

14 Razdiobe i simulacije - DOMAĆA ZADAĆA

0036538677 Lorena Jakić

2024-01-22

Domaća zadaća

1. Ako pretpostavimo da je vjerojatnost dobivanja na ruletu $1/36$, koja je vjerojatnost da ćete u 50 igranja dobiti barem 2 puta?

```
1 - pbinom(1, 50, 1/36)
```

```
## [1] 0.4062166
```

2. Procijenite vjerojatnost dobivanja 4 asa u pokeru, pri čemu pretpostavljate da će igrač prvo nasumično vući 5 karata (od mogućih 52), a potom odbaciti sve karte koje nisu asevi i ponovo vući toliko karata. Vjerojatnost procijenite simulirajući ovaj uzorak ponašanja 100,000 puta i računajući postotak pozitivnih ishoda.

NAPUTAK: smislite kako na praktičan način oblikovati objekt koji predstavlja špil karata (uzimajući u obzir da nam je relevantno samo da li je karta as ili ne). Potom napravite funkciju koja simulira gore opisano ponašanje igrača i vraća informaciju da li je postignut pozitivan ishod ili ne.

```
spil <- c(rep(1, 4), rep(0, 48))
```

```
izvzlacenja <- function() {  
  prvoIzvzlacenje <- sample(spil, 5)  
  brAseva <- sum(prvoIzvzlacenje)  
  novispil <- c(rep(1, 4 - brAseva), rep(0, 48))  
  drugoIzvzlacenje <- sample(novispil, 5 - brAseva)  
  brAseva2 <- sum(drugoIzvzlacenje)  
  if (brAseva2 + brAseva == 4)  
    return(1)  
  else  
    return(0)  
}  
replicate(100000, izvzlacenja()) %>% mean()
```

```
## [1] 0.00026
```

3. Jedan od stožernih teorema statistike je “centralni granični teorem” koji (pojednostavljeno rečeno) kaže da ako kontinuirano ponavljamo uzorkovanje populacije sredine ovih uzoraka će se ravnati po normalnoj razdiobi kojoj je sredina jednaka sredini originalne razdiobe populacije a varijanca jednaka varijanci populacije podijeljenoj sa veličinom uzorka.

Simulacijom dokažite ovu tvrdnju. Ponovite 1000 puta eksperiment nasumičnog uzorkovanja 100 mjera iz normalne razdiobe sa sredinom 50 i standardnom devijacijom 10 (rezultate pohranite u matricu 1000×100 , svaki redak predstavlja jedno provedeno uzorkovanje). Potom na istom grafu nacrtajte

- histogram svih 10,000 mjera
- histogram aritmetičkih sredina svakog pojedinog uzorkovanja (tj. histogram 1000 dobivenih sredina)

```

matrica_uzoraka <- replicate(1000, rnorm(100, 50, 10)) %>% t
sredine <- rowMeans(matrica_uzoraka)

podaci <- data.frame(
  tip = "Sve mjere",
  vrijednost = as.vector(matrica_uzoraka)
)
podaci2 <- data.frame(
  tip = "Aritmeticka sredina",
  vrijednost = as.vector(sredine)
)
podaci <- rbind(podaci, podaci2)

ggplot(podaci, aes(x = vrijednost, fill = tip)) +
  geom_histogram(binwidth = 1, position = "identity", alpha = 0.7, color = "black") +
  labs(title = "Histogram svih mjera i aritmetickih sredina", x = "Vrijednosti", y = "Broj promatranja") +
  scale_fill_manual(values = c("darkblue", "lightgreen")) +
  theme_minimal()

```

