Критерии для проверки данных на соответствие нормальному закону распределения (реализация в R)

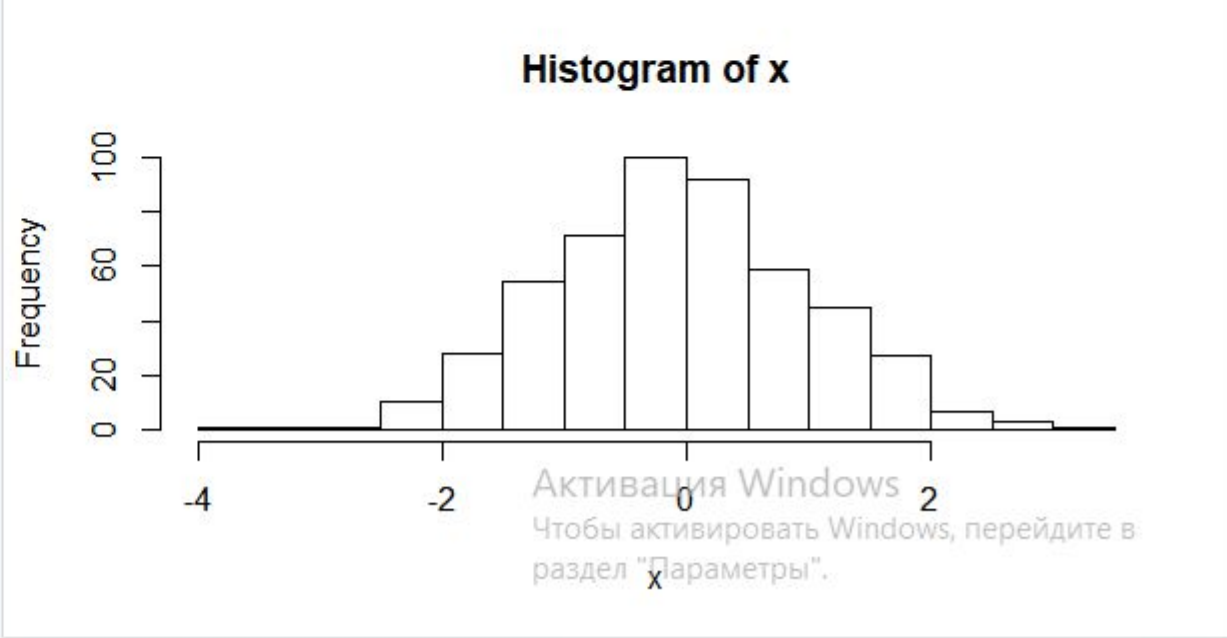
Подготовила: Лайзан А.В., гр 4607

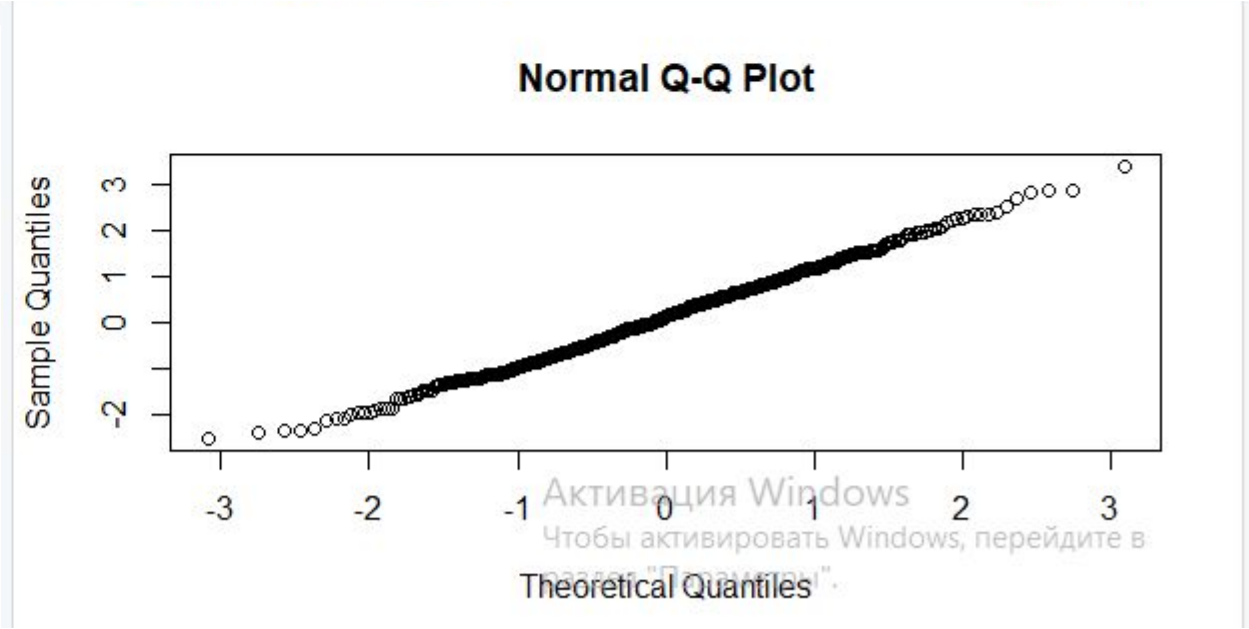
```r
x<-rnorm(500)
qqnorm(x)
```

Создание вектора с генерированной выборкой 500 объектов

Построение графика квантилей

```r
hist(x)
```





*для **hist()** важен шаг, для **qqnorm()** важно n. Поэтому данные функции только для визуализации, не для проверки

```r
shapiro.test(x)
shapiro.test(rnorm(500))
```

```
> shapiro.test(x)

        Shapiro-Wilk normality test

data:  x
W = 0.99687, p-value = 0.4516

> shapiro.test(rnorm(500))

        Shapiro-Wilk normality test

data:  rnorm(500)
W = 0.9974, p-value = 0.6246
```

```r
install.packages("nortest")
library(nortest)

#p-level>0.05 => H0
lillie.test(x)
lillie.test(rnorm(500))
```

Установка пакета и подключение библиотеки

```
> lillie.test(x)

        Lilliefors (Kolmogorov-Smirnov) normality test

data:  x
D = 0.026942, p-value = 0.5084

> lillie.test(rnorm(500))

        Lilliefors (Kolmogorov-Smirnov) normality test

data:  rnorm(500)
D = 0.029619, p-value = 0.3559
```

```r
pearson.test(x)
pearson.test(rnorm(500))
```

```
> pearson.test(x)

        Pearson chi-square normality test

data:  x
P = 17.8, p-value = 0.7178

> pearson.test(rnorm(500))

        Pearson chi-square normality test

data:  rnorm(500)
P = 22.6, p-value = 0.4246
```

```r
t<-runif(50, min =2, max=4)
shapiro.test(t)
lillie.test(t)
pearson.test(t)
```

runif генерирует выборку n элементов в диапазоне [min,max]

```
> t<-runif(50, min =2, max=4)
> shapiro.test(t)

        Shapiro-Wilk normality test

data:  t
W = 0.92557, p-value = 0.003771

> lillie.test(t)

        Lilliefors (Kolmogorov-Smirnov) normality test

data:  t
D = 0.13431, p-value = 0.02463

> pearson.test(t)

        Pearson chi-square normality test

data:  t
P = 14, p-value = 0.05118
```

Другие тесты

- ad.test() – тест Андерсона-Дарлинга
- cvm.test() – Крамера фон Мизеса
- sf.test() – тест Шапиро-Франсия

Литература

https://star.inp.nsk.su/~baldin/DataAnalysis/R/R-05-2var.pdf

https://r-analytics.blogspot.com/2012/06/blog-post_14.html