

Final

Lewis Collum

Updated: May 7, 2020

PART 1:

- 9: RBD Anova
 - 13: Exponential
 - 14: Normal
 - 15: Exponential
 - 16: Poisson
 - 17: Binomial
-

1

Hypotheses:

C: Clarkson

R: RPI

$$H_0 : \mu_C = \mu_R$$

$$H_A : \mu_C > \mu_R$$

Two Sample T-Test:

```
import pandas
from scipy import stats

data = pandas.read_excel('data.xlsx')
print(f'variance p-value = {stats.levene(data.Clarkson, data.RPI).pvalue:.3f}')
tScore, pTwoTailed = stats.ttest_ind(data.Clarkson, data.RPI, equal_var = False)

pOneTailed = pTwoTailed/2
print(f"\nt-test p-value = {pOneTailed:.2f}")
```

variance p-value = 0.000

t-test p-value = 0.19

Conclusion

Since the p-value is greater than the significance level (0.05), we fail to reject the null hypothesis. That is, we cannot conclude that Clarkson students have a higher starting salary on average than those at RPI.

2

Hypotheses:

$$H_0 : \mu = 30$$

$$H_A : \mu > 30$$

One sample T-Test:

```
import pandas
from scipy import stats

data = pandas.read_excel('data.xlsx', sheet_name=1)

tScore, pTwoTailed = stats.ttest_1samp(data.Mileage, 30)
pOneTailed = pTwoTailed/2
print(f't-score = {tScore:.2f}')
print(f"\np-value = {pOneTailed:.10f}")
```

t-score = 6.11

p-value = 0.0000013019

Conclusion

Since the p-value is less than the significance level (0.05), we reject the null hypothesis. Evidence suggests that the mileage is over 30 mpg for the sedan.

3

Hypotheses:

$$H_0 : \mu_D = \mu_P$$

$$H_A : \mu_D < \mu_P$$

Paired T-Test:

```
import pandas
from scipy import stats

data = pandas.read_excel('data.xlsx', sheet_name=2)

tScore, pTwoTailed = stats.ttest_rel(data.Placebo, data.Drug)
pOneTailed = pTwoTailed/2
print(f't-score = {tScore:.2f}')
print(f"\np-value = {pOneTailed:.10f}")
```

t-score = 10.33

p-value = 0.0000000000

Conclusion

Since the p-value is less than the significance level (0.05), we reject the null hypothesis. Evidence suggests that the drug does aid in weight loss.

4

Hypotheses:

I: Idaho

W: Washington

O: Oregon

N: Wisconsin

D: North Dakota

$$H_0 : \mu_I = \mu_W = \mu_O = \mu_N = \mu_D$$

$$H_A : \mu_I > \mu_i \text{ for } i \text{ in } [W, O, N, D]$$

Multiple Two Sample T-Test:

```
import pandas
from scipy import stats

data = pandas.read_excel('data.xlsx', sheet_name = 3)
#print(f'variance p-value = {stats.levene(data.Clarkson, data.RPI).pvalue:.3f}')
#tScore, pTwoTailed = stats.ttest_ind(data.Clarkson, data.RPI, equal_var = False)

#pOneTailed = pTwoTailed/2
#print(f"\nt-test p-value = {pOneTailed:.2f}")
```

Ran out of time

Conclusion

Since the p-value is greater than the significance level (0.05), we fail to reject the null hypothesis. Evidence suggests that Idaho Potato Growers Association potatoes have the highest satisfaction.

5

(a)

$$P(D) = 0.01$$

$$P(\sim D) = 1 - 0.01 = 0.99$$

$$P(A) = (0.99)^{15}$$

$$= 0.86$$

(b)

$$P(D) = 0.10$$

$$P(\sim D) = 1 - 0.10 = 0.9$$

$$P(A) = (0.9)^{15}$$

$$= 0.206$$

6

(a)

$$\text{Test 1: } P(D) = 0.25 \cdot 0.79$$

$$\text{Test 2: } P(D) = 0.25 \cdot 0.76$$

(b) ran out of time