

"Q1"

```
set.seed(10221)
```

```
X = rlnorm(1000,5,2.5)
```

```
a = quantile(X,probs = c(0.25,0.50,0.75))
```

```
b = qlnorm(p = c(0.25,0.50,0.75),5,2.5)
```

```
diff = a - b
```

```
print(diff)
```

```
> print(diff)
      25%      50%      75%
3.593005 10.952475 49.482344
```

Interpretation = there is very less difference between empirical quartiles and true quartiles

"Q2"

```
data = mtcars
```

```
model0 = lm(mpg ~ 1, data = data)
```

```
model1 = update(model0, ~. + wt)
```

```
model2 = update(model1, ~. + disp)
```

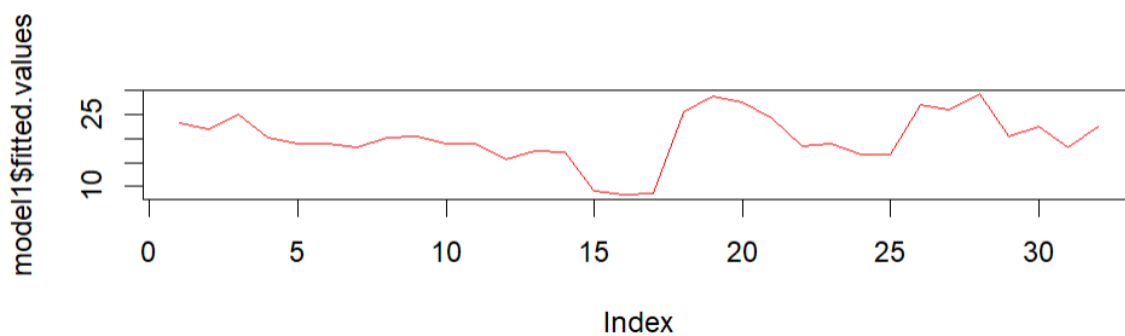
```
plot(model1$fitted.values, type = "l", col = "red")
```

```
summary(model0)
```

```
summary(model1)
```

```
summary(model2)
```

```
anova(model1, model2)
```



COMMENTS = "model0 there was no r sq as model without any independent variable
model1 r sq = 0.7528 which states it's a significant model , adj r sq = 0.7446
model2 r sq = 0.7809 that's not a significant increase at 5% los according to
anova(model1,model2) "

```
Residual standard error: 3.7616 on 30 degrees of freedom  
Multiple R-squared:  0.7528,    Adjusted R-squared:  0.7446  
F-statistic: 91.38 on 1 and 30 DF,  p-value: 1.294e-10
```

```
Multiple R-squared:  0.7809,    Adjusted R-squared:  0.7658
```

"Q3"

```
set.seed(2919)
```

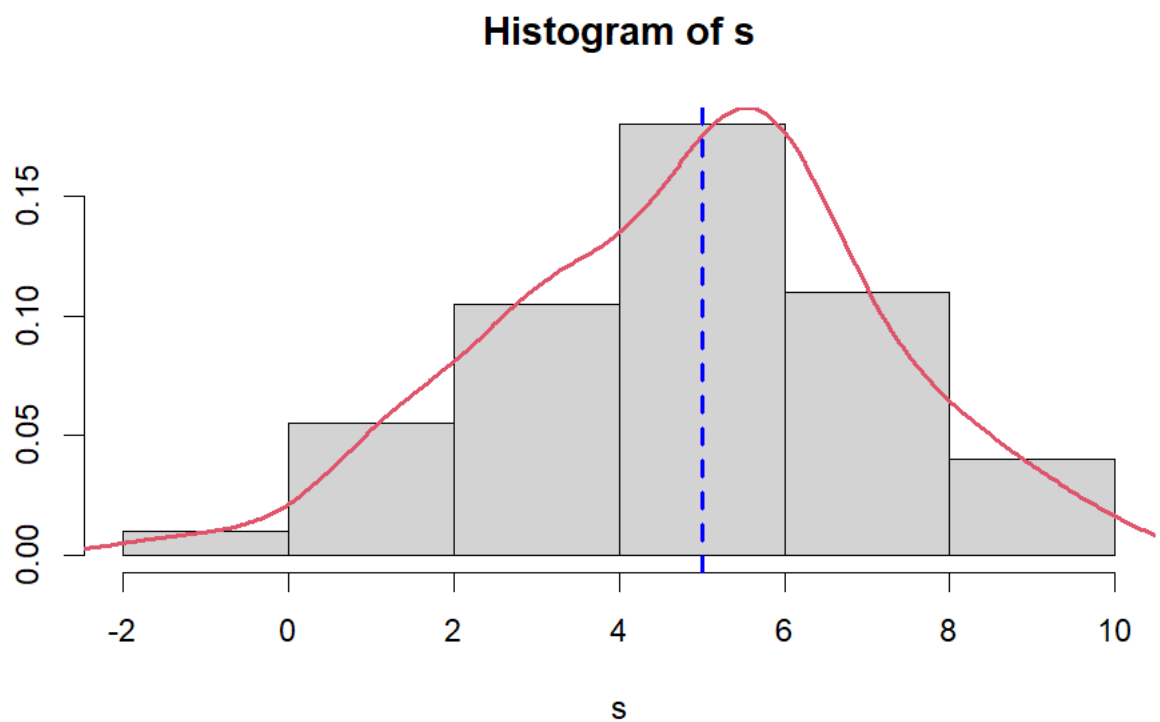
```
s = rnorm(100,5,2.5)
```

```
x = density(s)
```

```
hist(s, freq = FALSE)
```

```
lines(x$x, x$y, col = 2, lwd = 2)
```

```
abline(v = 5,col = "blue" , lwd = 2, lty = 2 )
```



“We can reduce difference between both by increasing the sample size as larger the sample size more approximately it follows normal distribution”

"Q4"

```
m = matrix(c(0,739280,  
             1,185309,  
             2,23349,  
             3,1937,  
             4,114,  
             5,1,  
             6,3),ncol = 2,byrow = T)
```

```
colnames(m) = c("X","Frequency")
```

```
chisq.test(m)
```

```
Pearson's Chi-squared test  
data:  m  
X-squared = 375789, df = 6, p-value < 2.2e-16
```

“As p value is less than 1% therefore this distribution comes from poisson distribution”

"Q5"

"a"

```
library(MASS)
```

```
data = Animals
```

```
cor(data)
```

```
      body      brain  
body 1.000000000 -0.005341163  
brain -0.005341163 1.000000000
```

Comment = "there is a weak negative correlation "

"b"

```
cor(log(data))
```

```
      body      brain  
body 1.0000000 0.7794935  
brain 0.7794935 1.0000000
```

Comment = "There is a strong positive correlation"

"d"

```
t.test(data$body,data$brain,alternative = "greater")
```

```
data: data$body and data$brain  
t = 1.1854, df = 27.354, p-value = 0.123  
alternative hypothesis: true difference in means is greater than 0  
95 percent confidence interval:  
-1615.936      Inf
```

Comment = "the mean Body weight is greater than mean Brain weight."

"e"

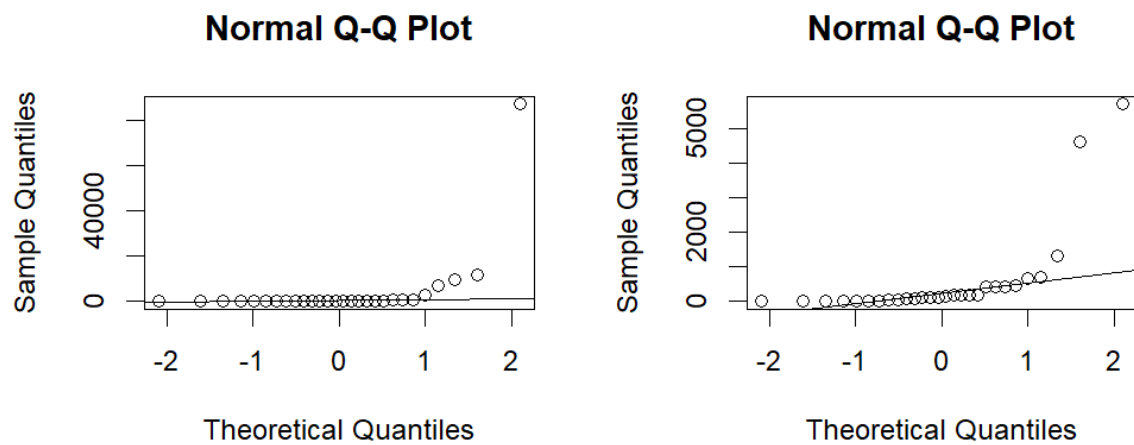
```
par(mfrow = c(1,2))
```

```
qqnorm(data$body)
```

```
qqnorm(data$brain)
```

```
qqline(data$body)
```

```
qqline(data$brain)
```



Comment = “as we can see that qqline is almost covering all the points of body weight therefore we can conclude that body weight follows normal more than brain weight”

"Q6"

```
m = matrix(c(40,20,10,30,35,15,30,45,25),nrow = 3,byrow = T)
```

```
colnames(m) = c("A","B","C")
```

```
rownames(m) = c("SSC","GRADUATE","PG")
```

```
cor(m)
```

	A	B	C
A	1.0000000	-0.9176629	-0.7559289
B	-0.9176629	1.0000000	0.9538210
C	-0.7559289	0.9538210	1.0000000

Comment = “we can see it is dependent on each other”

"Q7"

```
spam = read.csv(file.choose())
```

```
spam$Spam = as.factor(spam$Spam)
```

```
model = glm(Spam ~ ., data = spam, family = binomial())
```

```
summary(model)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-5.3006	-0.6469	0.0000	0.7918	3.9864

Coefficients:

	Estimate	Std. Error	z	value	Pr(> z)	
(Intercept)	0.332279	0.050606	6.566	5.17e-11	***	
the	-0.164586	0.018233	-9.027	< 2e-16	***	
to	-0.024602	0.019595	-1.255	0.209300		
and	0.219322	0.024289	9.030	< 2e-16	***	
of	0.181715	0.027009	6.728	1.72e-11	***	
a	0.162628	0.027082	6.005	1.91e-09	***	
in.	0.056937	0.030434	1.871	0.061369	.	
for.	-0.101996	0.028459	-3.584	0.000338	***	
you	0.173672	0.023386	7.426	1.12e-13	***	
is	-0.074714	0.033662	-2.220	0.026449	*	
on	-0.393930	0.037303	-10.560	< 2e-16	***	
this	0.213083	0.035686	5.971	2.36e-09	***	
that	-0.135693	0.040352	-3.363	0.000772	***	
s	-0.003784	0.041093	-0.092	0.926640		
enron	-21.098030	243.275658	-0.087	0.930890		
be	-0.182656	0.046072	-3.965	7.35e-05	***	
with	-0.020856	0.048726	-0.428	0.668628		
i	-0.170753	0.025139	-6.792	1.10e-11	***	
your	0.522094	0.038343	13.616	< 2e-16	***	
have	-0.293822	0.048427	-6.067	1.30e-09	***	
as	0.068130	0.044897	1.517	0.129148		
will	-0.332093	0.044793	-7.414	1.23e-13	***	
it	0.169437	0.047200	3.590	0.000331	***	
we	-0.144296	0.037909	-3.806	0.000141	***	
are	0.067673	0.046464	1.456	0.145262		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1