Sklop: Normalni model z znano varianco

1 Primer

Imamo dva vzorca, v katerih je podano število ur, ki so jih dijaki dveh srednjih šol potrebovali za pripravo domače naloge.

```
school1 <- c(2.11, 9.75, 13.88, 11.3, 8.93, 15.66, 16.38, 4.54, 8.86, 11.94,
  12.47, 11.11, 11.65, 14.53, 9.61, 7.38, 3.34, 9.06, 9.45, 5.98,
  7.44, 8.5, 1.55, 11.45, 9.73)
school2 <- c(0.29, 1.13, 6.52, 11.72, 6.54, 5.63, 14.59, 11.74, 9.12, 9.43,
             10.64, 12.28, 9.5, 0.63, 15.35, 5.31, 8.49, 3.04, 3.77, 6.22,
             2.14, 6.58, 1.11)
```

Privzemimo normalni model z znano varianco $\sigma^2 = 4$, torej $(X_i|\mu) \sim N(\mu, \sigma^2 = 4)$, medtem ko apriorna porazdelitev za μ naj bo $\mu \sim N(\mu_0 = 6, \tau_0^2 = 9)$. Zanimala nas aposteriorna porazdelitev μ za vsako šolo.

```
var <- 2
mu0 <- 6
tau0 <- 3
```

Uporabimo naslednjo oznako:

- $\vec{X}_1 = (X_{(1,1)}, X_{(2,1)}, ..., X_{(n_1,1)})$ - stevilo ur v prvem vzorcu
- $\vec{X}_2 = (X_{(1,2)}, X_{(2,2)}, ..., X_{(n_2,2)})$ - stevilo ur v drugem vzorcu

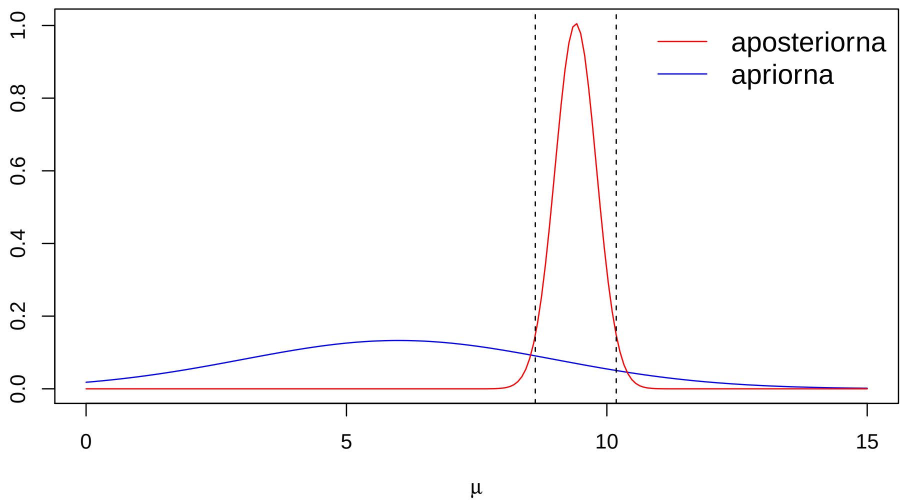
2 Naloge

2.1 Izračunajte aposteriorno porazdelitev za vsako šolo. Narišite apriorno in aposteriorno porazdelitev na istem grafu (spet za vsako šolo posebej). Na vsakem grafu dodajte se 95% centralni kredibilnostni interval (pomagajte si s funkcijo `qnorm`).

```
mus <- seq(0, 15, length=200)
apriorna <- dnorm(mus, mu0, tau0)
```

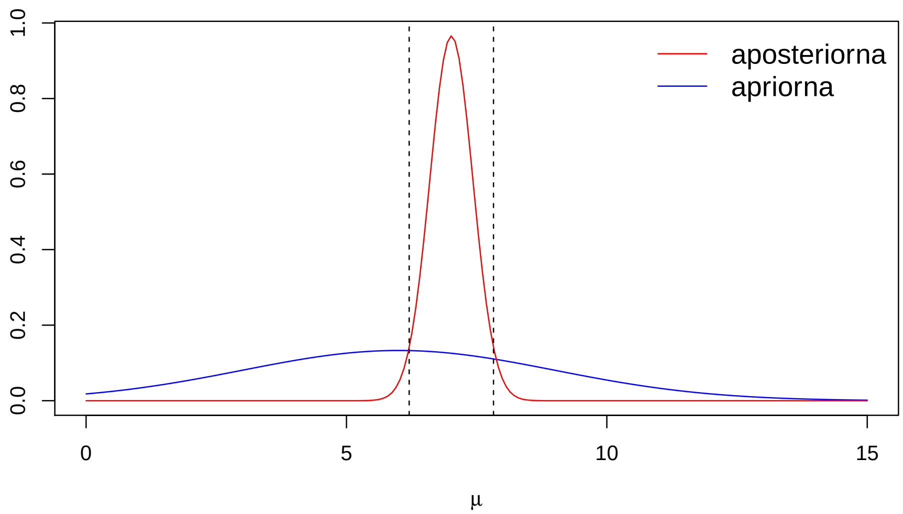
Sola 1:

```r
n1 <- length(school1)
mu1 <- tau0^2/(var^2/n1 + tau0^2) * mean(school1) + var^2/n1/(var^2/n1 + tau0^2) * mu0
tau1 <- var^2/n1 * tau0^2 /(var^2/n1 + tau0^2)
aposteriorna1 <- dnorm(mus, mu1, sqrt(tau1))
I1 <- qnorm(c(0.025,0.975), mu1, sqrt(tau1))
```



Sola 2:

```r
n2 <- length(school2)
mu2 <- tau0^2/(var^2/n2 + tau0^2) * mean(school2) + var^2/n2/(var^2/n2 + tau0^2) * mu0
tau2 <- var^2/n2 * tau0^2 /(var^2/n2 + tau0^2)
aposteriorna2 <- dnorm(mus, mu2, sqrt(tau2))
I2 <- qnorm(c(0.025,0.975), mu2, sqrt(tau2))
```



2.2 Izracunajte verjetnost:

$$P[(\mu|\vec{X}_1) > (\mu|\vec{X}_2)]$$

Verjetnost lahko izračunate eksaktno, ali pa jo ocenite s pomočjo simulacije.

Kaj nam pove ta verjetnost?

```
1 - pnorm(0, mu2 - mu1, sqrt(tau1 + tau2))
```

```
## [1] 1.499305e-05
```

2.3 Naj bosta $\tilde{X}_1$, $\tilde{X}_2$ nove vrednosti iz obeh vzorcev, za katere želimo podati napovedi. Na grafih iz naloge 2.1 dodajte še aposteriorne napovedne porazdelitve $(\tilde{X}_1|\vec{X}_1)$ in $(\tilde{X}_2|\vec{X}_2)$.

2.4 Izracunajte verjetnost:

$$P((\tilde{X}_1|\vec{X}_1) > (\tilde{X}_2|\vec{X}_2))$$

Verjetnost lahko izračunate eksaktno, ali pa jo ocenite s pomočjo simulacije.

Kaj nam pove ta verjetnost?

2.5 Dodatna naloga

Dobili smo podatke za še eno šolo:

```r
school3 <- c(4.33, 7.77, 4.15, 5.64, 7.69, 5.04, 10.01, 13.43, 13.63, 9.9,
             5.72, 5.16, 4.33, 12.9, 11.27, 6.05, 0.95, 6.02, 12.22, 12.85)
```

Označimo:

- $\vec{X}_3 = (X_{(1,3)}, X_{(2,3)}, ..., X_{(n_1,3)})$ - stevilo ur v tretjem vzorcu

Izračunajte verjetnost, da je aposteriorna porazdelitev μ v prvem vzorcu večja kot aposteriorna porazdelitev μ v drugem in tretjem vzorcu:

$$P[(\mu|\vec{X}_1) > (\mu|\vec{X}_2) \wedge (\mu|\vec{X}_1) > (\mu|\vec{X}_3)]$$

Izračunajte še verjetnost:

$$P((\tilde{X}_1|\vec{X}_1) > (\tilde{X}_2|\vec{X}_2) \wedge (\tilde{X}_1|\vec{X}_1) > (\tilde{X}_3|\vec{X}_3))$$

Obe verjetnosti lahko izračunate eksaktno, ali oceno podate s pomočjo simulacije.