7 Testowanie hipotez statystycznych

7.1 Przykłady

Test t-Studenta dla jednej próby

Przykład. Automat produkuje tabliczki czekolady o nominalnej wadze $250g$. Podczas kontroli technicznej pobrano 16-elementową próbę tabliczek czekolady otrzymując wyniki:

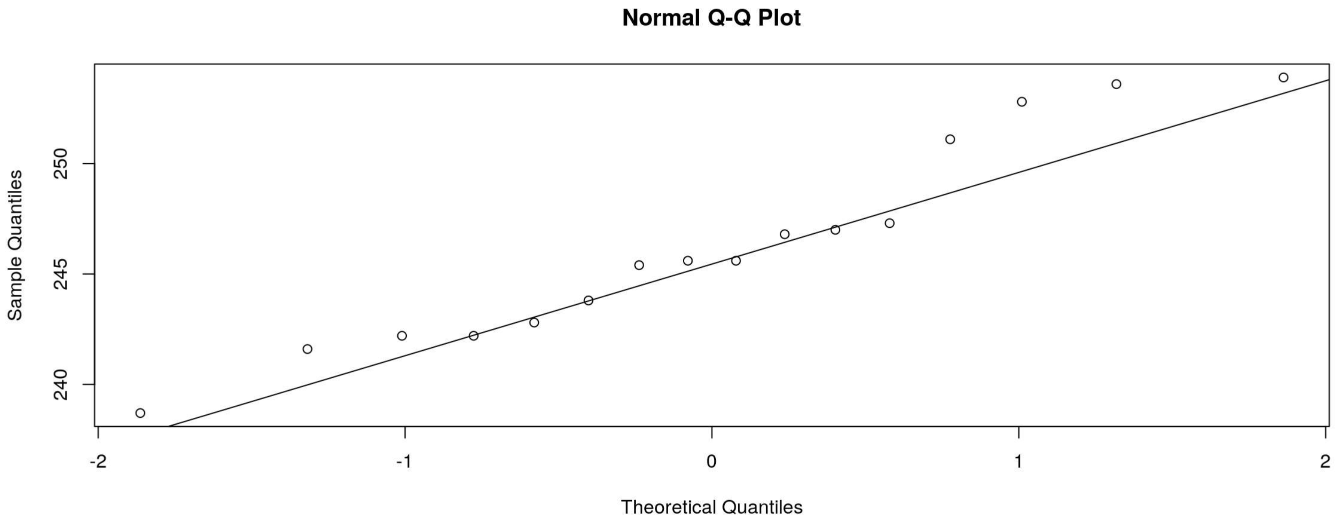
```
242.2 243.8 252.8 245.4 245.6 253.6 247.3 238.7 241.6 242.8 251.1 246.8 247 245.6 242.2
```

Na poziomie istotności $\alpha = 0{,}05$ zweryfikuj hipotezę, że automat rozlegulował się i produkuje tabliczki czekolady o istotnie różnej wadze od nominalnej wagi.

```r
x <- c(242.2, 243.8, 252.8, 245.4, 245.6, 253.6, 247.3, 238.7,
       241.6, 242.8, 251.1, 246.8, 247.0, 245.6, 242.2, 253.9)
shapiro.test(x)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  x
## W = 0.93622, p-value = 0.3052
```

```r
qqnorm(x)
qqline(x)
```




```r
mean(x)
```

```
## [1] 246.275
```

```r
t.test(x, mu = 250, alternative = "less")
```

```
##
##  One Sample t-test
##
## data:  x
## t = -3.2679, df = 15, p-value = 0.002595
## alternative hypothesis: true mean is less than 250
## 95 percent confidence interval:
##      -Inf 248.2732
## sample estimates:
## mean of x
##   246.275
```

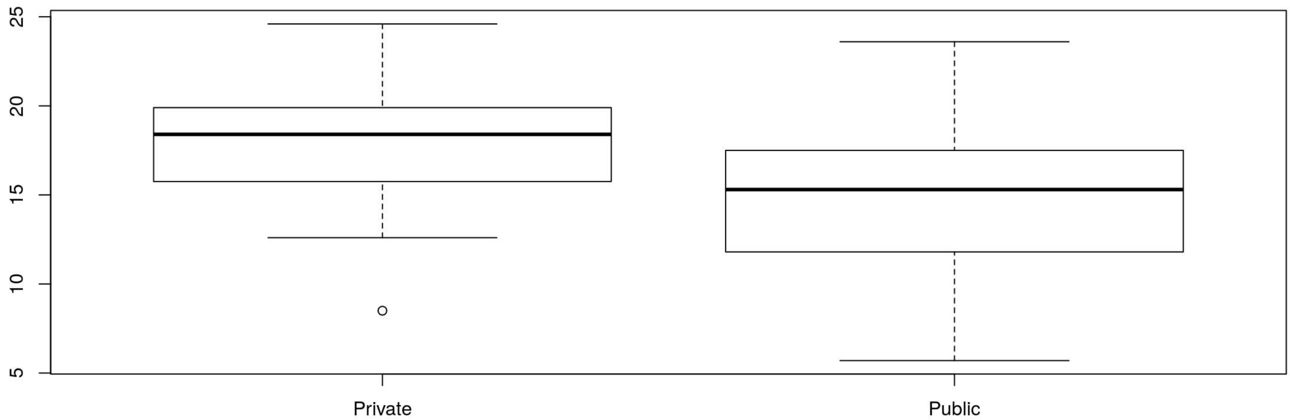
Test t-Studenta dla dwóch prób niezależnych i test F-Snedecora dla wariancji w dwóch próbach niezależnych

Przykład. Zbiór danych `homework` z pakietu `UsingR` zawiera informacje o ilości czasu poświęconego na odrabianie pracy domowej przez uczniów szkół prywatnych i publicznych. Naszym celem jest sprawdzenie, czy uczniowie obu typów szkół spędzają tyle samo czasu na odrabianiu zadań domowych.

```r
library(UsingR)
head(homework)
```

```
##   Private Public
## 1    21.3   15.3
## 2    16.8   17.4
## 3     8.5   12.3
## 4    12.6   10.7
## 5    15.8   16.4
## 6    19.3   11.3
```

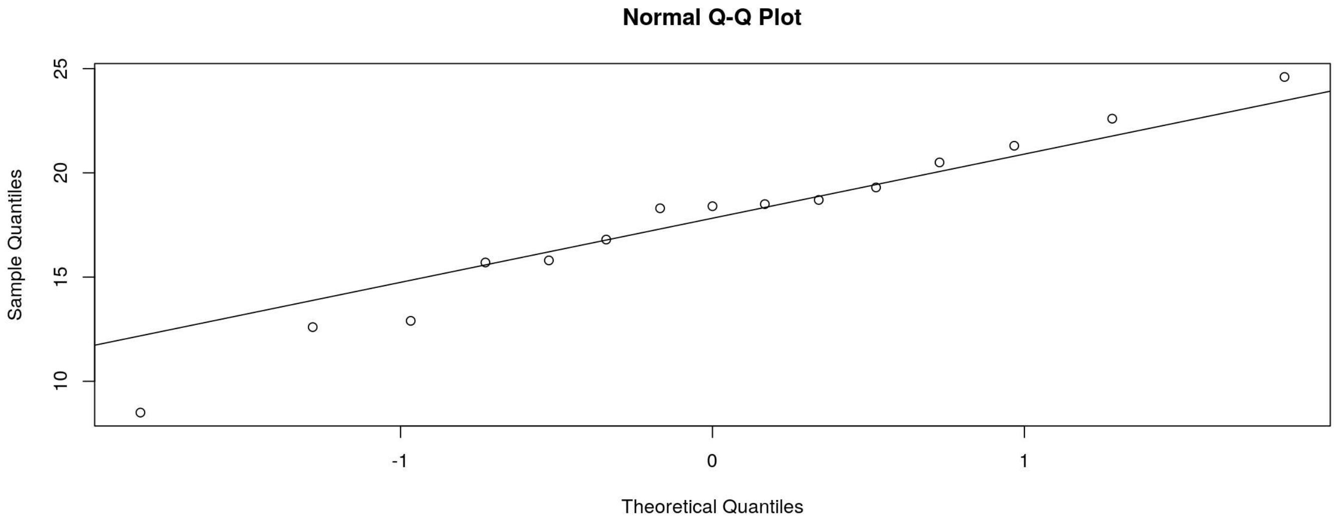
```r
boxplot(homework)
```




```r
shapiro.test(homework$Private)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  homework$Private
## W = 0.97017, p-value = 0.8606
```

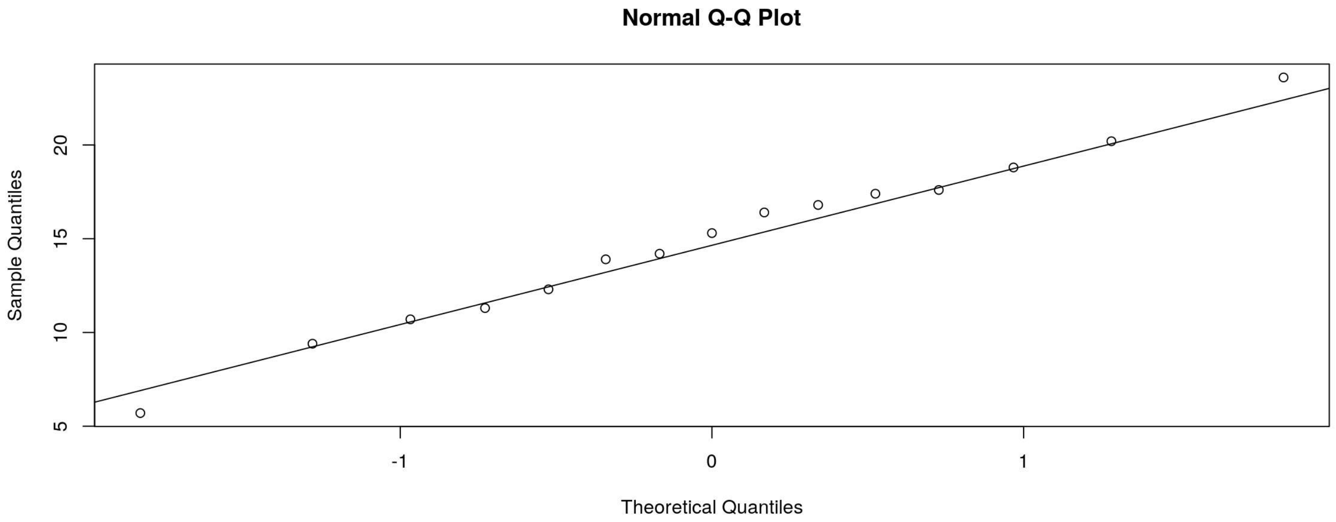
```r
qqnorm(homework$Private)
qqline(homework$Private)
```




```r
shapiro.test(homework$Public)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  homework$Public
## W = 0.99275, p-value = 0.9999
```

```r
qqnorm(homework$Public)
qqline(homework$Public)
```




```r
var(homework$Private)
```

```
## [1] 17.1081
```

```r
var(homework$Public)
```

```
## [1] 20.87781
```

```r
var.test(homework$Private, homework$Public, alternative = "less")
```

```
##
##  F test to compare two variances
##
## data:  homework$Private and homework$Public
## F = 0.81944, num df = 14, denom df = 14, p-value = 0.3573
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 2.035262
## sample estimates:
## ratio of variances
##          0.8194392
```

```r
mean(homework$Private)
```

```
## [1] 17.63333
```

```r
mean(homework$Public)
```

```
## [1] 14.90667
```

```r
t.test(homework$Private, homework$Public,
       var.equal = TRUE, alternative = 'greater')
```

```
##
##  Two Sample t-test
##
## data:  homework$Private and homework$Public
## t = 1.7134, df = 28, p-value = 0.04884
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.01957252        Inf
## sample estimates:
## mean of x mean of y
##  17.63333  14.90667
```

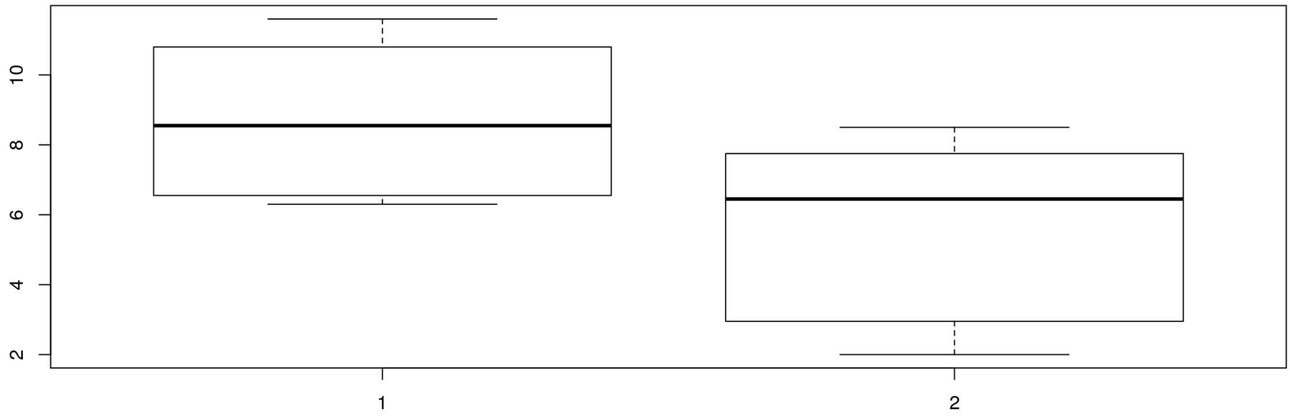
Test t-Studenta dla prób zależnych

Przykład. Badano wpływ hipnozy na redukcję bólu. Notowano poziom odczuwalnego bólu:

- przed hipnozą: 6.6, 6.5, 9.0, 10.3, 11.3, 8.1, 6.3, 11.6,
- po hipnozie: 6.8, 2.5, 7.4, 8.5, 8.1, 6.1, 3.4, 2.0.

Czy na poziomie istotności $0{,}05$ możemy stwierdzić, że hipnoza redukuje poziom odczuwalnego bólu?

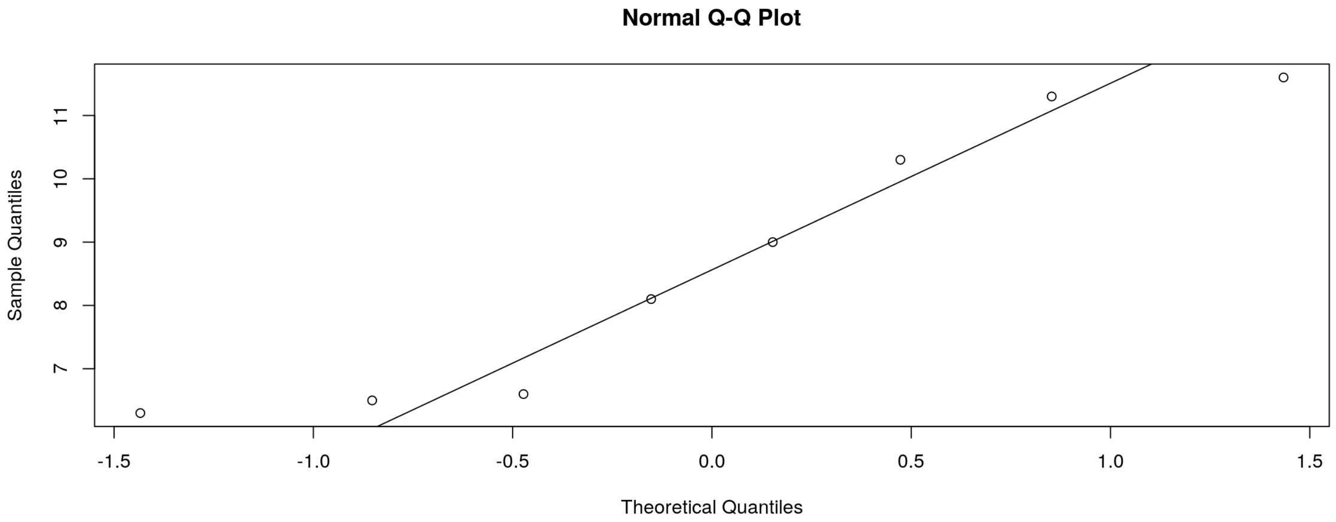
```r
a <- c(6.6, 6.5, 9.0, 10.3, 11.3, 8.1, 6.3, 11.6)
b <- c(6.8, 2.5, 7.4, 8.5, 8.1, 6.1, 3.4, 2.0)
boxplot(a, b)
```




```r
shapiro.test(a)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  a
## W = 0.88638, p-value = 0.2165
```

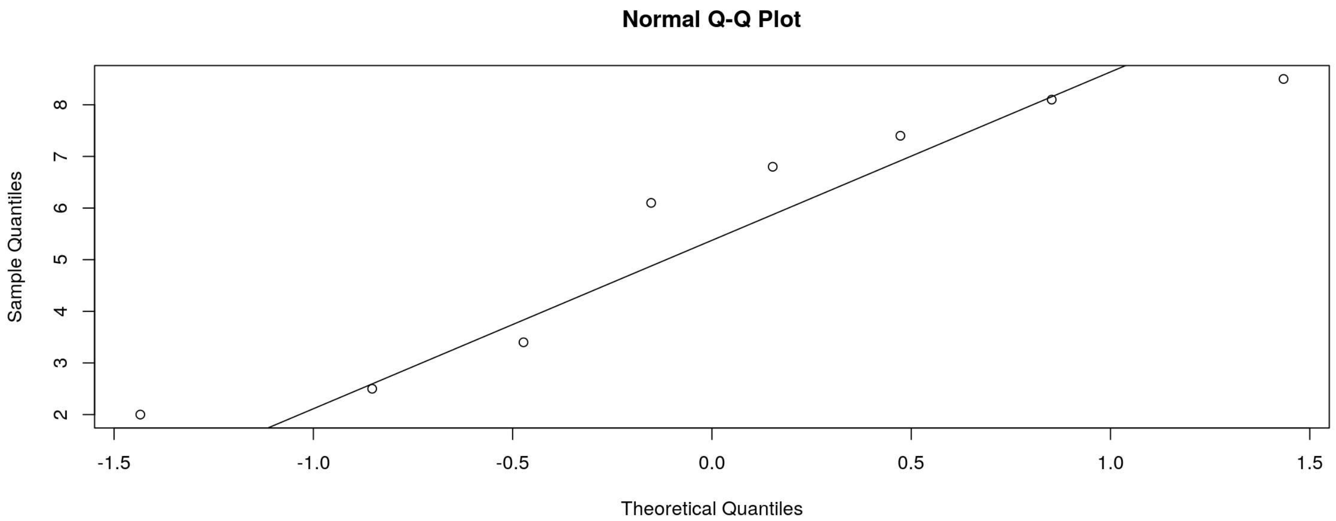
```r
qqnorm(a)
qqline(a)
```




```r
shapiro.test(b)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  b
## W = 0.88356, p-value = 0.2036
```

```r
qqnorm(b)
qqline(b)
```




```r
mean(a)
```

```
## [1] 8.7125
```

```r
mean(b)
```

```
## [1] 5.6
```

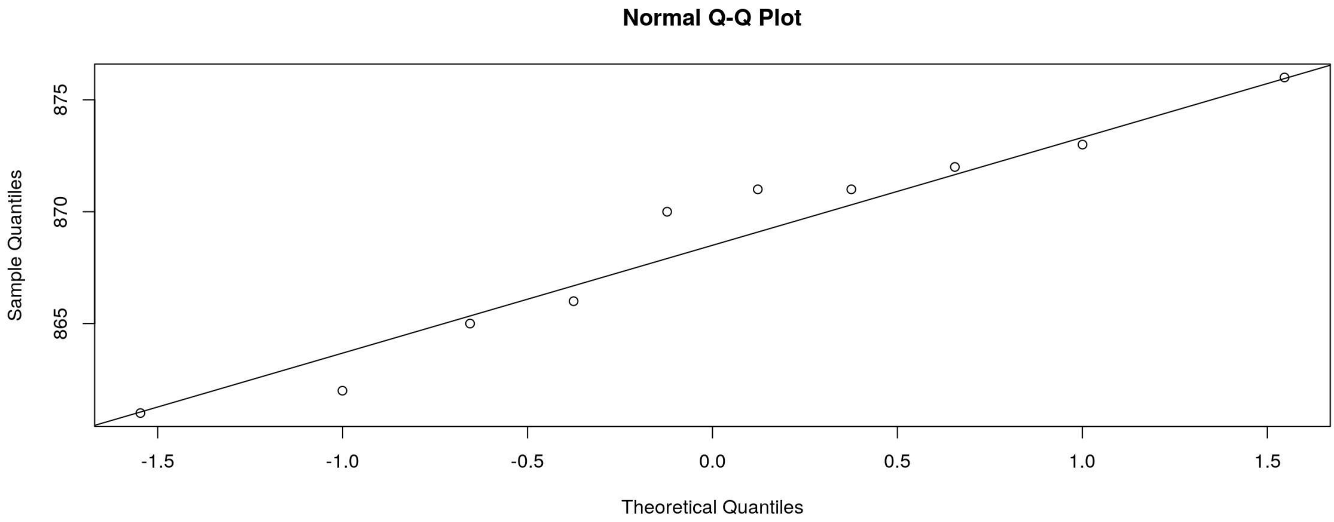
```r
t.test(a, b, alternative = 'greater', paired = TRUE)
```

```
## 
##  Paired t-test
## 
## data:  a and b
## t = 3.0285, df = 7, p-value = 0.009577
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  1.165386      Inf
## sample estimates:
## mean of the differences 
##                  3.1125
```

7.2 Zadania

Zadanie 1. W pewnym regionie wykonano dziesięć niezależnych pomiarów głębokości morza i uzyskano następujące wyniki: 862, 870, 876, 866, 871, 865, 861, 873, 871, 872. Na poziomie istotności $\alpha = 0{,}05$ zweryfikuj hipotezę, że średnia głębokość morza w tym regionie wynosi $870m$.

```
## [1] 0.545861
```




```
## [1] 868.7
```

```
## [1] 0.2136555
```

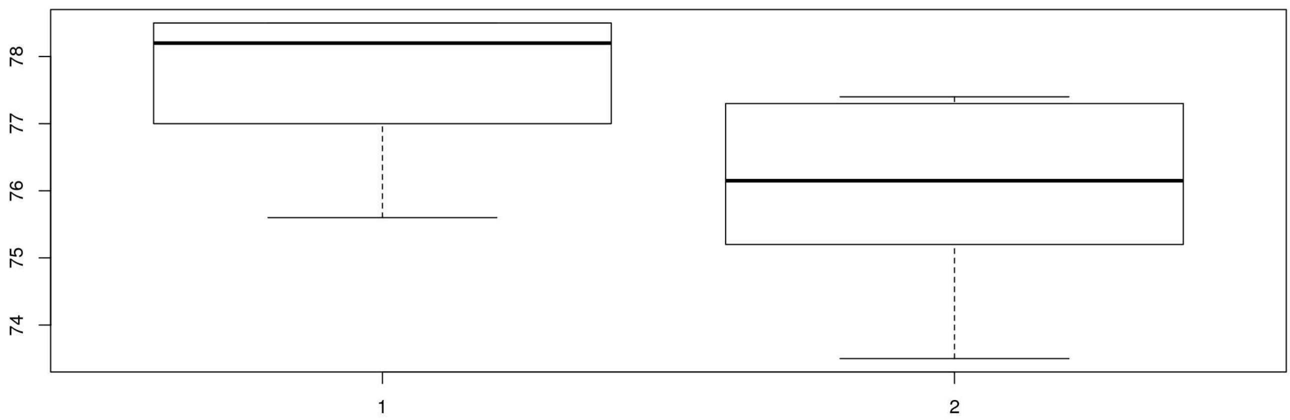
Zadanie 2. Producent proszku do prania A twierdzi, że jego produkt jest znacznie lepszy niż konkurencyjny proszek B. Aby zweryfikować to zapewnienie, CTA (Consumer Test Agency) przetestowało oba proszki do prania. W tym celu przeprowadzono pomiary stopnia wyprania 7 kawałków tkaniny z proszkiem A i uzyskano wyniki (w $\%$):

$$78{,}2; 78{,}5; 75{,}6; 78{,}5; 78{,}5; 77{,}4; 76{,}6,$$

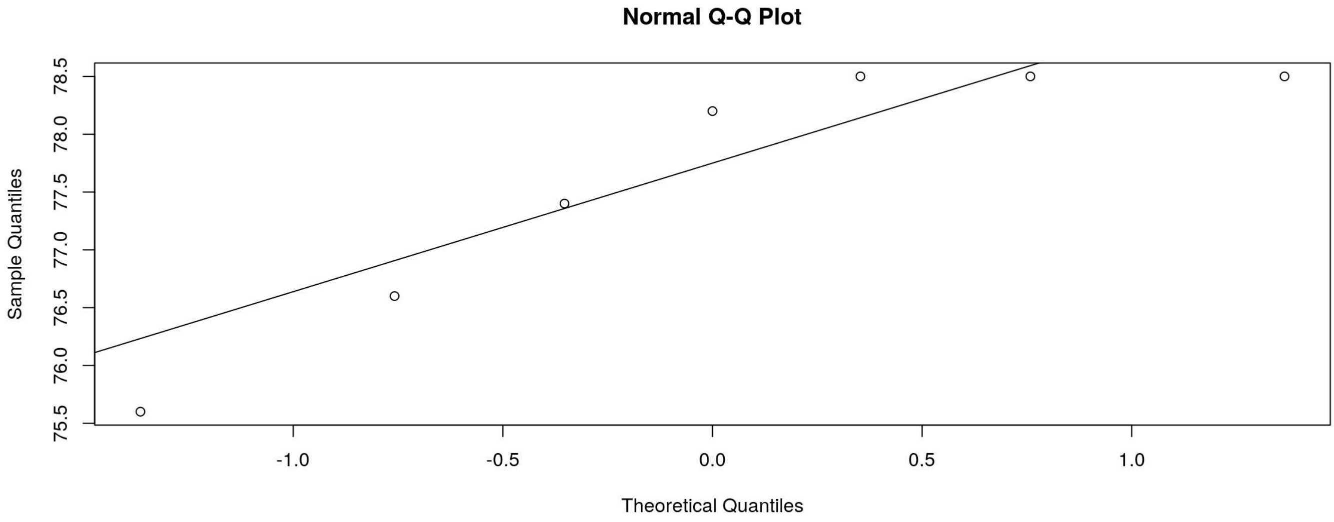
i 10 kawałków tkaniny z proszkiem B otrzymując wyniki (w $\%$):

$$76{,}1; 75{,}2; 75{,}8; 77{,}3; 77{,}3; 77{,}0; 74{,}4; 76{,}2; 73{,}5; 77{,}4.$$

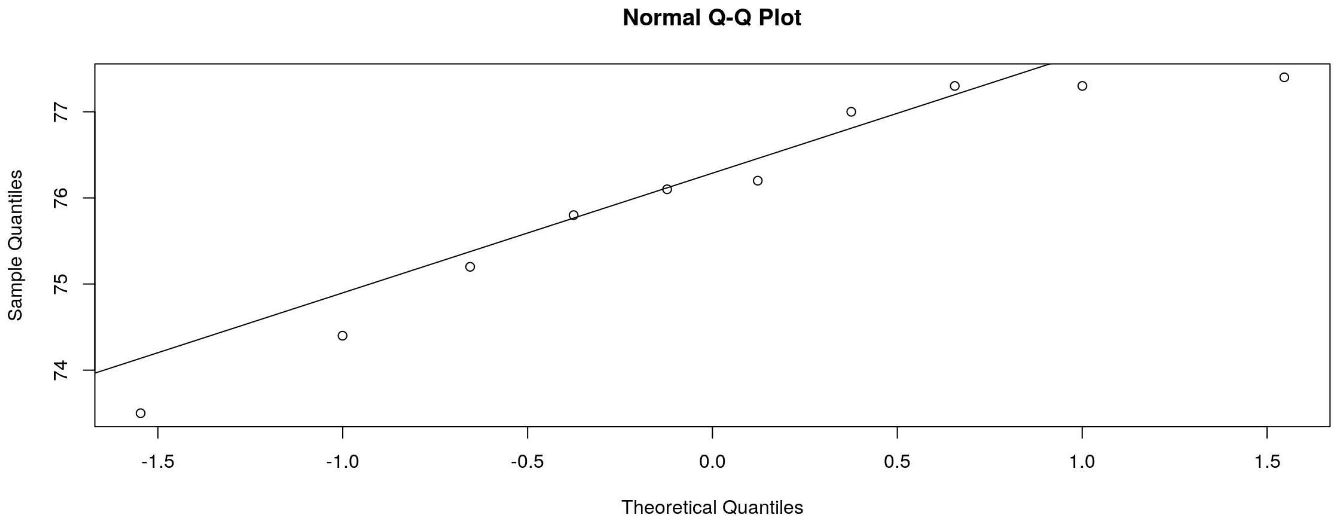
Jaki powinien być wniosek CTA na temat jakości tych proszków?



[1] 0.06832755



[1] 0.2558752



[1] 1.304762

[1] 1.764

[1] 0.3683809

[1] 77.61429

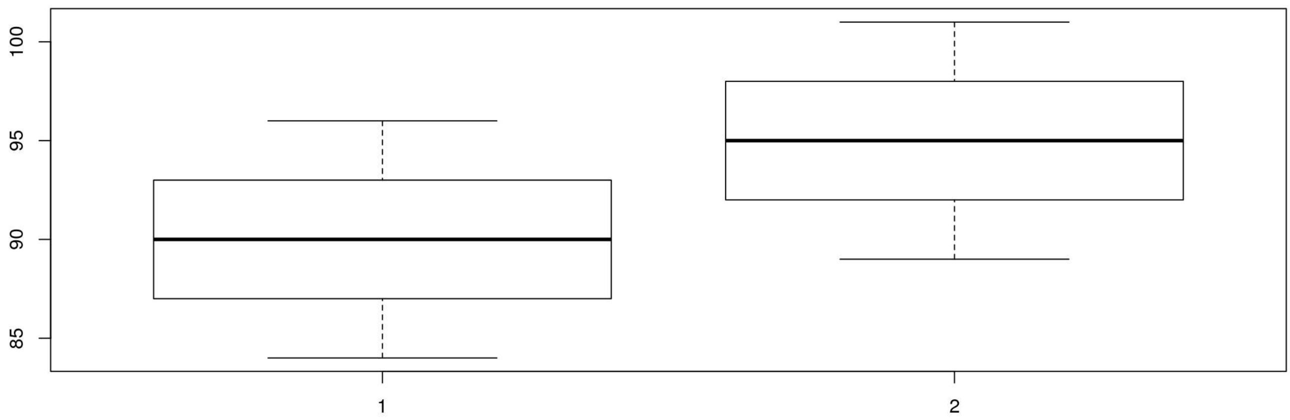
```
## [1] 76.02
```

```
## [1] 0.01059375
```

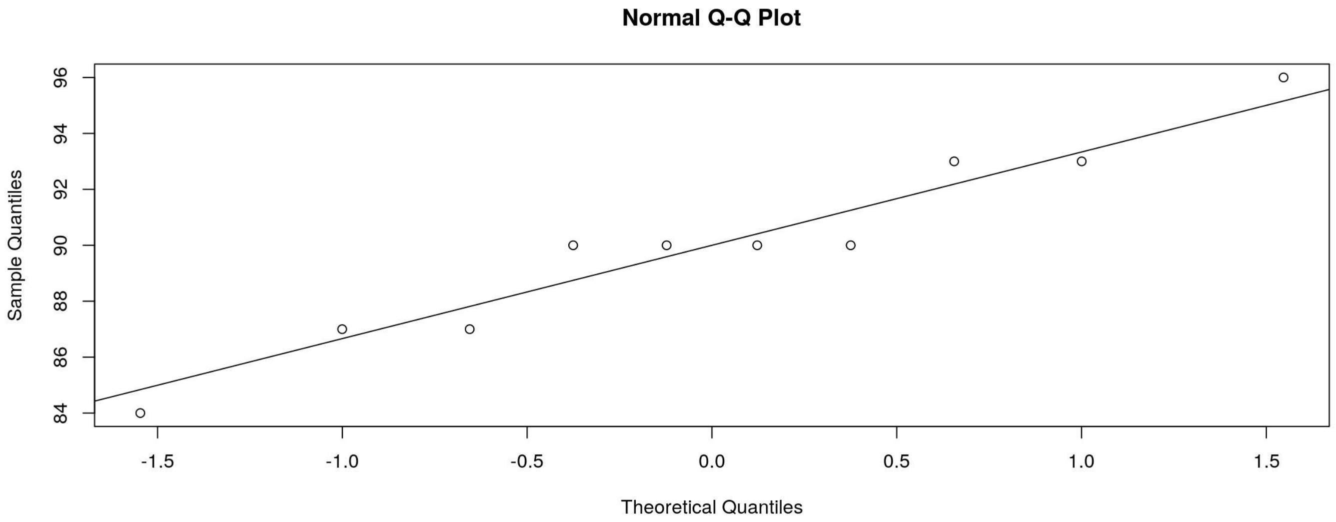
Zadanie 3. Grupa 10 osób została poddana badaniu mającym na celu zbadanie stosunku do szkół publicznych. Następnie grupa obejrzała film edukacyjny mający na celu poprawę podejścia do tego typu szkół. Wyniki są następujące (wyższa wartość oznacza lepsze podejście):

- przed: 84, 87, 87, 90, 90, 90, 90, 93, 93, 96,
- po: 89, 92, 98, 95, 95, 92, 95, 92, 98, 101.

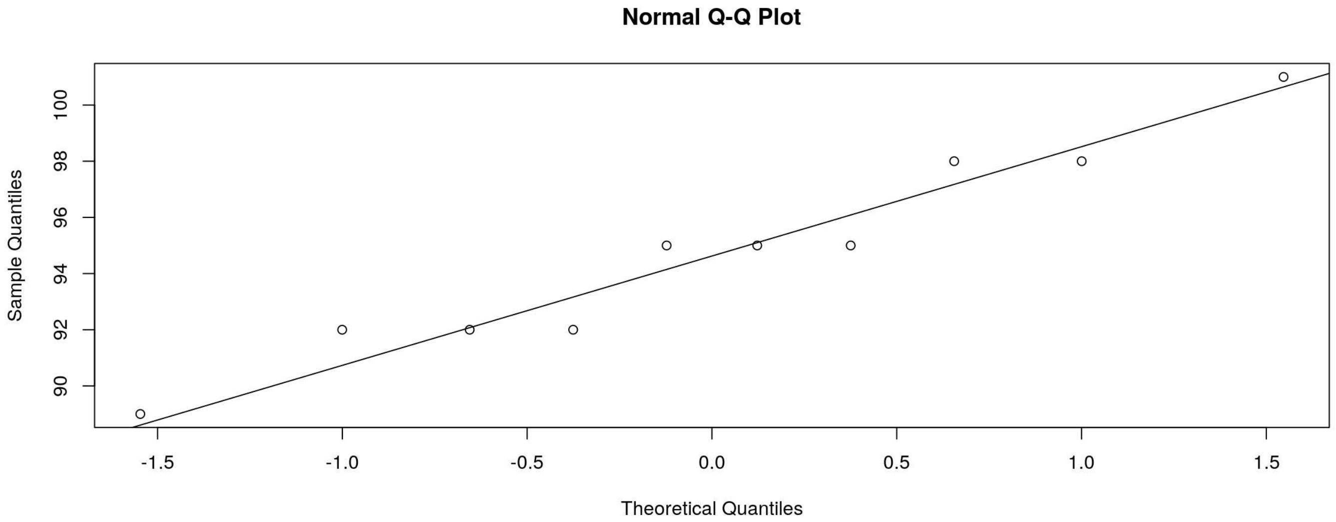
Zweryfikuj, czy film znacznie poprawia stosunek do szkół publicznych.




```
## [1] 0.7025892
```




```
## [1] 0.691489
```




```
## [1] 90
```

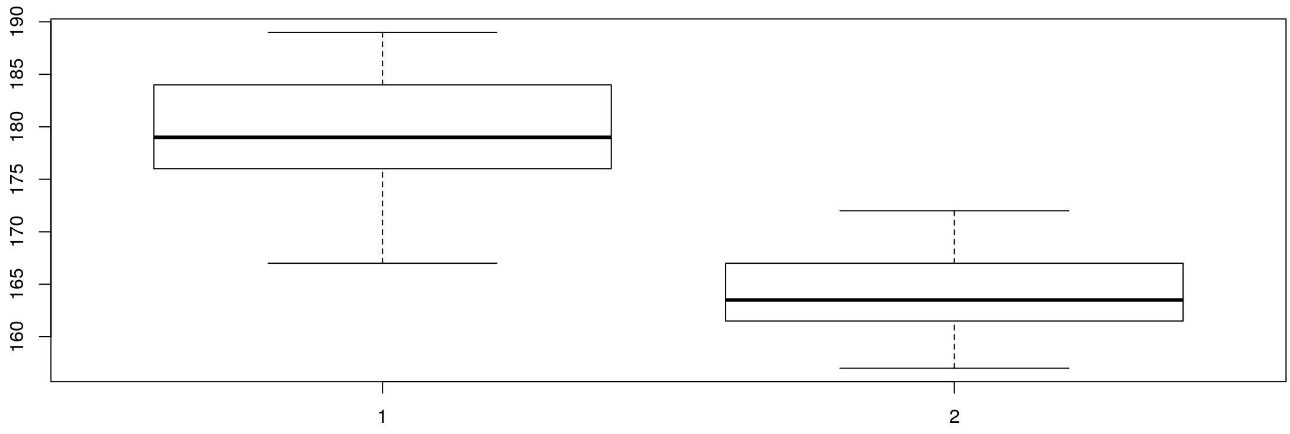
```
## [1] 94.7
```

```
## [1] 0.0003786878
```

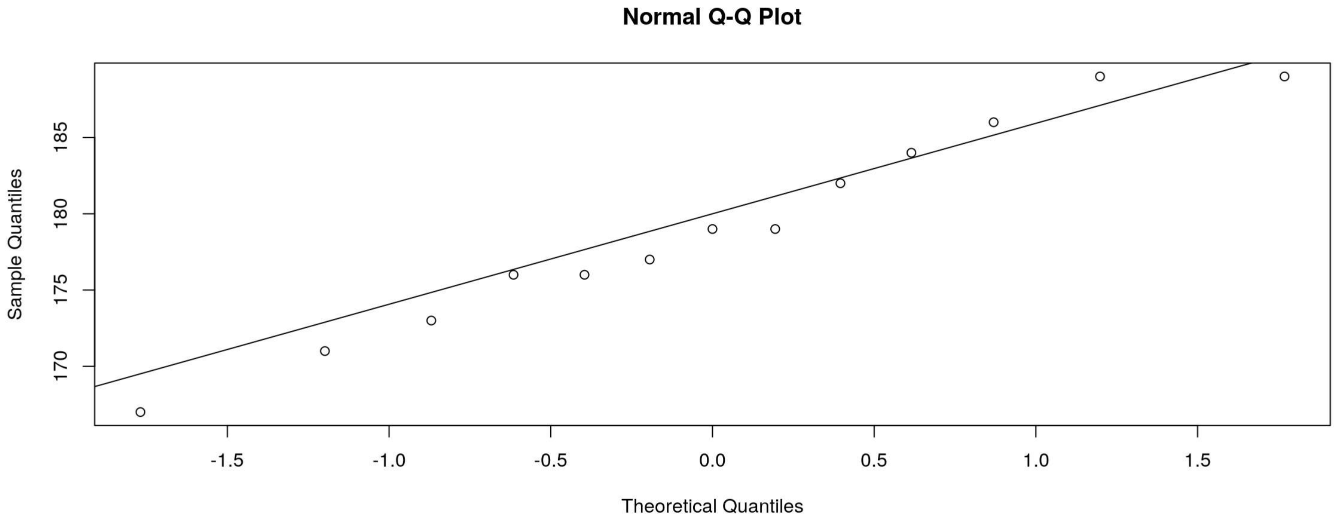
Zadanie 4. Zbadano wzrost 13 mężczyzn i 12 kobiet w pewnym centrum sportowym. Wyniki są następujące:

- mężczyźni: 171, 176, 179, 189, 176, 182, 173, 179, 184, 186, 189, 167, 177,
- kobiety: 161, 162, 163, 162, 166, 164, 168, 165, 168, 157, 161, 172.

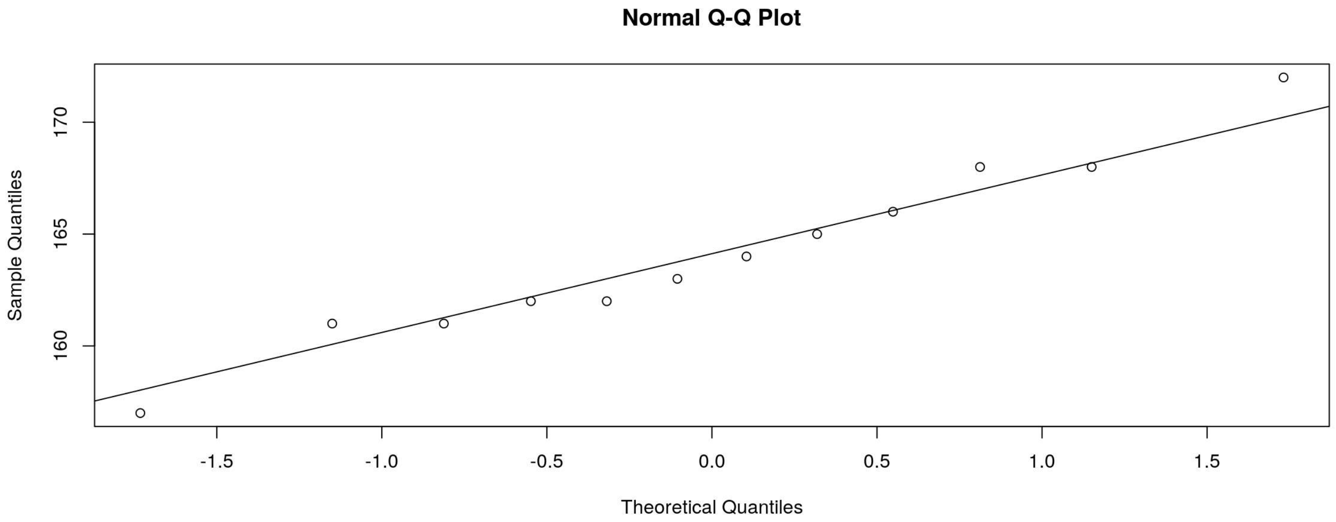
Czy możemy stwierdzić, że średni wzrost mężczyzn jest znacznie większy niż wzrost kobiet?




```
## [1] 0.8595396
```




```
## [1] 0.9447828
```




```
## [1] 45.74359
```

```
## [1] 16.08333
```

```
## [1] 0.04689163
```

```
## [1] 179.0769
```

```
## [1] 164.0833
```

```
## [1] 6.928802e-07
```

Zadanie 5.

a. Napisz funkcję `w_test()` implementującą test χ^2 w modelu wykładniczym, który jest opisany we wskazówce. Funkcja ta powinna mieć trzy argumenty: `x` - wektor zawierający dane, `lambda_zero` - wartość λ_0 w hipotezie zerowej oraz `alternative` - typ hipotezy alternatywnej, która może mieć trzy możliwe wartości: `"two.sided"` (wartość domyślna), `"greater"` , `"less"` . Funkcja zwraca obiekt będący listą klasy `htest` o elementach: `statistic` - wartość statystyki testowej, `parameter` - liczba stopni swobody, `p.value` - p-wartość, `alternative` - wybrana hipoteza alternatywna, `method` - nazwa testu, `data.name` - nazwa zbioru danych (użyj `deparse(substitute(x))`). Dla obiektów klasy `htest` funkcja `print()` istnieje w programie R, więc nie trzeba jej tworzyć.

Wskazówka. Niech $\mathbf{X} = (X_1, \ldots, X_n)^\top$ będzie próbą prostą z populacji o rozkładzie wykładniczym $Ex(\lambda)$, gdzie $\lambda > 0$ jest nieznanym parametrem. Testy χ^2 w modelu wykładniczym weryfikują hipotezę zerową $H_0 : \lambda = \lambda_0$, gdzie $\lambda_0 > 0$ jest ustaloną liczbą. Ich obszary krytyczne są następujące:

1. dla $H_1^{(1)} : \lambda > \lambda_0$

$$R = \left\{\mathbf{x} : T(\mathbf{x}) \leqslant \chi^2(\alpha, 2n)\right\},$$

2. dla $H_1^{(2)} : \lambda < \lambda_0$

$$R = \left\{\mathbf{x} : T(\mathbf{x}) \geqslant \chi^2(1-\alpha, 2n)\right\},$$

3. dla $H_1^{(3)} : \lambda \neq \lambda_0$

$$R = \left\{\mathbf{x} : T(\mathbf{x}) \geqslant \chi^2(1-\alpha/2, 2n)\ or\ T(\mathbf{x}) \leqslant \chi^2(\alpha/2, 2n)\right\},$$

gdzie

$$T(\mathbf{X}) = 2\lambda_0 n\bar{X}\Big|_{H_0} \sim \chi^2(2n)$$

jest statystyką testową, a $\chi^2(\beta, m)$ oznacza kwantyl rzędu β z rozkładu chi-kwadrat $\chi^2(m)$ z m stopniami swobody.

b. Wykorzystując funkcję `w_test()` zastosuj test χ^2 w modelu wykładniczym do danych dotyczących czasu bezawarynej pracy dostępnych w pliku awarie.txt i hipotezy zerowej $H_0 : \lambda = 0.001$.

```
## [1] 0.0009079683
```

```
##
##  Test chi-kwadrat w modelu wykładniczym
##
## data:  awarie$V1
## T = 110.14, num df = 100, p-value = 0.2295
## alternative hypothesis: less
```