hypothesis and we conclude that the population variances are equal.

We can now access the t-Test: Two sample assuming equal variances

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Equal Variances | | |
|  | *External Contact* | *Internal Operations* |
| Mean | 65.333333 | 49.4666667 |
| Variance | 43.69697 | 87.1238095 |
| Observations | 12 | 15 |
| Pooled Variance | 68.016 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 25 |  |
| t Stat | 4.9674617 |  |
| P(T<=t) one-tail | 2.027E-05 |  |
| t Critical one-tail | 1.7081408 |  |
| P(T<=t) two-tail | 4.055E-05 |  |
| t Critical two-tail | 2.0595386 |  |

Going back to our decision rule

Reject Ho if p-value < α(0.05)

The p-value for our one tail test is 0.0000203

Hence the probability the null hypothesis is correct is very small. We can reject the null hypothesis and conclude that the average score for managers with external contacts is greater than that of those who operate mostly internally.

**Question 17**

A question often posed is whether or not Males or Females perform better in Data Analysis for Business. A random sample of 13 males and 17 females produced the following results with regards to their final scores in the unit:.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Males | Females |  |
| Average | 63.63 | 70.73 | (Yes gentlemen , this is real data collected from one of the last semesters I taught) |
| Standard deviation | 21.35 | 17.13 |  |

Data for this question can be found in the Excel file. Using that data conduct the test to determine if females perform better than males i.e. score more on average.

Hypotheses

Here we wish to test if the is a average score of one group (males) is less than another, hence it is a one tail test of the DIFFERENCE IN MEANS.

Define population 1 as Males and population 2 as Females

H0: µ1 ≥ µ2 or rather µ1 - µ2 ≥ 0

H1: µ1 < µ2 or rather µ1 - µ2 < 0

Decision Rule

Reject H0 if p – value < α (not defined here)

Calculate p – value

Step 1 is to determine if the variances are equal or not which requires an F test

H0:

H1:

Reject Ho if p-value < α (0.05)

Using the Excel Data / Data Analysis F Test Two Sample for Variances function

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances |  |  |
|  |  |  |
|  | *Females* | *Males* |
| Mean | 70.72941 | 63.63077 |
| Variance | 293.3797 | 455.904 |
| Observations | 17 | 13 |
| df | 16 | 12 |
| F | 0.643512 |  |
| P(F<=f) one-tail | 0.202461 |  |
| F Critical one-tail | 0.50368 |  |

Here the p – value for the one tail test is 0.202

Hence two tail p -value = 0.202 x 2 = 0.404

We can not reject the null hypothesis and we conclude that the population variances are equal.

We can now complete the test using t test: Two sample assuming equal variances

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Equal Variances |  |  |
|  |  |  |
|  | *Males* | *Females* |
| Mean | 63.63077 | 70.72941 |
| Variance | 455.904 | 293.3797 |
| Observations | 13 | 17 |
| Pooled Variance | 363.033 |  |
| Hypothesized Mean Difference | 0 |  |
| Df | 28 |  |
| t Stat | -1.0112 |  |
| P(T<=t) one-tail | 0.160289 |  |
| t Critical one-tail | 1.312527 |  |
| P(T<=t) two-tail | 0.320579 |  |
| t Critical two-tail | 1.701131 |  |

p-value for a one tail test is p = 0.1603,

This p – value implies there is a 16% chance the null is true. We would not reject H0 at any reasonable value. As given in the lecture a p – value of more than 0.1 is considered to be weak evidence that the null is not true, and hence in this instance we conclude there is insufficient evidence that females perform better on average than males.

Question 18

One way of examining the effectiveness of the teaching and assessment process is to reflect on the overall grades of students. Not only should this be done with regard to average scores, but also the variation in scores. Not all students are equally intelligent, or work equally hard, and the variance of scores on a test should reflect this. In one course a lecturer sets an end of semester examination which typically as a standard deviation of about 15 marks (or a variance of 225). A random sample of 30 students results are selected to test whether the variance for this group is less than 225 (which would imply the test did not appropriately differentiate students). The variance for the sample was 169. Conduct the test at a 10% level of significance.

Hypotheses

In this test we are testing the variance of student scores and we wish to see if it is less than 225. This is lower one tail test on variance

H0: σ2 ≥ 225

H1: σ2 < 225

Decision Rule:

Reject H0 if p – value < α (0.1)

Calculate p – value

Test of variance is a χ2 test provided the data has been selected from a normally distributed population. Here we have not been told this, however student scores are usually reasonably normally distributed (similar to IQ’s) hence this is a reasonable assumption to make.

For a lower one tail test of variances, the p – value is:

P – value = chisq.dist(χ2calc, degrees of freedom , cumulative)

= chisq.dist(χ2calc, v, true)

S2 = 169

σ2 = 225 (from hypotheses)

v = n-1

n = 30

v = 29

P – value = = chisq.dist(21.782, 29, true) = 0.171

p-value = 0.171 is not less than α = 0.1 and hence there is insufficient evidence at the 10% level of significance to conclude the variance of scores on the examination are less than 169.

Question 19

One technique that historians use to identify the decline of any civilization is their use of gold. It is thought that the decline of the Roman Empire started around the time of Emperor Constantine. A sample of six coins dated before the time of Constantine were compared with a sample of 8 dated after the rule of Constantine. The percentage of gold contained in each coin is given below:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Gold Content (%) | | | | | | | |
| Pre Constantine: | 5.4 | 5.2 | 6.1 | 6.4 | 5.1 | 5.9 |  |  |
| Post Constantine: | 6.0 | 5.2 | 4.6 | 4.9 | 5.0 | 5.1 | 4.3 | 4.2 |

Assuming the distribution of gold content is normal, is there evidence at the 1% level of significance that the average gold content post Constantine was lower than pre Constantine?

Hypotheses

Here we wish to test if the average gold content Pre Constantine is higher than post Constantine.

Define population 1 as Pre and population 2 as Post

H0: µ1 ≤ µ2 or rather µ1 - µ2 ≤ 0

H1: µ1 > µ2 or rather µ1 - µ2 > 0 (which implies post Constantine is lower than pre)

Decision Rule:

Reject H0 if p – value < α (0.01)

Calculate p – value

The test will be conducted using a t distribution because the standard deviations come from the two samples. We can do this because we have the assumption that the population distributions are normal.

Before doing the test of means we first need to check if the population variances are equal or not which is an F test

H0:

H1:

Reject Ho if p-value < α (0.01)

Using the Excel Data / Data Analysis F Test Two Sample for Variances function

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances | | |
|  |  |  |
|  | *Pre Constantine* | *Post Constantine* |
| Mean | 5.683333333 | 4.9125 |
| Variance | 0.277666667 | 0.326964286 |
| Observations | 6 | 8 |
| df | 5 | 7 |
| F | 0.849226288 |  |
| P(F<=f) one-tail | 0.443770362 |  |
| F Critical one-tail | 0.095643342 |  |

Here the p – value for the one tail test is 0.444

Hence p -value = 0.444 x 2 = 0.888

Decision Rule is Reject H0 if p – value < α (0.01 from the question)

We do not reject the null hypothesis and conclude that the two variances are equal.

Accessing the t-Test: Two sample assuming equal variances function in Excel we get:

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Equal Variances | | |
|  |  |  |
|  | *Pre Constantine* | *Post Constantine* |
| Mean | 5.683333333 | 4.9125 |
| Variance | 0.277666667 | 0.326964286 |
| Observations | 6 | 8 |
| Pooled Variance | 0.306423611 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 12 |  |
| t Stat | 2.578433733 |  |
| P(T<=t) one-tail | 0.012083682 |  |
| t Critical one-tail | 2.680997993 |  |
| P(T<=t) two-tail | 0.024167363 |  |
| t Critical two-tail | 3.054539589 |  |

Here the p-value for the one tail test we are conducting is p-value = 0.0121

This is not less than our value of α = 0.01 and hence at the 1% level of significance we can not reject the null hypothesis and conclude there is insufficient evidence to say that the gold content has decreased.

Note that if we had used α = 0.05 we would have made a different decision. This p – value falls in the strong evidence region that the null is not true, and hence there is strong evidence that the gold content has fallen on average – but not conclusive at the evidence level we set.

**Question 20**

In order to meet established standards, it is important that the variance of the chemical impurity level is not greater than 4%. A random sample of 20 consignments had a sample variance of 5.62%. Conduct a test at the 5% level to **determine if the variance** of the chemical impurity level is **more** than 4%.

Hypotheses

In this test we are testing the variance of chemical impurity and we want to see if it is more than 4%. This is an upper tail, one tail test on variance

H0: σ2 ≤ 4

H1: σ2 > 4

Decision Rule:

Reject Ho if p – value < α (0.05)

Calculate p – value

Test of variances are conducted using a χ2 distribution provided the data is selected from a normally distributed population. We will need to make that assumption for this question as it is not given.

For an upper one tail test of variances, the p – value is calculated as:

P – value = chisq.dist.rt(χ2calc, degrees of freedom , cumulative)

v = n-1

n = 20

v = 19

s2 = 5.62

σ2 = 4 as given in hypotheses

P – value = chisq.dist.rt = (26.695, 19) = 0.112

P – value = 0.112 is not less than α = 0.05 and hence there is insufficient evidence at the 5% level of significance to conclude the variance of chemical impurities is greater than 4%.

**Question 21**

It is important that filling processes have as little variation as possible to ensure that containers are not overfilled (resulting in waste) or under filled (resulting in the customer getting less than what they paid for). One company manufacturing filling machinery boasts that their machinery is so accurate that the variance of fill for a 1 litre container is less than 1 millilitre. A potential customer interested in testing the accuracy of this claim selected 25 containers from a days run, and measured the fill. From each measurement they took away 1 litre. The 25 containers gave the following results (recorded in ml):

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.3 | -0.5 | 1.0 | -0.3 | -0.7 | -0.2 | -1.3 | 0.6 | -0.6 | 1.0 | -0.6 | -0.5 | -1.5 |
| 1.3 | -0.4 | -0.2 | 0.0 | -1.5 | 1.4 | -1.9 | 0.7 | -0.9 | 1.1 | 0.7 | -0.6 |  |

Assuming the filling process works according to a normal distribution conduct a test at the 5% level of significance to determine whether or not the manufacturing companies claim is true that the variance of fill is less than 1ml.

Hypotheses

In this test we are testing the variance of the filling mechanism is less than 1ml. This is a lower tail, one tail test on variance

H0: σ2 ≥ 1

H1: σ2 < 1

Decision Rule:

Reject Ho if p – value < α (0.05)

Calculate p – value

Test of variance is conducted using a χ2 distribution provided the data is selected from a normally distributed population. This question indicates the process does fill according to a normal distribution so the chi-squared test is valid.

For a lower one tail test of variance the p – value is calculated as:

P – value = chisq.dist(χ2calc, degrees of freedom , cumulative)

= chisq.dist(χ2calc, v, true)

Using descriptive statistics we get

|  |  |
| --- | --- |
| Mean | -0.144 |
| Standard Error | 0.186107 |
| Median | -0.3 |
| Mode | -0.6 |
| Standard Deviation | 0.930537 |
| Sample Variance | 0.8659 |
| Kurtosis | -0.85533 |
| Skewness | 0.020814 |
| Range | 3.3 |
| Minimum | -1.9 |
| Maximum | 1.4 |
| Sum | -3.6 |
| Count | 25 |

v = n-1

n = 25

v = 24

s2 = 0.8659

σ2 = 1 (from hypotheses)

P – value = chisq.dist(χ2calc, v, true) = chisq.dist(20.78, 24, true) = 0.3484

p-value = 0.3484 is not less than α = 0.05 and hence there is insufficient evidence at the 5% level of significance to support the companies claim that the machinery is accurate enough to reduce variance to less than 1ml.

In fact p – value of 0.34 is very weak evidence the null is false and as such can not support any claims of better variation.