SUJEET KUMAR
BITS ID 2018AD04062

Statistics Assignment : 24 Sept 2019

Problem No 1: **Testing of Hypothesis**

A manufacturing company claims that the mean breaking strength of the cables supplied s 1500 with a standard deviation of 100.A sample of size 9 cables tested is having mean breaking strength of 1600 with a standard deviation of 30. If the distribution of the mean breaking strength is normal, does this support the claim at 1% level of significance?

- Use z – statistic to validate the claim
- Use t – statistics to validate the claim
- Repeat the same at 5% level of significance
- Based on (a),(b) and (c) write your observations , comments and suggestions

t-test implementation(Python)

```python
from numpy import mean
from scipy.stats import sem
from scipy.stats import t

# function for calculating the t-test
def independent_ttest(mean1,se1,mean2,se2,samplesize, alpha):
    alpha=alpha
# standard error on the difference between the samples
    sed = se2/sqrt(samplesize)
# calculate the t statistic
    t_stat = (mean2 - mean1) / sed
# degrees of freedom
    df = samplesize-1
# calculate the critical value
    cv = t.ppf(1-alpha/2, df)
    return t_stat, df, cv
t_stat, df, cv= independent_ttest(1500, 100,1600,30,9, .01)
print('t=%.3f, df=%d, cv=%.3f' % (t_stat, df, cv))
```

t-test output : alpha =1% and 5%

```python
# Run the t-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.01
print('Run the t-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.01')
t_stat, df, cv = independent_ttest(1500, 100,1600,30,9, .01)
print('t=%.3f, df=%d, cv=%.3f '% (t_stat, df, cv))
interpret(t_stat,cv)
```

```
Run the t-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.01
t=10.000, df=8, cv=3.355
Reject the null hypothesis
```

```python
# Run the t-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.05
print('Run the t-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.05')
t_stat, df, cv = independent_ttest(1500, 100,1600,30,9, .05)
print('t=%.3f, df=%d, cv=%.3f' % (t_stat, df, cv))
interpret(t_stat,cv)
```

```
Run the t-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.05
t=10.000, df=8, cv=2.306
Reject the null hypothesis
```

Z-test implementation(Python)

lit View Insert Cell Kernel Widgets Help

Code

```python
def independent_Ztest(mean1,se1,mean2,se2,samplesize, alpha):
    alpha=alpha
# standard error on the difference between the samples
    sed = se1/sqrt(samplesize)
# calculate the t statistic
    z_stat = (mean2 - mean1) / sed
# calculate the critical value
    cv1=st.norm.cdf(z_stat)
# calculate the p-value
    p = (1.0 - st.norm.cdf(z_stat)) * 2.0
    return z_stat, cv1,p
z_stat, cv1,P = independent_Ztest(1500, 100,1600,30,9, .01)
print('Z=%.3f, cv=%.3f, p=%3f' % (z_stat, cv1,p))
# interpret via critical value
def interpret(alpha,cv1):
    if abs(1-alpha/2) > cv1:
        print('Accept null hypothesis .')
    else:
        print('Reject the null hypothesis ')
# interpret via p-value
def interpret_using_PValue(p, alpha):
    print('Interpreting Using P value........ ')
    if p > alpha:
        print('Accept null hypothesis ')
    else:
        print('P < alpha : Reject the null hypothesis ')
```

```
Z=3.000, cv=0.999, p=0.002700
```

Z-test output : alpha =1% and 5%

File Edit View Insert Cell Kernel Widgets Help

Code

In [107]:

```
# Run the Z-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.01
print('Run the Z-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.01')
z_stat, cv, p = independent_Ztest(1500, 100,1600,30,9, .01)
x=abs(1-.01/2)
print('z=%.3f, cv=%.3f, p=%.5f ,alpha=%.3f '% (z_stat, cv, p,x))
interpret(z_stat,cv)
interpret_using_PValue(p,.01)
```

```
Run the Z-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.01
z=3.000, cv=0.999, p=0.00270 ,alpha=0.995
Reject the null hypothesis
Interpreting Using P value........
P < alpha : Reject the null hypothesis
```

In [118]:

```
# Run the Z-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.05
print('Run the Z-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.05')
z_stat, cv, p = independent_Ztest(1500, 100,1600,30,9, .05)
x=abs(1-.05/2)
print('z=%.3f, cv=%.3f, p=%.5f ,alpha=%.3f '% (z_stat, cv, p,x))
interpret(z_stat,cv)
interpret_using_PValue(p,.05)
```

```
Run the Z-test with mean 1500,SD=100,Mean 1600 SD= 30 sample = 9 and alpha =.05
z=3.000, cv=0.999, p=0.00270 ,alpha=0.975
Reject the null hypothesis
Interpreting Using P value........
P < alpha : Reject the null hypothesis
```

T-test : Observation/Comments/Suggestions

t-test (alpha =5% and 1%)

P values equals 0.00000848.This means change of type-1 error is small. Smaller p value the more it supports Ha.

t-test(alpha =5%)

Test statistics t equals 10 is not in the 95% critical value accepted range [-2.306,2.306]. Mean =1600 is not acceptance in 95% range.

t-test (alpha = 1%)

Test statistics t equals =10 is not in range of 99% critical value accepted range [-3.3554,3,3554].Mean =1600 is not accepted in 99 % range

(Sample size =9 < 30), suggests we should go for t-statistical test

Ps: Scan document of solution Is attached

Z-test : Observation/Comments/Suggestions

z-test (alpha =5% and 1%)

P values equals 0.0026.This means change of type-1 error is small. Smaller p value the more it supports Ha.

z-test(alpha =5%)

Test statistics z=3.0 is not in the 95% critical value. Mean =1600 is not accepted in 95% range.

z-test (alpha = 1%)

Test statistics z=3.0 is not in range of 99% critical value. Mean =1600 is not accepted in 99 % range

(Sample size =9 < 30), suggests we should go for t-statistical test

Ps: Scan document of solution Is attached

Problem No 2: Effectiveness of the drug.

- An experiment was conducted to test the effectiveness of the drug. 300 patients were treated with new drug and 200 were not treated with the drug. The data is given below. Use an appropriate test and comment on the effectiveness of the drug.

Details	Cured	Condition worsened	No Effect	Total
Treated with the drug	180	20	100	300
Not treated with the drug	140	50	10	200
Total	**320**	**70**	**110**	**500**

Chi-square implementation(Python)

ter AssignmentQ3 Last Checkpoint: 10 hours ago (autosaved)

Edit View Insert Cell Kernel Widgets Help

Code

5]:

```
# chi-squared test with similar proportions
from scipy.stats import chi2_contingency
from scipy.stats import chi2
# contingency table
table = [[180, 20, 100],[140, 50,  10]]
print(table)
stat, p, dof, expected = chi2_contingency(table)
print('dof=%d' % dof)
print(expected)
# interpret test-statistic
prob = 0.95
critical = chi2.ppf(prob, dof)
print('probability=%.3f, critical=%.3f, stat=%.3f' % (prob, critical, sta
if abs(stat) >= critical:
    print('Dependent (reject H0)')
else:
    print('Independent (fail to reject H0)')
# interpret p-value
alpha = 1.0 - prob
print('significance=%.3f, p=%.3f' % (alpha, p))
if p <= alpha:
    print('Dependent (reject H0)')
else:
    print('Independent (fail to reject H0)')
```

Chi-square: output

```
[[180, 20, 100], [140, 50, 10]]
dof=2
[[ 192.   42.   66.]
 [ 128.   28.   44.]]
probability=0.950, critical=5.991, stat=74.472
Dependent (reject H0)
significance=0.050, p=0.000
Dependent (reject H0)
```

Ps: Solution scan copy is attached