HW1 STT810 Tiancheng Liu

Tiancheng Liu

2022-09-12

Table of Contents

## Question 1

### a

samp <- sample(c(1:6),100,replace = TRUE)  
samp

## [1] 2 5 2 1 5 6 6 5 1 3 5 6 4 5 3 1 6 4 3 5 5 1 2 2 5 4 3 4 3 1 1 3 4 6 3 4 3  
## [38] 1 2 4 5 4 2 5 3 6 1 5 1 4 6 1 3 1 2 1 2 4 6 2 3 6 5 1 5 4 2 3 3 4 5 6 6 3  
## [75] 2 5 2 1 4 6 4 1 6 4 5 4 1 4 4 3 4 1 3 2 2 4 5 4 5 1

### b

mean(samp)

## [1] 3.46

### c

samp1000 <- sample(c(1:6),1000,replace = TRUE)  
mean(samp1000)

## [1] 3.623

samp10000 <- sample(c(1:6),10000,replace = TRUE)  
mean(samp10000)

## [1] 3.504

samp100000 <- sample(c(1:6),100000,replace = TRUE)  
mean(samp100000)

## [1] 3.49897

samp1000000 <- sample(c(1:6),1000000,replace = TRUE)  
mean(samp1000000)

## [1] 3.499426

Yes, the mean is approaching the true value.

### d

samp1000 <- sample(c(0,1),1000,replace = TRUE)  
mean(samp1000)

## [1] 0.498

samp10000 <- sample(c(0,1),10000,replace = TRUE)  
mean(samp10000)

## [1] 0.4989

samp100000 <- sample(c(0,1),100000,replace = TRUE)  
mean(samp100000)

## [1] 0.50018

samp1000000 <- sample(c(0,1),1000000,replace = TRUE)  
mean(samp1000000)

## [1] 0.500358

Yes,same convergence.

### e

samp1000 <- sample(c(0,1,1),1000,replace = TRUE)  
mean(samp1000)

## [1] 0.664

samp10000 <- sample(c(0,1,1),10000,replace = TRUE)  
mean(samp10000)

## [1] 0.6627

samp100000 <- sample(c(0,1,1),100000,replace = TRUE)  
mean(samp100000)

## [1] 0.66757

samp1000000 <- sample(c(0,1,1),1000000,replace = TRUE)  
mean(samp1000000)

## [1] 0.666797

Yes, the result is converging to 0.667.

## Question 2

### a

The sample space is {0,1,2,3,4,5}.

### b

6\*6 there are a total of 36 possible outcomes, and if the difference is 1 then it has to be 12,23,34,45,56,65,54,43,32,21, a total of 10 outcomes. Then the probability is 10/36 = 0.278

### c

fir <- sample(c(1:6),10000,replace = TRUE)  
sec <- sample(c(1:6),10000,replace = TRUE)  
dif <- abs(fir-sec)  
sum(dif == 1)/length(dif)

## [1] 0.2763

We can see the result agrees on what we have in b.

## Qustion 3

### a

Dice 1, coin 1 head\*2 Dice 2, coin 2 head.

(1+2)/4\*6 = 0.125 chance to win the game.

### b

count <- 0  
  
for (i in 1:10000){  
 dice <- sample(c(1:6),1,replace = T)  
 coin <- sample(c(0,1),2,replace = T)  
 if (dice == sum(coin)){  
 count = count + 1  
 }  
}  
count/10000

## [1] 0.1198

The probability to win the game is about 0.125.

## Question 4

Let

A : a viewer watched gymnastics B : a viewer watched baseball C : a viewer watched soccer

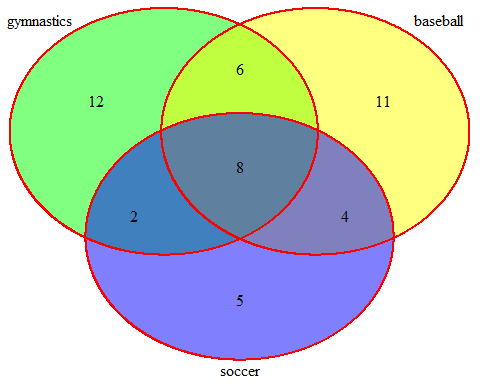
We have

library(VennDiagram)

## 载入需要的程辑包：grid

## 载入需要的程辑包：futile.logger

draw.triple.venn(area1=28, area2=29, area3=19,  
 n12=14, n23=12, n13=10, n123=8,  
 category=c("gymnastics","baseball","soccer"),  
 col="Red",fill=c("Green","Yellow","Blue"))



## (polygon[GRID.polygon.1], polygon[GRID.polygon.2], polygon[GRID.polygon.3], polygon[GRID.polygon.4], polygon[GRID.polygon.5], polygon[GRID.polygon.6], text[GRID.text.7], text[GRID.text.8], text[GRID.text.9], text[GRID.text.10], text[GRID.text.11], text[GRID.text.12], text[GRID.text.13], text[GRID.text.14], text[GRID.text.15], text[GRID.text.16])

Thus the result should be (100-28-29-19+14+12+10+8)/100 = 0.68.

68 percent of the group watched none of the three sports during the last year.

## Question 5 Matloff 1.9

Skipped for the question is incorrect.

## Question 6 Matloff 1.10

### a

Pn3n4 <- 2/4  
Pn3n4

## [1] 0.5

Thus the answer for a is 0.5.

### b

pn43 <- (1/2)\*(1/4)+(1/2)\*(1/4) #p(n3=1 and n4=3) + p(n3=2 and n4=3)  
pn43

## [1] 0.25

Thus the answer is 0.25.

## Question 7 Matloff 2.1

minpiece <- function(k) {  
 breakpts <- sort(runif(k-1))  
 lengths <- diff(c(0,breakpts ,1))  
 min(lengths)  
}  
# returns the approximate probability  
# that the smallest of k pieces will  
# have length less than q  
bkrod <- function(nreps ,k,q) {  
 minpieces <- replicate(nreps ,minpiece(k))  
 mean(minpieces < q)  
}  
bkrod(10000,5,0.02)

## [1] 0.3352

#here is the modified function  
bkrod(10000,sample(c(2,3,4),1,replace = T,prob = c(0.3,0.3,0.4)),0.02)

## [1] 0.2126

The probability is shown above.

## question 8 Matloff 2.6

nreps <- 10000  
nstops <- 9 # because the 10th arrival is the same as 9th departure  
count <- 0  
pas\_tt <- rep(0,10000)  
for (i in 1:nreps){  
 pas\_lis <- 0  
 passengers <- 1  
 for (j in 1:nstops) {  
 if (passengers > 0){  
 for (k in 1:passengers){  
 if (runif(1) < 0.2)  
 passengers <- passengers - 1  
 }  
 }  
 newpass <- sample(0:2,1,prob=c(0.5,0.4,0.1))  
 passengers <- passengers + newpass  
 if (passengers == 0){  
 pas\_lis <- 1  
 }  
 }  
 pas\_tt[i] <- pas\_lis  
}  
  
sum(pas\_tt)/10000

## [1] 0.3021

Thus we can see that, it about 30% chance the bus will be empty for at least one stop.