

9. vježbe

1 Zadatak

Učitajte datoteku `podaci.dat` u kojoj prvi stupac predstavlja podatke za maksimalne širine lubanja Etruščana, a drugi stupac maksimalne širine lubanja današnjih Talijana.

1. Procjenite prosječnu maksimalnu širinu lubanja Etruščana.
2. Procjenite prosječnu maksimalnu širinu lubanja Talijana.

Rješenje:

1.

```
> (etr<-mean(podaci$etruscani))  
[1] 143.8429
```
2.

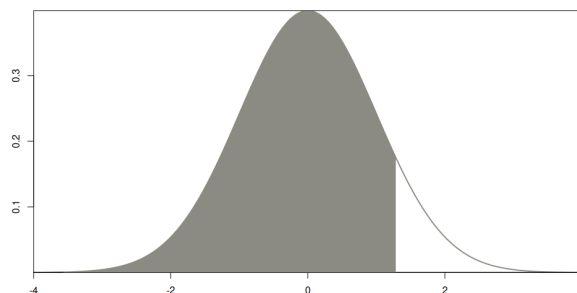
```
> (tal<-mean(podaci$talijani))  
[1] 132.4429
```

2 Zadatak

Neka je $X \sim N(0, 1)$ slučajna varijabla.

1. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(X \leq a) = 0.9$
2. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(X \leq a) = 0.95$
3. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(X \leq a) = 0.99$
4. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(X \geq a) = 0.9$
5. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(X \geq a) = 0.95$
6. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(X \geq a) = 0.99$

Rješenje: U prva tri podzadatka tražimo točku a na x-osi tako da je do te točke površina ispod grafa funkcije gustoće standardne normalne (normalne s očekivanjem 0 i standardnom varijacijom 1) slučajne varijable lijevo od te točke a dani broj 0.9, 0.95, 0.99. U R-u koristimo naredbu `qnorm()` koja je inverzna (suprotna) naredbi `pnorm()`.



```
1. > qnorm(p=0.9, mean=0, sd=1)
[1] 1.281552
> pnorm(q=1.281552, mean = 0, sd = 1)
[1] 0.9000001
```

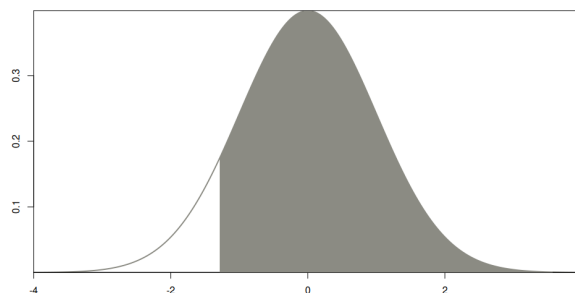
Ovdje drugu naredbu nije bilo potrebno pozvati, ali nam demonstrira vezu između `qnorm()` i `pnorm()` funkciji. Prva za danu površinu daje točku na x-osi, a druga za danu točku na x-osi daje površinu ispod grafa.

```
2. > qnorm(p=0.95, mean=0, sd=1)
[1] 1.644854
```

```
3. > qnorm(p=0.99, mean=0, sd=1)
[1] 2.326348
```

```
4. > qnorm(p=0.1, mean=0, sd=1)
[1] -1.281552
> -qnorm(p=0.9, mean=0, sd=1)
[1] -1.281552
> qnorm(p=0.9, mean=0, sd=1, lower.tail = FALSE)
[1] -1.281552
```

Sve ove naredbe daju isto rješenje, točku a takvu da je površina ispod grafa desno od nje 0.9



```
5. > qnorm(p=0.95, mean=0, sd=1, lower.tail = FALSE)
[1] -1.644854
```

```
6. > qnorm(p=0.99, mean=0, sd=1, lower.tail = FALSE)
[1] -2.326348
```

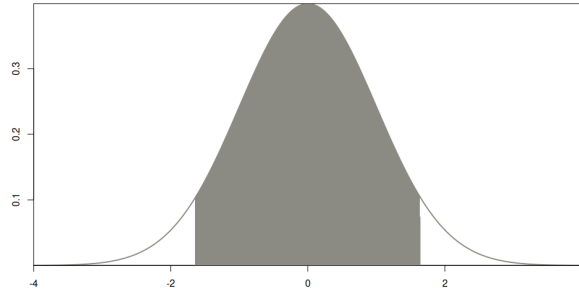
3 Zadatak

Neka je $X \sim N(0, 1)$ slučajna varijabla.

1. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(-a \leq X \leq a) = 0.9$
2. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(-a \leq X \leq a) = 0.95$
3. Odredite $a \in \mathbb{R}$ takav da je $\mathbb{P}(-a \leq X \leq a) = 0.99$

Rješenje: Tražimo točku a na x-osi tako da je do te točke površina ispod grafa funkcije gustoće standardne normalne slučajne varijable između točke $-a$ i a dani broj 0.9, 0.95, 0.99.

1. Podijelimo površinu 0.1 na dva jednaka dijela, od kojih će jedan biti lijevo od točke $-a$, a drugi desno od točke a



```
> qnorm(p=0.05, mean=0, sd=1)
[1] -1.644854
> qnorm(p=0.95, mean=0, sd=1)
[1] 1.644854
```

Dakle $a = 1.644854$.

```
2. > alpha<-0.95
> qnorm(p=(1-alpha)/2, mean=0, sd=1)
[1] -1.959964
```

Dakle $a = 1.959964$.

```
3. > alpha<-0.99
> qnorm(p=(1-alpha)/2, mean=0, sd=1)
[1] -2.575829
```

Dakle $a = 2.575829$.

4 Zadatak

Neka je $X \sim N(\mu, \sigma^2)$ slučajna varijabla.

1. Koja je distribucija slučajne varijable $X - \mu$?
2. Koja je distribucija slučajne varijable $\frac{X - \mu}{\sigma}$?

Rješenje: Ako je X normalna slučajna varijabla tada ćemo zbrajanjem X s realnim brojem i množenjem X s realnim brojem ponovno dobiti normalno distribuiranu slučajnu varijablu. Ona će imati različite parametre od X . Prvi parametar μ je očekivanje od X , a drugi σ^2 je varijanca

1. $\mathbb{E}[X - \mu] = \mathbb{E}[X] - \mathbb{E}[\mu] = \mu - \mu = 0$, dok je
 $\text{Var}(X - \mu) = \mathbb{E}[(X - \mu)^2] - \mathbb{E}[X - \mu]^2 = \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mu^2 + 0^2 = \mathbb{E}[X^2] - \mu^2 = \text{Var}(X)$,
 pa je $X - \mu \sim N(0, \sigma^2)$
2. $\mathbb{E}[aX] = a\mathbb{E}[X]$, dok je $\text{Var}(aX) = \mathbb{E}[(aX)^2] - \mathbb{E}[aX]^2 = a^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2) = a^2\text{Var}(X)$,
 pa kad uzmemo $a = \frac{1}{\sigma}$ dobivamo da je $\frac{X - \mu}{\sigma} \sim N(0, 1)$

Ovaj zadatak ćemo koristiti da odredimo pouzdane intervale u idućem zadatku, gdje pretpostavljamo da je svaka vrijednost maksimalne širine lubanje realizacija normalne slučajne varijable, koje sve imaju isto očekivanje i varijancu ali su međusobno nezavisne. Budući da imamo 70 mjerenja to nam daje $X_1, X_2, \dots, X_{70} \sim N(\mu, \sigma^2)$ kako je aritmetička sredina $\bar{X} = \frac{X_1 + X_2 + \dots + X_{70}}{70}$ normalna slučajna varijabla s očekivanjem $\mathbb{E}[\bar{X}] = \frac{1}{70}\mathbb{E}[X_1 + X_2 + \dots + X_{70}] = \frac{70\mu}{70} = \mu$ i varijancom $\text{Var}(\bar{X}) = \frac{1}{70^2}\text{Var}(X_1 + X_2 + \dots + X_{70}) = \frac{70\sigma^2}{70^2} = \frac{\sigma^2}{70}$ odnosno zaključujemo da je $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ (kod nas je $n = 70$).

Sada po prethodnom zadatku vidimo da je $Y = \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0, 1)$ (kod nas je $n = 70$). Na Y ćemo primijeniti zadatak 3 kako bi dobili traženi interval. Imamo da je

$$\begin{aligned}\mathbb{P}(-1.644854 \leq Y \leq 1.644854) &= 0.9 \\ \mathbb{P}\left(-\frac{1.644854}{\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma} \leq \frac{1.644854}{\sqrt{n}}\right) &= 0.9 \\ \mathbb{P}\left(-\frac{1.644854}{\sqrt{n}}\sigma \leq \bar{X} - \mu \leq \frac{1.644854}{\sqrt{n}}\sigma\right) &= 0.9 \\ \mathbb{P}\left(\bar{X} - \frac{1.644854}{\sqrt{n}}\sigma \leq \mu \leq \bar{X} + \frac{1.644854}{\sqrt{n}}\sigma\right) &= 0.9\end{aligned}$$

Dakle 90% pouzdani interval za očekivanje μ će biti $\left[\bar{X} - \frac{1.644854}{\sqrt{n}}\sigma, \bar{X} + \frac{1.644854}{\sqrt{n}}\sigma\right]$. Prosijek smo izračunali u 1. zadatku, rekli smo da imamo $n = 70$ podataka, jedino još ne znamo koliko je σ . Budući da je veličina uzorka $n = 70$ velika (veće od 20) možemo za σ uzeti standardnu devijaciju uzorka.

5 Zadatak

1. Odredite 90% pouzdani interval za maksimalnu širinu lubanja Etruščana.
2. Odredite 95% pouzdani interval za maksimalnu širinu lubanja Etruščana.
3. Odredite 99% pouzdani interval za maksimalnu širinu lubanja Etruščana.
4. Odredite 90% pouzdani interval za maksimalnu širinu lubanja Talijana.
5. Odredite 95% pouzdani interval za maksimalnu širinu lubanja Talijana.
6. Odredite 99% pouzdani interval za maksimalnu širinu lubanja Talijana.

Rješenje:

```
1. > n<-length(podaci$etruscani)
   > alpha<-0.9
   > etr+ qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
   [1] 141.5022
   > etr- qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
   [1] 146.1835
```

Dakle traženi interval je [141.5022, 146.1835].

```
2. > alpha<-0.95
   > etr+ qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
   [1] 141.0538
   > etr- qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
   [1] 146.6319
```

Dakle traženi interval je [141.0538, 146.6319]. Vidimo da se širina intervala povećava. To ima smisla, jer ako želimo biti više sigurni da će očekivana širina lubanje biti u tom intervalu moramo imati širi interval.

```
3. > alpha<-0.99
   > etr+ qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
   [1] 140.1775
   > etr- qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
   [1] 147.5083
```

Dakle traženi interval je [140.1775, 147.5083]. Najširi od ova tri.

```

4. > n<-length(podaci$talijani)
> alpha<-0.9
> tal+ qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
[1] 130.1022
> tal- qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
[1] 134.7835

```

Dakle traženi interval je [130.1022, 134.7835].

```

5. > alpha<-0.95
> tal+ qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
[1] 129.6538
> tal- qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
[1] 135.2319

```

Dakle traženi interval je [129.6538, 135.2319]. Ponovno se širina intervala povećava.

```

6. > alpha<-0.99
> tal+ qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
[1] 128.7775
> tal- qnorm(p=(1-alpha)/2, mean=0, sd=1)*sqrt(n)/sd(podaci$etruscani)
[1] 136.1083

```

Dakle traženi interval je [128.7775, 136.1083]. Vidimo da se 99% pouzdani intervali za očekivanu maksimalnu širinu lubanje talijana i etruščana nimalo ne preklapaju. Zaključujemo da se širina glave značajno promijenila kroz vrijeme.