

Projekt WdAD

Filip Cebula 151410

19 grudnia 2024

1 Zadanie 1

1.1 Podpunkt A

Zapisujemy dane do naszego zadania.

```
1 earnings = c(45617,7166,18594,2236,1278,19828,4033,28151,2414,3800);  
2 earnings.sd = 15000;
```

Estymujemy średnią z próbki, oraz błąd standardowy (korzystając z danego odchylenia standardowego).

```
1 earnings.mean = mean(earnings);  
2 earnings.se = earnings.sd / sqrt(length(earnings));  
3  
4 # Output
```

Obliczamy przedziały ufności 90%, przy założeniu rozkładu normalnego próbki.

```
1 earnings.interval1 = 0.1;  
2 earnings.ci1.margin = qnorm(p = (1 - earnings.interval1 / 2)) * earnings.se;  
3 earnings.ci1 = earnings.mean + c(-1,1) * earnings.ci1.margin;  
4  
5 # Output
```

Obliczamy przedziały ufności 95%, przy założeniu rozkładu normalnego próbki.

```
1 earnings.interval2 = 0.05;  
2 earnings.ci2.margin = qnorm(p = (1 - earnings.interval2 / 2)) * earnings.se;  
3 earnings.ci2 = earnings.mean + c(-1,1) * earnings.ci2.margin;  
4  
5 # Output
```

1.2 Podpunkt B

Odrzucamy dane nam wcześniej odchylenie standardowe i estymujemy nowe odchylenie standardowe, oraz błąd standardowy na podstawie danej próbki.

```
1 earnings.sd = sd(earnings);
2 earnings.se = earnings.sd / sqrt(length(earnings));
3
4 # Output
```

Używamy rozkładu t Studenta, żeby obliczyć przedziały ufności dla naszej próbki.
Przedział ufności 90%.

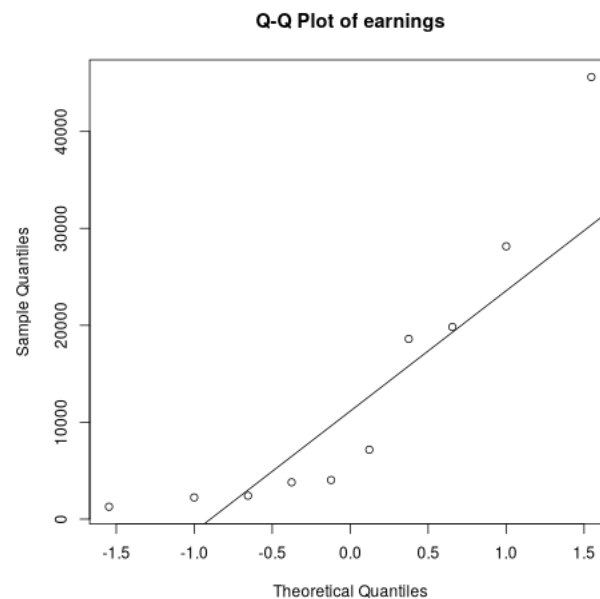
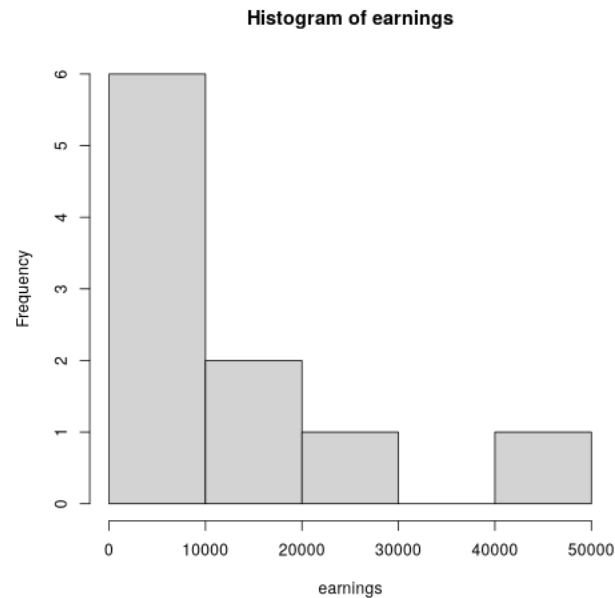
```
1 earnings.interval1 = 0.1;
2 earnings.ci1.margin = qt(p = (1 - earnings.interval1 / 2),
3                           df = length(earnings) - 1) * earnings.se;
4 earnings.ci1 = earnings.mean + c(-1,1) * earnings.ci1.margin;
5
6 # Output
```

Przedział ufności 95%.

```
1 earnings.interval2 = 0.05;
2 earnings.ci2.margin = qt(p = (1 - earnings.interval2 / 2),
3                           df = length(earnings) - 1) * earnings.se;
4 earnings.ci2 = earnings.mean + c(-1,1) * earnings.ci2.margin;
5
6 # Output
```

1.3 Podpunkt C

Odrzucamy założenie o normalności rozkładu naszej próbki. Możemy użyć histogramu, oraz wykresu kwantyl-kwantyl, aby sprawdzić jak blisko rozkładu normalnego jest rozkład naszej próbki.



Na podstawie powyższych wykresów, a szczególnie wykresu kwantyl-kwantyl, możemy zauważyć, że rozkład naszej próbki nie jest rozkładem normalnym. Wynika z tego, że obliczone przez nas powyżej przedziały ufności nie są dokładne. Aby otrzymać dokładniejsze wyniki, użyjemy metody bootstrap do obliczenia przedziałów ufności.

Tworzymy funkcję bootstrap i używamy jej na naszej próbce.

```
1 my_bootstrap = function(data) {  
2   n = length(data);  
3   means = c();  
4  
5   for (i in 1:10000) {  
6     rands = sample(1:n, n, replace = T);  
7     xs = data[rands];  
8  
9     means = append(means, mean(xs));  
10  }  
11  
12  return(means);  
13 }  
14  
15 earnings.bootstrap = my_bootstrap(earnings);
```

Wyliczamy przedział ufności 90%.

```
1 quantile(earnings.bootstrap, probs = c(0.05, 0.95));  
2 # Output
```

Wyliczamy przedział ufności 95%.

```
1 quantile(earnings.bootstrap, probs = c(0.025, 0.975));  
2 # Output
```

2 Zadanie 2

2.1 Shrimp

2.2 Sitka89

2.3 Quine