Зад.1 Генерирайте n на брой наблюдения над сл.в $X \in N(2, 4)$. Работете с $n = 10$, 30, 100. Проверете хипотезата:

б) $H_0 \ : \ EX = 3$

Ще разгледаме случая $n = 30$. Проверяваме хипотезата:

$H_0 \ : \ EX = 3$

$H_1 \ : \ EX \neq 3$

Данните са нормално разпределени затова използваме стандартния т-тест:

```
> x = rnorm(30, 2, 2)
> t.test( x, mu = 3, alternative = 'two.sided')
```

```
        One Sample t-test

data:  x
t = -2.5184, df = 29, p-value = 0.01756
alternative hypothesis: true mean is not equal to 3
95 percent confidence interval:
 1.216045 2.815041
sample estimates:
mean of x
 2.015543
```

Извод: Получената стойност за p-val е по-малка от 5%. Имаме основание да отхвърлим хипотезата H_0 и да приемем алтернативата H_1, т.е. $EX \neq 3$.

Зад.2 Данните 'vacation' от пакета 'UsingR' съдържат броя на платените почивни дни за работещите в текстилната индустрия. Проверете хипотезата, че те почиват 24 дни.

Проверяваме хипотезата:

$H_0 \ : \ \mu = 24$

$H_1 \ : \ \mu \neq 24$

Важно е да знаем как са разпределени данните. Извършваме проверка за нормалност с графични методи.

```
> boxplot( vacation, horizontal = T )
> hist( vacation )
> qqnorm( vacation )
> qqline( vacation )
```

Боксплота, хистограмата и Q-Q плота, показват по-скоро нормално разпределение.

Можем да изпълним и тест за нормалност на Шапиро - Уилк.

```
> shapiro.test( vacation )
```

```
        Shapiro-Wilk normality test

data: vacation
W = 0.95272, p-value = 0.1374
```

При p-val по-голямо от 5% приемаме, че данните са нормално разпределени. В такъв случай за проверка на първоначалната хипотеза може да приложим т-тест.

> t.test(vacation, mu = 24, alternative = 'two.sided')

```
        One Sample t-test

data: vacation
t = -2.2584, df = 34, p-value = 0.03045
...
```

Извод: Малката стойност на p-val ни дава право да отхвърлим хипотезата, работещите не почиват 24 дни.

Зад.3 Компания твърди, че 50% от клиентите са доволни от нейните услуги. Направено е проучване при 100 лица, 42 от които казват че са доволни. Може ли да се приеме твърдението на компанията.

Всеки клиент е отговарял с "Да" или "Не" на въпроса, следователно става дума за проверка на пропорция. Би следвало да приемем твърдението на компанията за вярно, ако 50% или повече са доволни и да отхвърлим, ако броят на доволните е съществено по-малък, т.е. проверяваме хипотеза:

$H_0 \ : \ p = 0.5$

$H_1 \ : \ p < 0.5$

за вероятността p клиента да е доволен.

> prop.test(42, 100, p =0.5, alternative = 'less')

```
        1-sample proportions test with continuity correction

data: 42 out of 100, null probability 0.5
X-squared = 2.25, df = 1, p-value = 0.06681
alternative hypothesis: true p is less than 0.5
...
```

Извод: При p-value = 0.067 нямаме основание да отхвърляме хипотезата $\Longrightarrow$ $H_0 : p = 0.5$ Възможно е 50% от клиентите да са доволни от услугите и това, че по малко са отговорили с "Да" да е плот на случайността.

Зад.4 Проверяваме хипотеза клиентите говорят 5 минути по телефона, срещу алтернатива говорят повече. Направените наблюдения са

12.8 3.5 2.9 9.4 8.7 0.7 0.2 2.8 1.9 2.8 3.1 15.8

Можем ли да отхвърлим хипотезата?

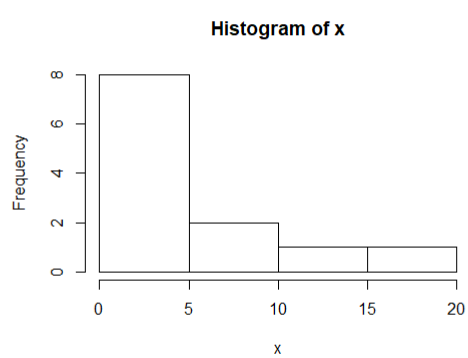
Проверяваме

$H_0 \ : \ \mu = 5$

$H_1 \ : \ \mu > 5$

Лесно се вижда, че данните не са нормално разпределени. Например:

> hist(x)



> shapiro.test(x)

```
        Shapiro-Wilk normality test

data: x
W = 0.83988, p-value = 0.0276
```

Поради малката стойност на p-val, отхвърляме възможносттта данните да са нормални. За проверка на хипотезата ще трябва да използваме теста на Уилкоксън.

> wilcox.test(x, mu = 5, alternative = 'greater')

```
        Wilcoxon signed rank test with continuity correction

data: x
V = 39, p-value = 0.5156
alternative hypothesis: true location is greater than 5
```

Извод: Приемаме хипотезата, клиентите найстина говорят 5 минути