Worksheet14

Sucheta Jhunjhunwala, MDS202151

22/12/2021

```
set.seed(51)

#Solution 1

p_score = c(30, 0, 50, 22, 55, 50, 55, 40, 44, 60) #Scores given by Prashant
i_score = c(20, 10, 40, 11, 44, 30, 33, 20, 33, 60) #scores given by Ishan

#Solution 1.a

#Let the level of significance alpha be 0.05
t.test(p_score,i_score,paired =TRUE)
```

```
##
##  Paired t-test
##
## data:  p_score and i_score
## t = 3.4278, df = 9, p-value = 0.007535
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   3.570509 17.429491
## sample estimates:
## mean of the differences
##                    10.5
```

```
#We take the two vectors representing the scores. The value of the t-statistic
#is calculated as t = 3.4278. Since the length of each vector if 10 (number of quizzes)
#degree of freedom = df = 10-1=9. The probability of not rejecting the null
#Hypothesis is the p-value which is 0.007535
#The null hypothesis says that the two scores have the same mean and the alternate
#hypothesis says that the scores have the different means, that is, the difference
#between the means is 0.
#The 95% confidence interval comes as (3.570509, 17.429491).
#Since this is a paired two sample t.test, the mean of the difference between
#the two scores comes out as 10.5.
#p-value = 0.007535<0.05 so we reject the null hypothesis.

#So we can say that the two scores are statistically different since the mean of
#the difference between the two is 10.5 which is not near to 0 and the p-value is
#also less than 0.05.


#solution 1b

#Let the level of significance alpha be 0.05
t.test(p_score,i_score,var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  p_score and i_score
## t = 1.3733, df = 18, p-value = 0.1865
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5.563316 26.563316
## sample estimates:
## mean of x mean of y
##      40.6      30.1
```

```
#We take the two vectors representing the scores. The value of the t-statistic
#is given by t = 1.3733. degree of freedom = df=18 and the probability of not
#rejecting the null Hypothesis is given by p-value=0.1865
#The null hypothesis says that the two scores have the same mean and the alternate
#hypothesis says that the scores have the different mean.
#The 95% confidence interval comes as (-5.563316, 26.563316). since this is an
#unpaired independent t.test we get two means for the two samples.
#The mean of the first student's score is 40.6 and the second student's score is 30.1.
#The p-value comes as 0.187>0.05 so we can't reject the null hypothesis.

#So we can't say that the two scores are statistically different since the p-value is
#not less than 0.05. It might be possible that the two scores are statistically the same.
#So the two scores are not statistically different.

#Solution 2

top1 = c(1,2,3,4,5,6)
no1 = c(77, 89, 76, 84, 80, 94)
no2= c(50, 199, 102, 52, 44, 53)


#Solution 2a

#Let the level of significance alpha be 0.05
c1 = chisq.test(no1)
c1
```

```
##
##  Chi-squared test for given probabilities
##
## data:  no1
## X-squared = 3.016, df = 5, p-value = 0.6975
```

```
if (c1$p.value<0.05){
  print("The Dice is not fair")
} else {
  print("The Dice may or may not be fair")
}
```

```
## [1] "The Dice may or may not be fair"
```

```
#for the results of the first dice we get the X-squared as 3.016 which is the test
#statistic. The degree of freedom is df = (6-1)(2-1) = 5. The probability of not
#rejecting the null Hypothesis is given by p-value which comes as 0.6975.
#The null hypothesis is that the variables, that is the observations for the two
#dice rolls are independent and the alternate hypothesis is that they are dependent.
#p-value>0.05 so we can't reject the null hypothesis, that is the variables may or
#may not be independent. This in turn implies that the observation made from the
#dice roll may or may not be independent and hence the dice may or may not be fair.
#since the observations are independent only if the dice is fair one.

#Since we fail to reject the null hypothesis, the dice is fair.

#Solution 2b

#Let the level of significance alpha be 0.05
c2 = chisq.test(no2)
c2
```

```
##
##  Chi-squared test for given probabilities
##
## data:  no2
## X-squared = 219.45, df = 5, p-value < 2.2e-16
```

```
if (c2$p.value<0.05){
  print("The Dice is not fair")
} else {
  print("The Dice may or may not be fair")
}
```

```
## [1] "The Dice is not fair"
```

```
#for the results of the first dice we get the X-squared as 219.45 which is the test
#statistic. The degree of freedom is df = (6-1)(2-1) = 5. The probability of not
#rejecting the null Hypothesis is given by p-value which comes less than 2.2e-16.
#The null hypothesis is that the variables, that is the observations for the
#dice roll are independent and the alternate hypothesis is that they are dependent.
#Since the p-value<0.05, we reject the null hypothesis, that is the variables are not
#independent. This in turn implies that the observation made from the
#dice roll are not independent and hence the dice is not fair.
#since the observations are independent only if the dice is fair one.

#We reject the null hypothesis and the dice is not fair.

#Solution 2c
df = data.frame(no1,no2)

#Let the level of significance alpha be 0.05
chisq.test(df)
```

```
##
##  Pearson's Chi-squared test
##
## data:  df
## X-squared = 80.968, df = 5, p-value = 5.263e-16
```

```
#for the results of the first dice we get the X-squared as 80.968 which is the test
#statistic. The degree of freedom is df = (6-1)(2-1) = 5. The probability of rejecting
#the null hypothesis is given by p-value which comes as
#5.263e-16.
#The null hypothesis is that the variables, that is the observations for the two
#dice rolls are independent and the alternate hypothesis is that they are dependent.
#p-value<0.05 so we reject the null hypothesis, that is the variables are
#not independent. This in turn implies that the observation made from the
#dice rolls do not have the same distribution.
```