STT810 ICA1 Tiancheng Liu

2022-09-06

Table of Contents

## Question 1

### a

library(zoo)

##   
## 载入程辑包：'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

x\_t <- rep(0,20)  
for (i in 1:20){  
 x\_t[i] <- system.time(sample(c(0, 1), 1000000\*(2\*i), replace = TRUE))[1]  
}  
x\_t

## [1] 0.25 0.44 0.69 0.83 1.04 1.17 1.29 1.52 1.53 1.88 1.94 2.22 2.47 2.64 2.44  
## [16] 2.62 3.02 3.14 2.89 3.43

We can see that 10000000 runs takes about 1 second.

### b

sample(1:10,10)

## [1] 5 1 6 10 4 9 8 2 7 3

We see now the 10 numbers are in random order.

## Question 2

### a

p\_b <- (0.6\*0.8+(1-0.6)\*0.1)  
p\_ab <- 0.6\*0.8/p\_b  
p\_ab

## [1] 0.9230769

The probability that a given message is spam, given that is marked as spam is 0.923.

### b

#A : not marked as spam  
#B : spam  
#P(A|B) = 0.5  
#0.5 = P(B|A)P(A / (P(A)\*P(B|A) + (1-P(A))\*P(A|B))  
#0.5 = (1-0.8)\*P(A)/P(A\*(1-0.8)+(1-P(A))\*0.9)  
p\_a <- 0.45/0.55  
p\_a

## [1] 0.8181818

Thus 81.818 percent of all emails have to be spam in order for 50% of emails which are not marked to be spam to be actually spam.