R Notebook For Hypothesis Testing- Shreeya Kamath : 220953004 ; CCE c

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

```
#Hypothesis Testing For Location Parameter
#ONE sample Gauss Test
#One sample z-test
#Hypothesized population mean
population_mean <- 2

# Sample size
 sample_size <- 20
 sample_mean=1.97
 sample_sd =0.1

#Calculate Z test statistic
 z_stat <- (sample_mean-population_mean)/(sample_sd/sqrt(sample_size))

 z_stat
```

```
## [1] -1.341641
```

```
  p_value <- 2 * (1- pnorm(abs(z_stat)))
 # Display the p-value
 print(p_value)
```

```
## [1] 0.1797125
```

```
#Given the sample
 sample = c(1.2, 1.45, 2.1, 1.78, 1.3, 2.2, 1.98, 1.7, 1.67, 2.4)
 population_mean <- 2
 # Sample size
 sample_size <- length(sample)
 sample_mean=mean(sample)
 sample_sd =sd(sample)
 z_stat
```

```
## [1] -1.341641
```

```
 # Calculate the Z test statistic
 z_stat <- (sample_mean- population_mean) / (sample_sd / sqrt(sample_size))

  p_value <- 2 * (1- pnorm(abs(z_stat)))
 # Display the p-value
 print(p_value)
```

```
## [1] 0.07480788
```

```r
#Using Z test function from BSDA package
library(BSDA)
```

```
## Warning: package 'BSDA' was built under R version 4.3.3

## Loading required package: lattice

##
## Attaching package: 'BSDA'

## The following object is masked from 'package:datasets':
##
##     Orange
```

```r
#perform one sample z-test
z.test(sample, mu=2, sigma.x=0.1)
```

```
##
##  One-sample z-Test
##
## data:  sample
## z = -7.0203, p-value = 2.215e-12
## alternative hypothesis: true mean is not equal to 2
## 95 percent confidence interval:
##  1.71602 1.83998
## sample estimates:
## mean of x
##     1.778
```

```r
#Unknown variance one sample
t.test(sample, mu=2)
```

```
##
##  One Sample t-test
##
## data:  sample
## t = -1.7816, df = 9, p-value = 0.1085
## alternative hypothesis: true mean is not equal to 2
## 95 percent confidence interval:
##  1.496126 2.059874
## sample estimates:
## mean of x
##     1.778
```

```r
#Two sample one tailed test (unknown varaince)
#Syntax

x=c(10,12,13,16,11,13,14,15)
y=c(9,8,7,5,4,10,11,12)
t.test(x,y,alternative = 'greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  x and y
## t = 3.89, df = 12.63, p-value = 0.0009798
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  2.582669      Inf
## sample estimates:
## mean of x mean of y
##     13.00      8.25
```

```r
#Two sample two tailed test

x=c(10,12,13,16,11,13,14,15)
y=c(9,8,7,5,4,10,11,12)
t.test(x,y,alternative = 'two.sided')
```

```
##
##  Welch Two Sample t-test
##
## data:  x and y
## t = 3.89, df = 12.63, p-value = 0.00196
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.10411 7.39589
## sample estimates:
## mean of x mean of y
##     13.00      8.25
```

```r
#Two sample Known varaince

cityA = c(82, 84, 85, 89, 91, 91, 92, 94, 99, 99,
105, 109, 109, 109, 110, 112, 112, 113, 114, 114)
cityB = c(90, 91, 91, 91, 95, 95, 99, 99, 108, 109,
109, 114, 115, 116, 117, 117, 128, 129, 130, 133)
```

```r
#perform two sample z-test
z.test(x=cityA, y=cityB, mu=0, sigma.x=15, sigma.y=15)
```

```
##
##  Two-sample z-Test
##
## data:  cityA and cityB
## z = -1.7182, p-value = 0.08577
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -17.446925   1.146925
## sample estimates:
## mean of x mean of y
##    100.65    108.80
```

```r
#z.test(x=cityA, y=cityB, mu=0,alternative = 'two.sided')
```

```r
# paired T test
pre <- c(85, 78, 92, 91, 72, 84, 99, 90, 96, 84)
post <- c(88, 90, 93, 91, 80, 97, 100, 93, 91, 90)
t.test(pre, post, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  pre and post
## t = -2.3744, df = 9, p-value = 0.04161
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -8.2014541 -0.1985459
## sample estimates:
## mean difference
##            -4.2
```

```r
x <- c(512,530,498,540,521,528,505,523)
y <- c(499,500,510,495,515,503,490,511)
t.test(x,y,alternative='greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  x and y
## t = 2.9058, df = 11.672, p-value = 0.006762
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.452174      Inf
## sample estimates:
## mean of x mean of y
##   519.625   502.875
```

```r
#Parametric Tests for Probabilities
#One-Sample Binomial Test for the Probability p

#Take decision based on the probability value.
qbinom(p=0.95, prob=0.3, size=2000)
```

```
## [1] 634
```

```r
binom.test(c(700,1300),p=0.3,alternative='greater')
```

```
##
##  Exact binomial test
##
## data:  c(700, 1300)
## number of successes = 700, number of trials = 2000, p-value = 8.395e-07
```

```
## alternative hypothesis: true probability of success is greater than 0.3
## 95 percent confidence interval:
##  0.332378 1.000000
## sample estimates:
## probability of success
##                   0.35
```

```r
binom.test(x=700,n=2000,p=0.3, alternative='greater')
```

```
##
##  Exact binomial test
##
## data:  700 and 2000
## number of successes = 700, number of trials = 2000, p-value = 8.395e-07
## alternative hypothesis: true probability of success is greater than 0.3
## 95 percent confidence interval:
##  0.332378 1.000000
## sample estimates:
## probability of success
##                   0.35
```

```r
binom.test(x=700,n=2000,p=0.3, alternative='less')
```

```
##
##  Exact binomial test
##
## data:  700 and 2000
## number of successes = 700, number of trials = 2000, p-value = 1
## alternative hypothesis: true probability of success is less than 0.3
## 95 percent confidence interval:
##  0.0000000 0.3679481
## sample estimates:
## probability of success
##                   0.35
```

```r
#Two-Sample Binomial Test
# Example data
successes <- c(14, 13) # Number of successes in each group
trials <- c(63, 45)
# Number of trials in each group
# Perform the two-sample binomial test
result <- prop.test(successes, trials)
# View the result
print(result)
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  successes out of trials
## X-squared = 0.31746, df = 1, p-value = 0.5731
## alternative hypothesis: two.sided
## 95 percent confidence interval:
```

```
##  -0.2532728  0.1199394
## sample estimates:
##    prop 1    prop 2
## 0.2222222 0.2888889
```

```r
#Nonparamatric Test
 #Wilcoxon-Mann-Whitney (WMW) U-Test
 coffee <- c(3.7, 4.9, 5.2, 6.3, 7.4,4.4,5.3,1.7, 2.9)
 water <- c(4.5, 5.1, 6.2,7.3,8.7,4.2,3.3,9.9,2.6, 4.8)
 wilcox.test(coffee, water)
```

```
##
##  Wilcoxon rank sum exact test
##
## data:  coffee and water
## W = 38, p-value = 0.6038
## alternative hypothesis: true location shift is not equal to 0
```

```r
# X2-Goodness of Fit Test
 chisq.test(c(315, 108, 101, 32),
 p=c(9/16,3/16,3/16,1/16))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  c(315, 108, 101, 32)
## X-squared = 0.47002, df = 3, p-value = 0.9254
```

```r
qchisq(df=3, p=0.95)
```

```
## [1] 7.814728
```

```r
# Solution to Que 10.4
x <- c(91,101,42,99,108,88,89,105,111,104)
y <- c(261,47,40,29,64,6,87,47,98,351)
t.test(x,y,alternative= 'greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  x and y
## t = -0.25522, df = 9.5635, p-value = 0.598
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -74.83902       Inf
## sample estimates:
## mean of x mean of y
##      93.8     103.0
```

```r
x <- c(91,101,42,99,108,88,89,105,111,104)
y <- c(261,47,40,29,64,6,87,47,98,351)
wilcox.test(x,y)
```

```
## Warning in wilcox.test.default(x, y): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  x and y
## W = 72, p-value = 0.104
## alternative hypothesis: true location shift is not equal to 0
```