Statistics Mini Project

Arun Kumar P

May 30, 2018

R Markdown

Question 1:

Describe the five percent significance test you would apply to these data to determine whether new scheme has significantly raised outputs? What conclusion does the test lead to?

Solution:

H0->Mean of new scheme = Mean of old scheme

H0-> mu = 68.03

HA-> Mean of new scheme != Mean of old scheme

HA-> mu != 68.03

alpha-> 0.05 (5%)

```
d<-read.csv("rawdata2.csv",header = TRUE)
summary(d)
```

```
##     OldData          NewData
##  Min.   : 28.00   Min.   : 32.00
##  1st Qu.: 54.00   1st Qu.: 55.00
##  Median : 67.00   Median : 74.00
##  Mean   : 68.03   Mean   : 72.03
##  3rd Qu.: 81.50   3rd Qu.: 85.75
##  Max.   :110.00   Max.   :122.00
```

```
oldData<-d$OldData
newData<-d$NewData
t.test(newData,oldData,alternative ="two.sided",
              conf.level = 0.95,paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  newData and oldData
## t = 1.5559, df = 29, p-value = 0.1306
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.257949  9.257949
## sample estimates:
## mean of the differences
##                       4
```

Result:

```
1.Rejecting Null hypothesis.

2.Yes, the new scheme has significantly raised outputs

3.The company should not abandon the new scheme.
```

Question 2:

Suppose it has been calculated that in order for Titan to break even, the average output must increase by £5000.If this figure is alternative hypothesis, what is:

(i) The probability of a type 1 error?

Solution:

The probability of a Type-I error is alpha value - 0.05 (5%) according to industrial standard.The rules can be made stricter if we decrease the alpha value.

(ii) What is the p- value of the hypothesis test if we test for a difference of $5000?

Solution:

H0->Mean of new scheme <= 5

HA-> Mean of new scheme > 5

alpha-> 0.05 (5%)

difference 5000 i.e mu =5

```
t.test(newData,oldData,mu=5,alternative ="greater",conf.level = 0.95,paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  newData and oldData
## t = -0.38898, df = 29, p-value = 0.6499
## alternative hypothesis: true difference in means is greater than 5
## 95 percent confidence interval:
##  -0.3681762        Inf
## sample estimates:
## mean of the differences
##                       4
```

Result:

```
p value is equal to 0.6499
```

(iii) Power of the test: Say, you specify the hypothesis as follows:

H0: The difference is zero

HA: The difference is 5000.

So you have clearly specified that either the difference is 0 or 5000. If you fail to reject the null hypothesis (null hypothesis of zero difference), you may be committing a type-II error.

What is the probability of committing a type-II error?

Solution:

H0->Mean of new scheme is equal to old scheme

H0 = 0

HA != 0 (Mean difference is 5000)

Calculation the power of test

```
d<-read.csv("rawdata2.csv",header = TRUE)
oldData<-d$OldData
newData<-d$NewData
#Delta
del<-mean(newData) - mean(oldData)
sd1<- sd(newData)
sd2<- sd(oldData)
#Pooled Standard Deviation
pooledSD<-(((30-1)*(sd1^2)+(30-1)*(sd2^2))/(30+30-2))^0.5
power.t.test(n=30, delta = del, sd=pooledSD,
             sig.level = 0.05,type = "paired",
             alternative = "two.sided" )
```

```
##
##      Paired t test power calculation
##
##               n = 30
##           delta = 4
##              sd = 22.33211
##       sig.level = 0.05
##           power = 0.1559796
##     alternative = two.sided
##
## NOTE: n is number of *pairs*, sd is std.dev. of *differences* within pairs
```

Calculation the Type II error Probability (beta)

Power of test = (1-beta)

0.1559796 = 1 - beta

beta = 1-0.155979

beta = 0.844021

Result:

```
The probability of committing a Type II error is  84% (beta = 0.844021)
```