MThambeliyagodage\_Data605\_W7\_Assign7

Matheesha Thambeliyagodage

March 28, 2018

Table of Contents

## Problem 1

Let be mutually independent random variables, each of which is uniformly distributed on the integers from to . Let denote the minimum of the ’s. Find the distribution of .

#### SOLUTION

Number of possible combinations of ’s is (choosing values out of options with replacement).

Consider number of combinations with at least one . It is equal to all combinations () minus all combinations with values between and (). So .

Consider number of combinations with at least one and no . It is euqal to all combinations () minus all combinations with at least one (see above: ) and minus all combinations with values between and (). So .

Similarly considering combinations without or and with at least one ,

.

More generally, we can see that .

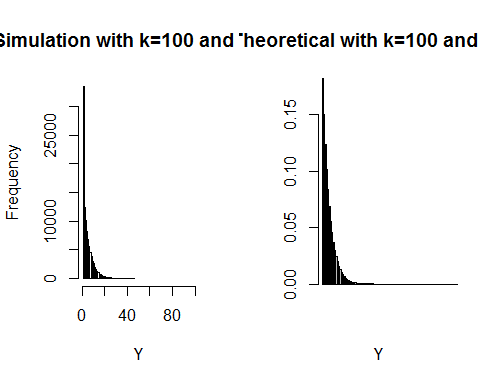
#### SIMULATION

Set up a function to run simulated trials.

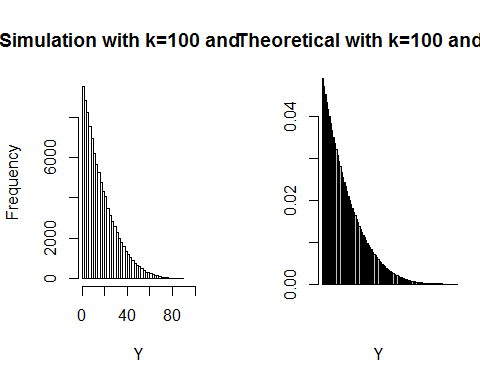
problem1sim <- function(k,n,trials=100000) {  
 Y<-rep(0,trials)  
 for (i in 1:trials) {  
 x<-sample.int(k,size=n,replace=TRUE)  
 Y[i]<-min(x)  
 }  
 return(Y)  
}

Plot distribution of simulated trials and theoretical probability distribution for several values of and .

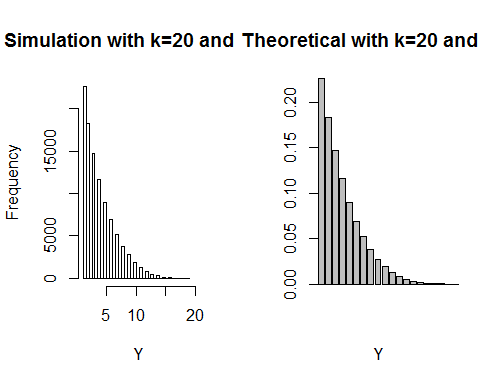
# Run 1  
par(mfrow=c(1,2))  
k<-100  
n<-20  
hist(problem1sim(k,n),breaks=60,  
 main=paste("Simulation with k=",k," and n=",n,sep=""),  
 xlab="Y",xlim=c(1,k))  
pY<-((k-1:k+1)^n-(k-1:k)^n)/k^n  
barplot(pY,main=paste("Theoretical with k=",k," and n=",n,sep=""),  
 xlab="Y",xlim=c(1,k))



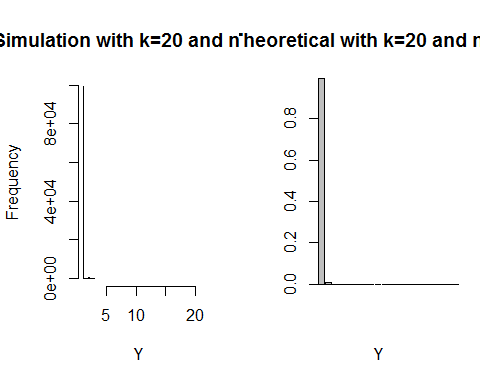
# Run 2  
par(mfrow=c(1,2))  
k<-100  
n<-5  
hist(problem1sim(k,n),breaks=60,  
 main=paste("Simulation with k=",k," and n=",n,sep=""),  
 xlab="Y",xlim=c(1,k))  
pY<-((k-1:k+1)^n-(k-1:k)^n)/k^n  
barplot(pY,main=paste("Theoretical with k=",k," and n=",n,sep=""),  
 xlab="Y",xlim=c(1,k))



# Run 3  
par(mfrow=c(1,2))  
k<-20  
n<-5  
hist(problem1sim(k,n),breaks=60,  
 main=paste("Simulation with k=",k," and n=",n,sep=""),  
 xlab="Y",xlim=c(1,k))  
pY<-((k-1:k+1)^n-(k-1:k)^n)/k^n  
barplot(pY,main=paste("Theoretical with k=",k," and n=",n,sep=""),  
 xlab="Y",xlim=c(1,k))



# Run 4  
par(mfrow=c(1,2))  
k<-20  
n<-100  
hist(problem1sim(k,n),breaks=60,  
 main=paste("Simulation with k=",k," and n=",n,sep=""),  
 xlab="Y",xlim=c(1,k))  
pY<-((k-1:k+1)^n-(k-1:k)^n)/k^n  
barplot(pY,main=paste("Theoretical with k=",k," and n=",n,sep=""),  
 xlab="Y",xlim=c(1,k))



## Problem 2

Your organization owns a copier (future lawyers, etc.) or MRI (future doctors). This machine has a manufacturer’s expected lifetime of 10 years. This means that we expect one failure every ten years. (Include the probability statements and R Code for each part.).

With one failure every ten years and . In this scenario, a failure of the machine is considered *success* in probability distributions.

### PART A

*What is the probability that the machine will fail after 8 years? Provide also the expected value and standard deviation. Model as a geometric. (Hint: the probability is equivalent to not failing during the first 8 years?)*

For geometric distribution, CDF , where is the number of failures before the first success (which is how R defines geometric distribution). Alternatively, . So for , .

# Calculating P(X>8) using geometric distribution  
pgeom(8, 0.1, lower.tail=FALSE)

## [1] 0.3874205

Expected number of years before the first machine failure is .

Standard deviation .

### PART B

*What is the probability that the machine will fail after 8 years? Provide also the expected value and standard deviation. Model as an exponential.*

For exponential distribution, CDF , where is the rate parameter. For this example, . . So for , .

# Calculating P(X>8) using exponential distribution  
pexp(8, 0.1, lower.tail=FALSE)

## [1] 0.449329

Expected value is .

Standard deviation .

### PART C

*What is the probability that the machine will fail after 8 years? Provide also the expected value and standard deviation. Model as a binomial. (Hint: 0 success in 8 years)*

For binomial distribution, . Probability of a machine failure after 8 years is the same as probability of 0 *successes* after 8 trials. So for and , .

# Calculating P(X=0) for n=8 using binomial distribution  
pbinom(0,8,0.1,lower.tail=TRUE)

## [1] 0.4304672

Expected value and standard deviation will depend on number of years/trials tracked. Consider first 8 years.

Expected value .

Standard deviation .

### PART D

*What is the probability that the machine will fail after 8 years? Provide also the expected value and standard deviation. Model as a Poisson.*

On average we observe machine failures per year. For Poisson distribution, . Probabilty of a machine failure after 8 years is the same as probability of 0 *successess* after 8 intervals (similarly to the binomial distribution).

# Calculating P(X=0) for 8 intervals using Poisson distribution  
ppois(0,0.1,lower.tail=TRUE)^8

## [1] 0.449329

For Poisson distribution, .