Assignment-9

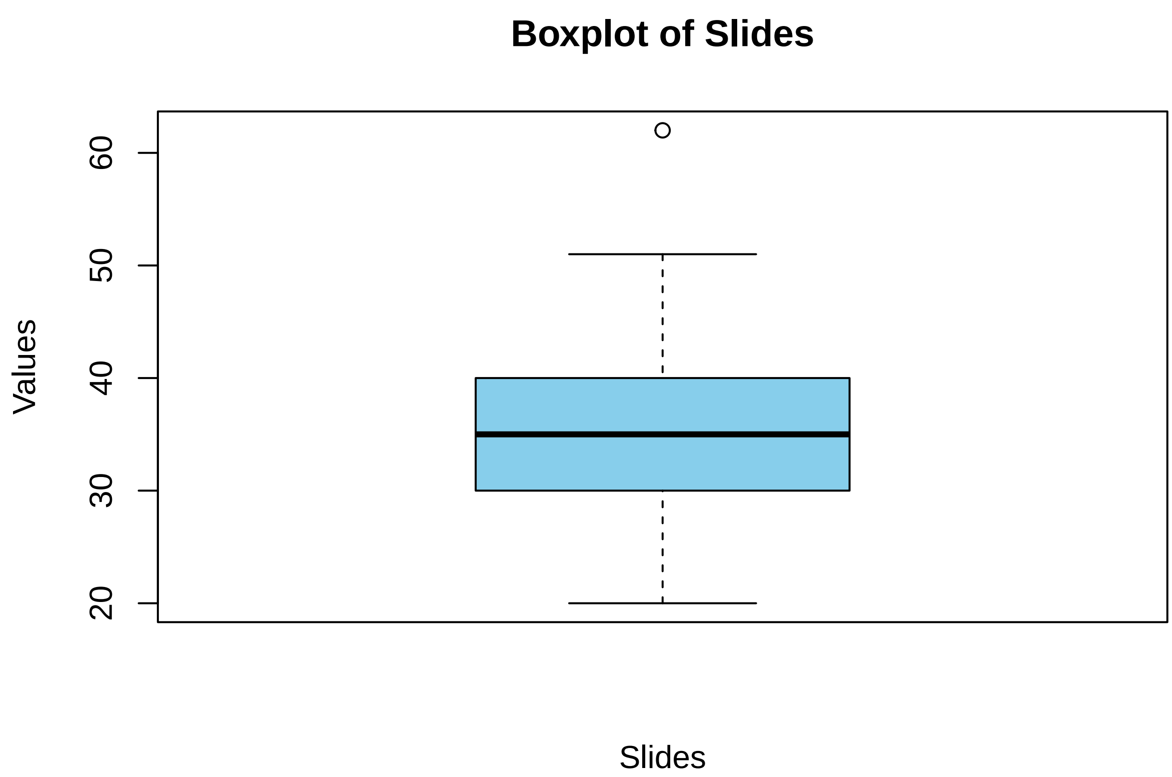
Ashish Verma

2024-04-02

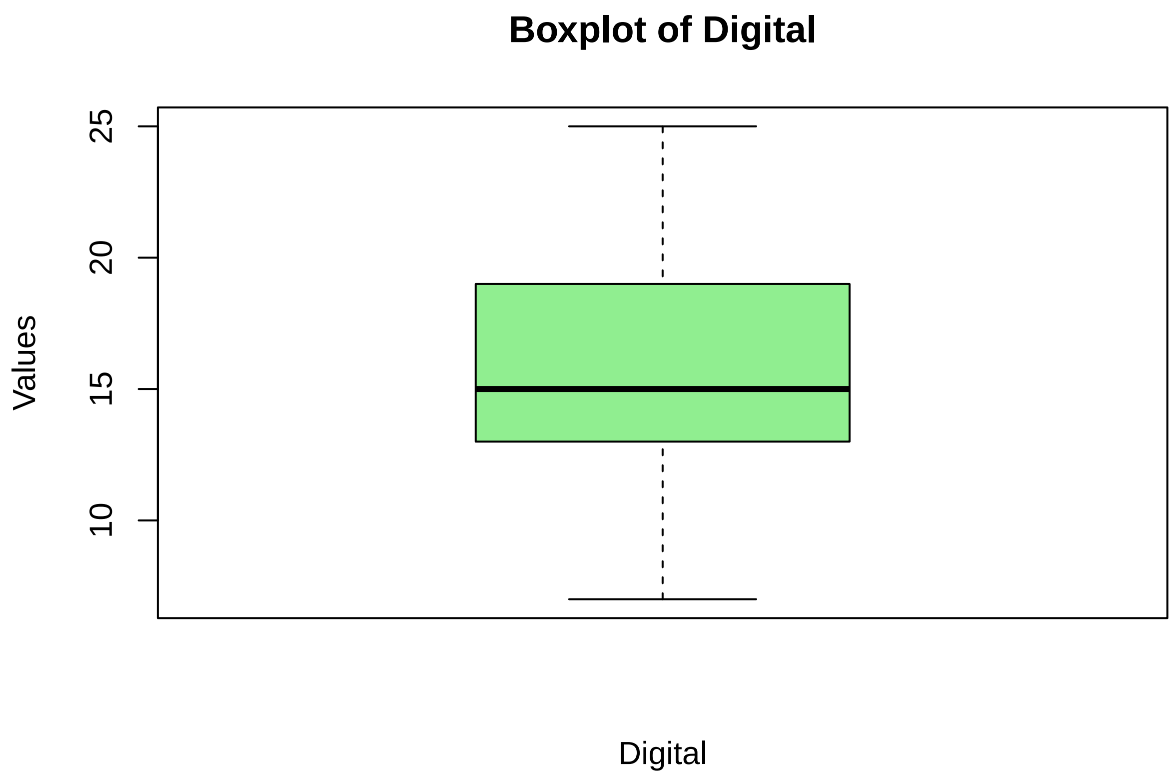
Section 9.3 Problem #38

Solution(a)

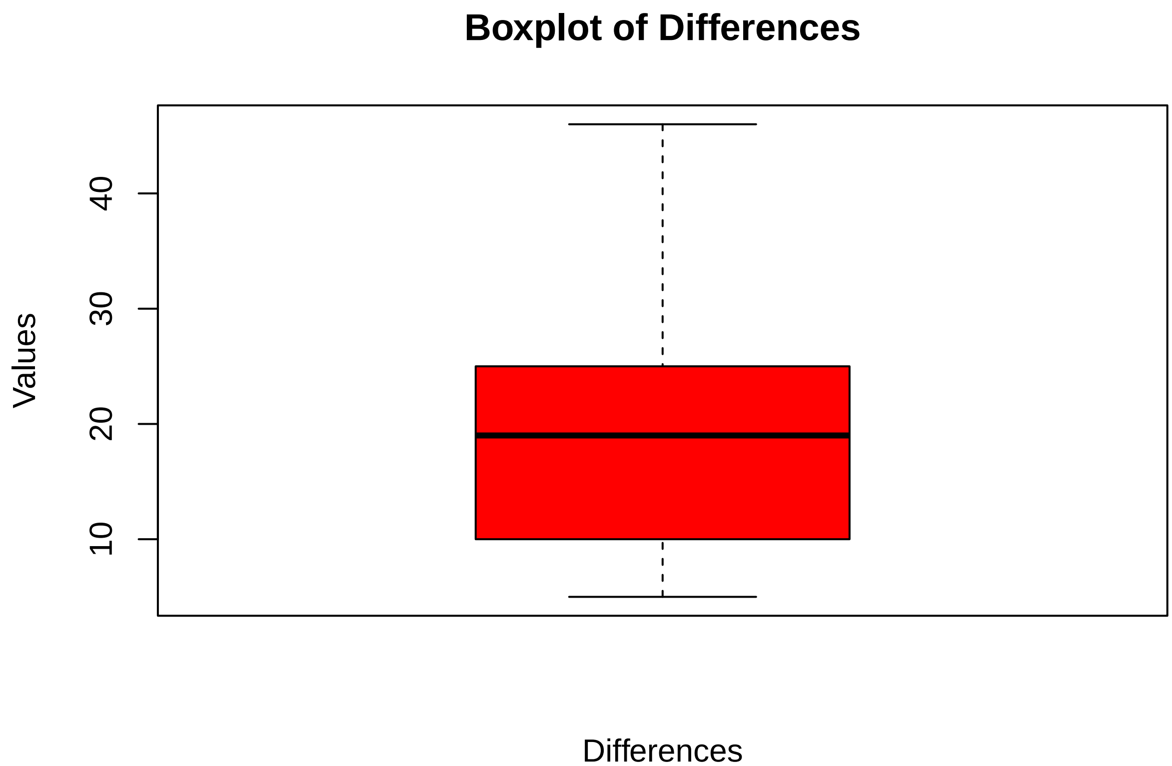
```
x_slide<-c(30,35,40,25,20,30,35,62,40,51,25,42,33)
y_digital<-c(25,16,15,15,10,20,7,16,15,13,11,19,19)
xy_diff<-c(5,19,25,10,10,10,28,46,25,38,14,23,14)

boxplot(x_slide,
        main = "Boxplot of Slides",  # Title of the plot
        xlab = "Slides",             # Label for x-axis
        ylab = "Values",             # Label for y-axis
        col = "skyblue",  # Colors for the boxes
        outline = TRUE)              # Show outliers
```



```r
boxplot(y_digital,
        main = "Boxplot of Digital",  # Title of the plot
        xlab = "Digital",             # Label for x-axis
        ylab = "Values",             # Label for y-axis
        col = "lightgreen",  # Colors for the boxes
        outline = TRUE)              # Show outliers
```




```r
boxplot(xy_diff,
        main = "Boxplot of Differences",  # Title of the plot
        xlab = "Differences",             # Label for x-axis
        ylab = "Values",             # Label for y-axis
        col = "red",  # Colors for the boxes
        outline = TRUE)              # Show outliers
```



Observations:

1.The box plot indicates that retrieval time for professional is much longer of images for slides than digital.

2.There is one outlier in the slides data.

3.The median value of slides is greater than median value of digital.

4.There are no outliers in digital and differences data.

Solution(b)

Null Hypothesis:H0:mu1=mu2

Alternate Hypothesis:Ha:mu1!=mu2

```
x_slide<-c(30,35,40,25,20,30,35,62,40,51,25,42,33)
y_digital<-c(25,16,15,15,10,20,7,16,15,13,11,19,19)
xy_diff<-c(5,19,25,10,10,10,28,46,25,38,14,23,14)
xy_diff_bar<-mean(xy_diff)
n<-13
alpha<-0.05
degree_of_freedom <- n -1
sqaure_sum_diff<-sum((xy_diff^2))
square_whole_diff<-(sum(xy_diff)^2)

paired_std_dev<- sqrt((sqaure_sum_diff -(square_whole_diff/n))/degree_of_freedom)
cat("The paired standard deviation",paired_std_dev)
```

```
## The paired standard deviation 11.96255
```

```r
std_error <- paired_std_dev/sqrt(n)
cat("The standard error",std_error)
```

```
## The standard error 3.317814
```

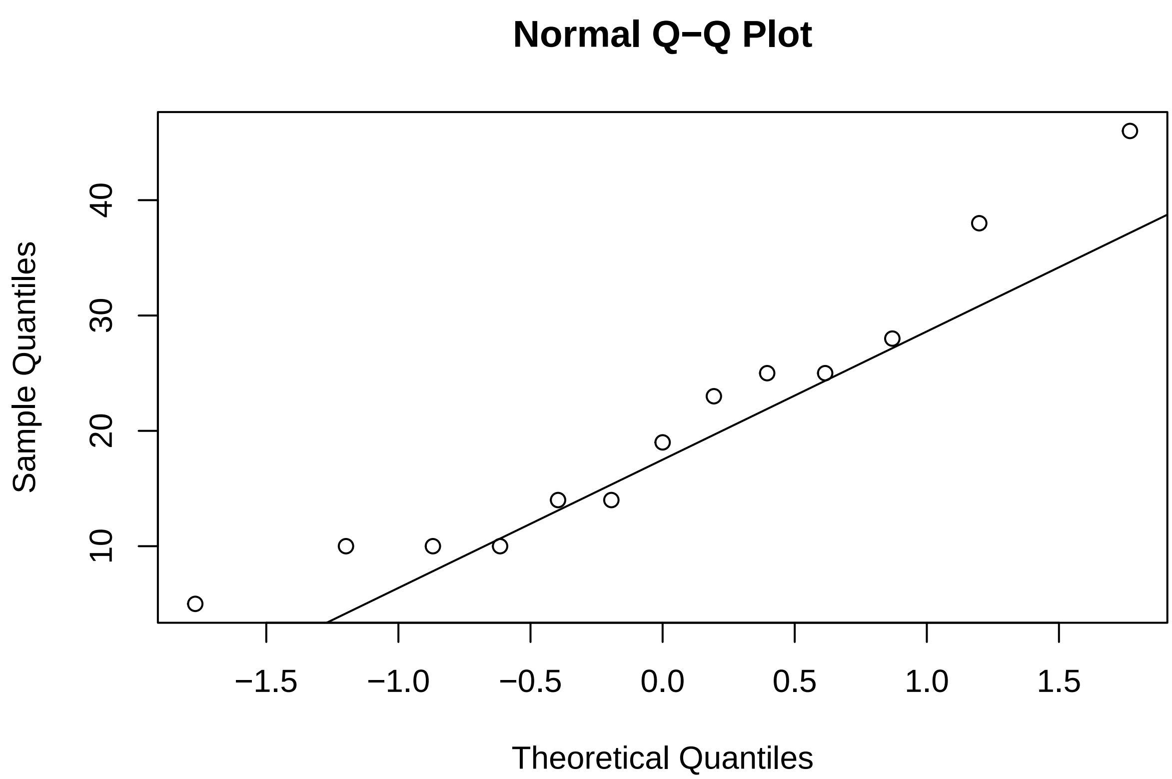
```r
t_crit<-qt(1 -(alpha/2),degree_of_freedom)
error_margin<-std_error*t_crit
cat("The margin of error",error_margin)
```

```
## The margin of error 7.228896
```

```r
lower_limit<-xy_diff_bar-error_margin
uppler_limit<-xy_diff_bar+error_margin
cat("The 95% confidence interval is given by",lower_limit,uppler_limit)
```

```
## The 95% confidence interval is given by 13.30957 27.76736
```

```r
qqnorm(xy_diff)
qqline(xy_diff)
```



(a).The normal plot is almost a straight, so its plausible that difference follows normal distribution and paired t distribution is valid.

(b).Since the xy_diff contains all the positive numbers it not plausible true mean difference is zero.

Section 9.3 Problem #40

Solution(a)

Null Hypothesis:H0:mu<=25

Alternate Hypothesis:Ha:mu>25

```
x_l<-c(1928,2549,2825,1924,1628,2175,2114,2621,1843,2541)
y_l<-c(2126,2885,2895,1942,1750,2184,2164,2626,2006,2627)
xy_diff<-c(198,336,70 ,18,122 ,9,50 ,5 ,163 ,86)
xy_diff_bar<-mean(xy_diff)
n<-10
alpha<-0.05
degree_of_freedom <-n -1
sqaure_sum_diff<-sum((xy_diff^2))
square_whole_diff<-(sum(xy_diff)^2)
paired_std_dev<- sqrt((sqaure_sum_diff -(square_whole_diff/n))/degree_of_freedom)
cat("The paired standard deviation",paired_std_dev)
```

```
## The paired standard deviation 103.845
```

```
t_testing<- ((xy_diff_bar - 25)*sqrt(n))/paired_std_dev
cat("The t-testing value",t_testing)
```

```
## The t-testing value 2.457468
```

```
p_value<-pt(t_testing,degree_of_freedom,lower.tail = FALSE)
cat("The P-value value",p_value)
```

```
## The P-value value 0.01815458
```

```
if (p_value <alpha) {
print("Reject the null hyptothesis")
} else
{
print ("Failed to reject the null hypothesis")
}
```

```
## [1] "Reject the null hyptothesis"
```

Solution(b)

```
x_l<-c(1928,2549,2825,1924,1628,2175,2114,2621,1843,2541)
y_l<-c(2126,2885,2895,1942,1750,2184,2164,2626,2006,2627)
xy_diff<-c(198,336,70 ,18,122 ,9,50 ,5 ,163 ,86)
xy_diff_bar<-mean(xy_diff)
n<-10
alpha<-0.05
degree_of_freedom <-n -1
sqaure_sum_diff<-sum((xy_diff^2))
```

```
square_whole_diff<-(sum(xy_diff)^2)
paired_std_dev<- sqrt((sqaure_sum_diff -(square_whole_diff/n))/degree_of_freedom)
t_crit<-qt(1- alpha,degree_of_freedom)
upper_side<-xy_diff_bar - t_crit*(paired_std_dev/sqrt(n))
cat("The upper bound confidence interval is given by",upper_side)
```

```
## The upper bound confidence interval is given by 45.50299
```

Solution(c)

Independent two sample t-test should only be used when we compare two separate groups.Here we would have used a two sample t-test only if different subjects were used in Lactation and Postweaning.

But here we used the same subject to obtain the measurements for the two situations.Therefore an two sample t-test would not be appropriate here.

The two sample t-tests are lot less powerful than an paired test because in a paired test we use the same participants and hence it elminates the variation between the samples.It might so happen that due to such large between the samples,we might not be able to reject the null hypothesis using the two sample t-test.

Section 9.4 Problem #50 part (a) only

Solution(a)

Null Hypothesis:H0:p1-p2=0

Alternate Hypothesis:Ha:p1!=p2

```
n1<-80
n2<-80
p1cap<-35/80
q1cap<-1 - p1cap
p2cap<- 66/80
q2cap<- 1 - p2cap
alpha<-0.01

numerator<-(p1cap - p2cap)- 0
denomiator<-sqrt(((p1cap*q1cap)/n1)+((p2cap*q2cap)/n2))
z_testing<-numerator/denomiator
cat("The z-statistic results",z_testing)
```

```
## The z-statistic results -5.546558
```

```
p_value<-2*pnorm(z_testing)
cat("The p value is given by",p_value)
```

```
## The p value is given by 2.91348e-08
```

```
if (p_value <alpha) {
print("Reject the null hyptothesis")
} else
{
print ("Failed to reject the null hypothesis")
}
```

```
## [1] "Reject the null hyptothesis"
```

Section 9.4 Problem #52

Solution

```r
n1<-395
n2<-266
p1cap<-224/n1
p2cap<-126/n2
q1cap<- 1 - p1cap
q2cap<- 1 - p2cap
alpha<-0.05
cat("Checking the condition",n1*p1cap,n1*q1cap,n2*p2cap,n2*q2cap)
```

```
## Checking the condition 224 171 126 140
```

```r
z_crit<-qnorm(1 - (alpha/2))
part1<-(p1cap*q1cap)/n1
part2<-(p2cap*q2cap)/n2
margin_of_error<-z_crit*sqrt(part1+part2)
lower_bound<-(p1cap - p2cap)-margin_of_error
upper_bound<-(p1cap - p2cap)+margin_of_error
cat("The 95% CI is given by",upper_bound,lower_bound)
```

```
## The 95% CI is given by 0.1707861 0.01602272
```

Section 9.5 Problem #64

Solution

Null Hypothesis:H0:simgasq1=simgasq2

Alternate Hypothesis:HA:sigmasq1!=sigmasq2

```r
x_energizer<-c(8.65,8.74,8.91,8.72,8.85,8.52,8.62,8.68,8.86)
y_ultracell<-c(8.76,8.81,8.81,8.70,8.73,8.76,8.68,8.64,8.79)
n<-9
simga_x<-sd(x_energizer)
simga_y<-sd(y_ultracell)
f_testing<-(simga_x^2)/(simga_y^2)
alpha<-0.05
cat("The f testing result",f_testing)
```

```
## The f testing result 4.547806
```

```r
# Calculate the p-value for a one-sided test
p_value_one_sided <- pf(f_testing, 8, 8)

# For a two-sided test, multiply the one-sided p-value by 2
p_value_two_sided <- 2 * min(p_value_one_sided, 1 - p_value_one_sided)
cat("The P value is given by",p_value_two_sided)
```

```
## The P value is given by 0.04648205
```

```r
if (p_value_two_sided <alpha) {
print("Reject the null hyptothesis")
} else
{
print ("Failed to reject the null hypothesis")
}
```

```
## [1] "Reject the null hyptothesis"
```

The energizer battery have larger variance, since its more expensive given same mean time and larger variance, hence I will not pay extra money.