

Домашно задание по СЕМ Практикум

Даниел Халачев, №62547, група 2

Задача 1

В тесте от 32 карти има 4 маркирани. Тестето е разбъркано и раздадено на 4 играчи. Чрез подходяща симулация намерете приближение на вероятността на събитието "на всеки играч се пада една маркирана карта".

```
simulate.deal = function(special, standard){
  # represent the deck and shuffle it
  cards = sample(c(rep(0, standard), rep(1, special)))
  # verify each of the four parts has one special card
  sum(cards[1:8])==1 && sum(cards[9:16])==1 && sum(cards[17:24])==1 && sum(cards[25:32])==1
}
probability.cards = function(timesToRepeat, special, standard){
  result = replicate(timesToRepeat, simulate.deal(special, standard))
  sum(result)/length(result)
}
probability.cards(10000, 4, 28)
```

Изпълнението на кода дава приближение на вероятността 0.11-0.12.

Задача 2

В кутия има 5 зелени, 7 сини и 7 жълти топки. Симулирайте n тегления на топка с връщане. Проверете хипотезата, че в кутията има еднакъв брой топки от трите цвята. Повторете 10000 пъти за $n = 50, 100, 200, 500, 1000$. Колко често заключението на теста е вярно? Представете чрез подходяща графика честотата на вярно заключение в зависимост от n .

```
simulate.take = function(quantity, green, blue, yellow){
  # represent the set of balls
  balls = sample(c(rep("g", green), rep("b", blue), rep("y", yellow)), quantity, TRUE)
  # assume they are an equal quantity - H0
  probs = c(1/3, 1/3, 1/3)
  # chi^2 test
  balls = table(balls)
  chi2.obs = sum ((balls-quantity*probs)^2/(quantity*probs))
  p.value = 1-pchisq(chi2.obs, length(probs)-1)
  # if p.value > 0.05, H0 is accepted
  p.value > 0.05
}
probability.balls = function(timesToRepeat, quantity, green, blue, yellow){
  result = replicate(timesToRepeat, simulate.take(quantity, green, blue, yellow))
  # how often the conclusion is right = how often H0 is discarded = 1 - how often H0 is accepted
  1-(sum(result)/length(result))
}
n.values = c(50, 100, 200, 500, 1000)
results = data.frame(n = integer(), probability = double())
for (n in n.values){
  probability = probability.balls(10000, n, 5, 7, 7)
  results <- rbind(results, data.frame(n, probability))
}
colnames = n.values
barplot(results$probability, xlab = "n", names.arg = results$n, ylim=c(0,1.0))
```

Резултатите са обобщени в графиката по-долу:



Графиката показва, че точността на извода за хипотезата нараства с броя тегления n .

Задача 3

В кутия има 5 зелени, 5 сини и 5 жълти топки. Симулирайте n тегления на топка с връщане. Проверете хипотезата, че в кутията има еднакъв брой топки от трите цвята. Повторете 10000 пъти за $n = 50, 100, 200, 500, 1000$. Колко често заключението на теста е вярно? Представете чрез подходяща графика честотата на вярно заключение в зависимост от n .

```
simulate.take = function(quantity, green, blue, yellow){
  # represent the set of balls
  balls = sample(c(rep("g", green), rep("b", blue), rep("y", yellow)), quantity, TRUE)
  # assume they are of equal quantity - H0
  probs = c(1/3, 1/3, 1/3)
  # chi^2 test
  balls = table(balls)
  chi2.obs = sum ((balls-quantity*probs)^2/(quantity*probs))
  p.value = 1-pchisq(chi2.obs, length(probs)-1)
  # if p.value > 0.05, H0 is accepted
  p.value > 0.05
}
probability.balls = function(timesToRepeat, quantity, green, blue, yellow){
  result = replicate(timesToRepeat, simulate.take(quantity, green, blue, yellow))
  # how often the conclusion is right = how often H0 is accepted
  (sum(result)/length(result))
}
n.values = c(50, 100, 200, 500, 1000)
results = data.frame(n = integer(), probability = double())
for (n in n.values){
  probability = probability.balls(10000, n, 5, 5, 5)
  results <- rbind(results, data.frame(n, probability))
}
colnames = n.values
barplot(results$probability, xlab = "n", names.arg = results$n, ylim=c(0,1.0))
```

Резултатите са обобщени в графиката по-долу:



Графиката показва, че точността на извода за хипотезата е 95% за всяко n .