**Design**

Question #1

overtData <- read.csv(file = "Traffic\_data\_orig.csv", header = TRUE,sep = ",")

secretMessage <- "this is a secret message"

messageLen = as.numeric(nchar(secretMessage))

CharToBinary = function(pchar) {

lhex = charToRaw(pchar)

lbits = rev(as.numeric(rawToBits(lhex)))

return(lbits)

}#Converts a character to Binary equivalent

StringToBinary = function(pstr, pstrlen) {

lbitstream = NULL

ltemp = NULL

for (i in 1:pstrlen) {

ltemp = CharToBinary(substring(pstr,i,i))

lbitstream <- c(lbitstream,ltemp)

}

return(lbitstream)

}#converts a string to Binary vector

#Q1. GENERATING MODIFIED PACKET STREAM

binaryMessage = StringToBinary(secretMessage, messageLen)

lenBinaryMessage = as.numeric(length(binaryMessage))

covertPackets = NULL

timeStream = 0

covertDataDelay = NULL

for (i in 1:lenBinaryMessage) {

if(binaryMessage[i] == 0){

timeStream <- timeStream + 0.25

covertDataDelay <- c(covertDataDelay,0.25)

}#if bit is 0, delay is 0.25

else{

timeStream <- timeStream + 0.75

covertDataDelay <- c(covertDataDelay,0.75)

}#if bit is 1, delay is 0.75

covertPackets <- c(covertPackets,timeStream)

}#loop to generate packet stream from binary message

#Q2. Generating the Histogram

temp = 0 #holds the previous timestamp

overtDataDelay = numeric(dim(overtData)[1])

for( i in 1:dim(overtData)[1])

{

overtDataDelay[i] = overtData[i,2] - temp

temp = overtData[i,2]

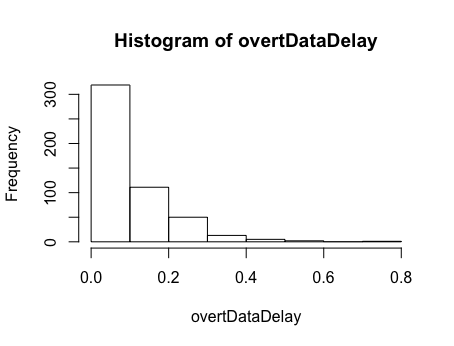
}

hist(overtDataDelay)

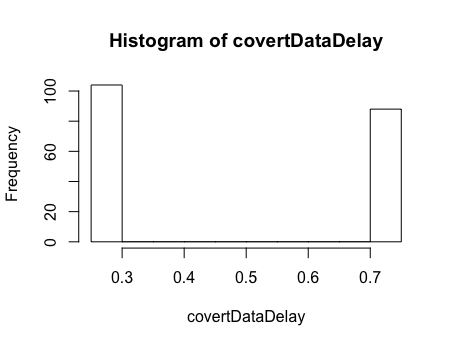
hist(covertDataDelay)

Question #2

**Overt Packet Delay Histogram**

****

**Covert Packet Delay Histogram**



Will Eve be suspicious?

Eve will be suspicious of the 0.75 & 0.25 because it's exactly two unique differences, whereas for the "Overt" data, given to us in the CSV, the histogram looks more like a negative exponential distribution, which is what you would expect of this type of time dependent data

Question #3

#Q3

m = median(overtDataDelay)

min = min(overtDataDelay)

max = max(overtDataDelay)

semiRandomPacketsDelay = NULL

semiRandomPackets = NULL

timeStream = 0

for (i in 1:lenBinaryMessage) {

if(binaryMessage[i] == 0){

timeLapse = runif(1, min, m)

}#if bit is 0, delay is 0.25

else{

timeLapse = runif(1, m, max)

}#if bit is 1, delay is 0.75

timeStream <- timeStream + timeLapse

semiRandomPacketsDelay <- c(semiRandomPacketsDelay,timeLapse)

semiRandomPackets <- c(semiRandomPackets,timeStream)

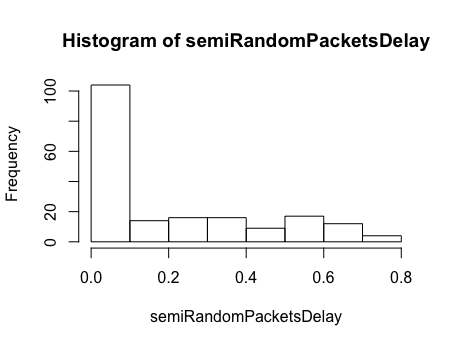
}#loop to generate packet stream from binary message

hist(semiRandomPacketsDelay)

hist(overtDataDelay)

**Note: Code is built up on code from question 1**

Question #4

****

Question #5

1. **Improved Method**

Since we know that the overt timing delay dataset produces an exponential distribution (pdf), we need to make the covert dataset look as exponential as possible so Eve doesn’t suspect us.

We can use the mean m of the overt data set to generate an exponential distribution (using rexp).

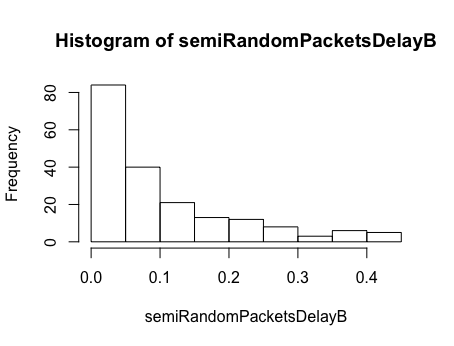
Alice can base the covert delay times on the generated exponential distribution with some alterations to specify the bit values. The cutoffs for 1s and 0s would be represented as a function of the packet number (the 1st packet of the two packets that the delay is between).

**Encoding Scheme:-**

* + Exp(i) - 0.0025 for bit = 1; where i is the packet number and Exp(i) is the density that corresponds to i in an exp(rate = m^-1)
  + Exp(i) + 0.0025 for bit = 0; where i is the packet number and Exp(i) is the density that corresponds to i in an exp(rate = m^-1)

**Assumptions:-**

* Bob knows the distribution of the overt timing dataset (exponential with rate = m)
* Bob knows the encoding scheme so he can decode the message.

****

**Code**

set.seed(1237)

#m = median(overtDataDelay)

#m = quantile(overtDataDelay, probs = c(0.75) ,names= FALSE)

m = mean(overtDataDelay)

min = min(overtDataDelay)

max = max(overtDataDelay)

semiRandomPacketsDelayB = NULL

semiRandomPacketsB = NULL

timeStream = 0

Nexp <- rexp(1000000000000000, m^-1)

for (i in 1:lenBinaryMessage) {

if(binaryMessage[i] == 0){

timelapse = Nexp[i] - 0.0025

#timeLapse = rexp(1, m)

}#if bit is 0, delay is 0.25

else{

timeLapse = Nexp[i] + 0.0025

#timeLapse = runif(1, m)

}#if bit is 1, delay is 0.75

timeStream <- timeStream + timeLapse

semiRandomPacketsDelayB <- c(semiRandomPacketsDelayB,timeLapse)

semiRandomPacketsB <- c(semiRandomPacketsB,timeStream)

}#loop to generate packet stream from binary message

hist(semiRandomPacketsDelayB)

1. If Alice buffers the packets, there will be a moment (in the beginning) when no packets are sent which would disrupt the continuity of the timing channel; this is obviously unrealistic in the case of a Skype call, which is based on real-time interaction. Overall, it would negatively affect the quality of the call, leading Eve to suspicion.
2. If the network modifies the inter packet delay, the message cannot be decoded correctly since the delays are changed and each bit is encoded via delay times. If Bob knows the distribution of the noise within the channel, he can subtract the mean of the noise’s distribution from the covert delay times while decoding to get a somewhat accurate message. This method is obviously flawed as it assumes that the semantic integrity of the message isn’t critical.

**Detection**

Step #1

*Code*

*x <- rnorm(30,mean = 0,sd = 1)*

*y <- rnorm(30,mean = 0,sd = 1)*

*qqplot(x,y, plot.it = TRUE)*

*abline(0,1)*

Step #2

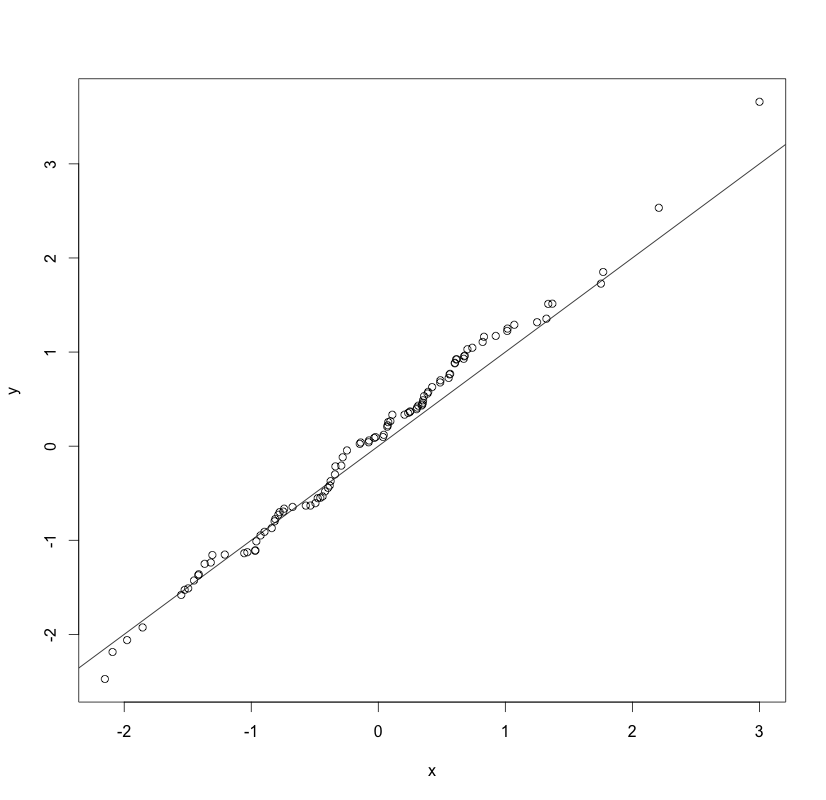
*Code*

*x <- rnorm(100,mean = 0,sd = 1)*

*y <- rnorm(100,mean = 0,sd = 1)*

*qqplot(x,y, plot.it = TRUE)*

*abline(0,1)*

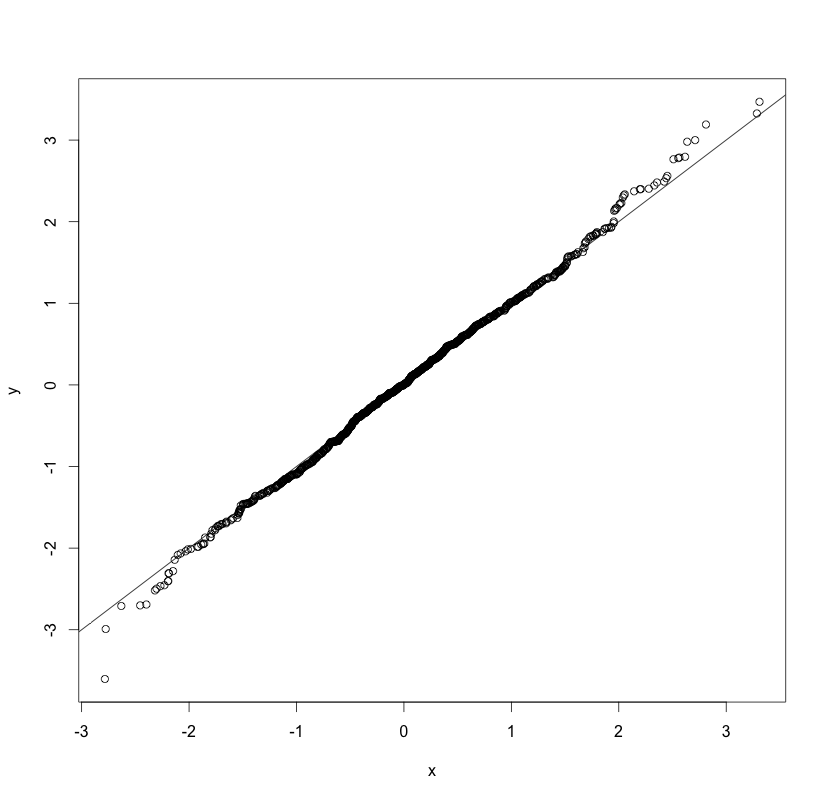
**

*x <- rnorm(1000,mean = 0,sd = 1)*

*y <- rnorm(1000,mean = 0,sd = 1)*

*qqplot(x,y, plot.it = TRUE)*

*abline(0,1)*

**

*Observation*

#It is clear that the shape of the qqplot becomes more linear as n increases.

#This makes sense, as the two distributions are exactly the same.

#The points are dense on the graph between the 1st and 3rd quartiles; they are especially dense around the mean.

Step #3

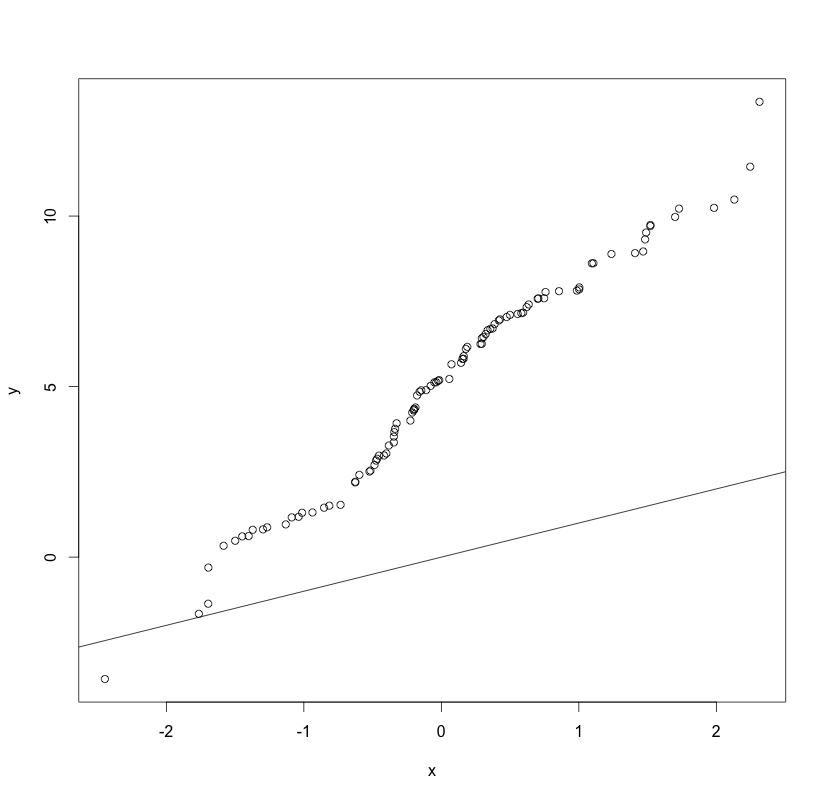
*Code*

*x <- rnorm(100,mean = 0,sd = 1)*

*y <- rnorm(100,mean = 5,sd = 3)*

*qqplot(x,y, plot.it = TRUE)*

*abline(0,1)*

**

*Observation*

The shape of this qqplot also appears to be linear, however the slope is much larger than 1 (which would have indicated perfect correlation). Also, the line itself is shifted up so that at x = 0, y = 5. This shows the relationship between the means: the mean of x is 0, and the mean of y is 5. Using this information, we can apply similar rational to the difference in slope. It is caused by the different standard deviations. The slope for y vs x is 3 in our graph, which also describes the ratio of the standard deviations between the two data sets. Despite the fact that the means and standard deviations are different, and that the numerical values for slope and intercepts change because of it, if we normalize y with respect to x, we see that the it does indeed fall on the slope=1 line.

Step #4

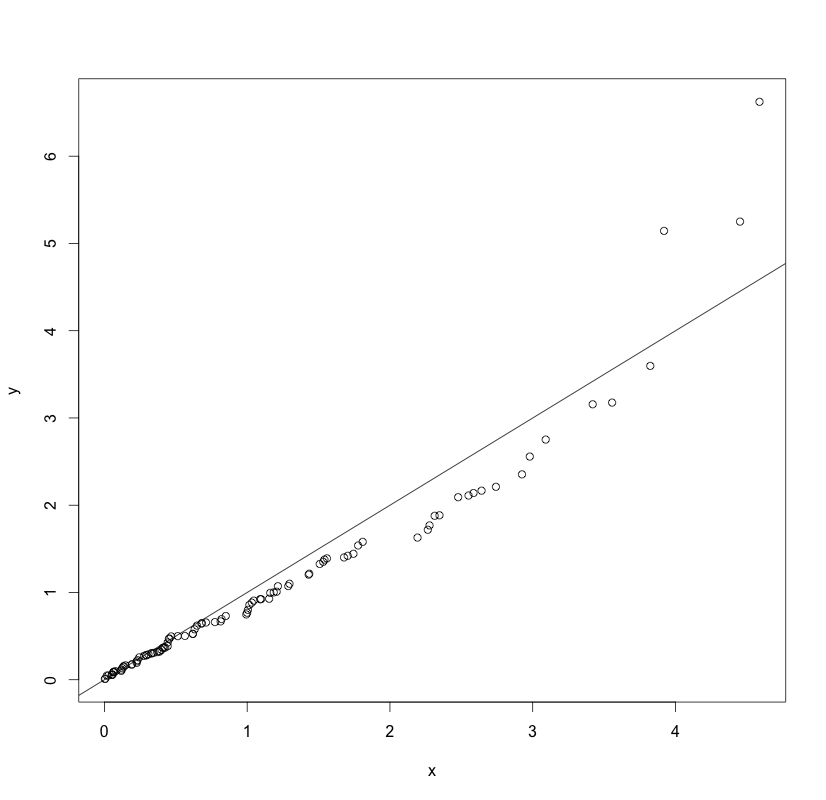
*Code*

*x <- rexp(100,rate = 1)*

*y <- rexp(100,rate = 1)*

*qqplot(x,y, plot.it = TRUE)*

*abline(0,1)*

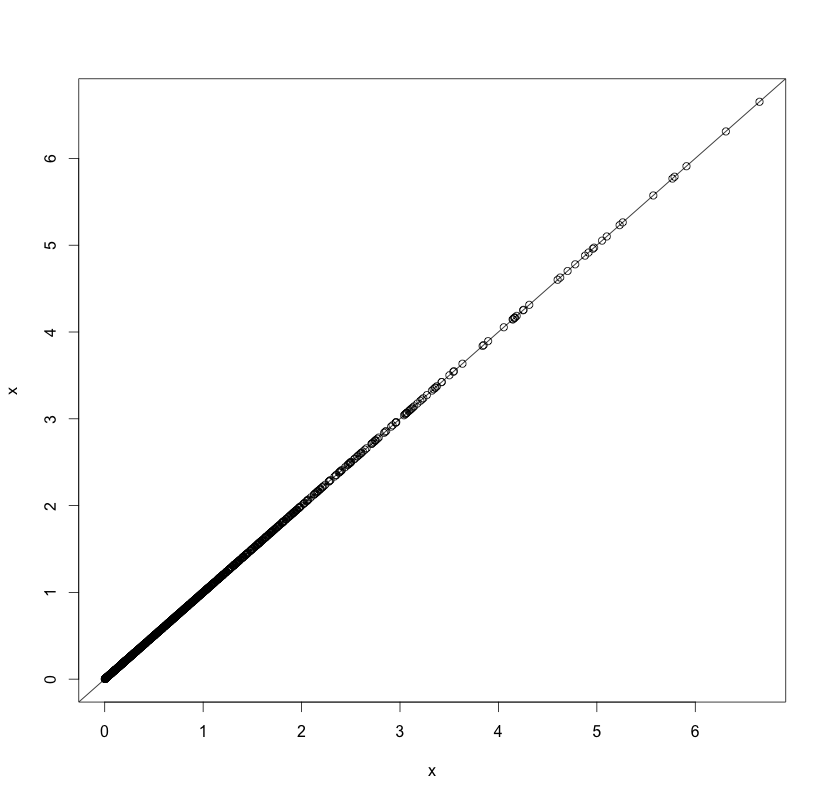
**

*x <- rexp(1000,rate = 1)*

*y <- rexp(1000,rate = 1)*

*qqplot(x,y, plot.it = TRUE)*

*abline(0,1)*

**

*Observation*

#It is clear that the shape of the qqplot becomes more linear as n increases.

#This makes sense as the two distributions are exactly the same.

#The points are dense around 1 which makes sense as the mean for an exponencial distribution is (lamda)^-1.

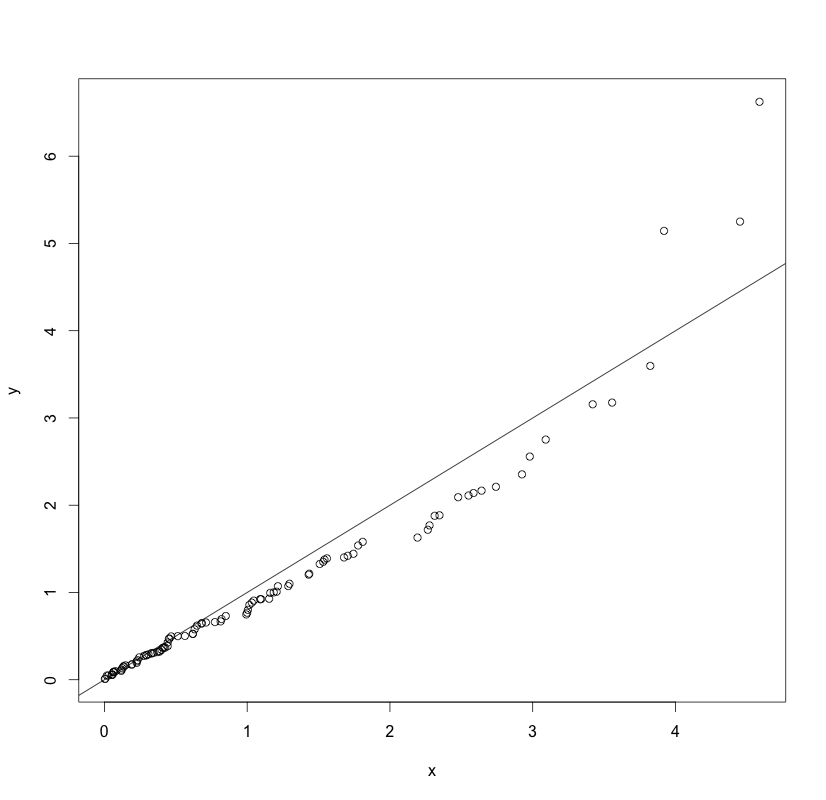
Step #5

*Code*

x <- rnorm(100,mean = 0,sd = 1)

y <- rexp(100,rate = 1)

qqplot(x,y, plot.it = TRUE)

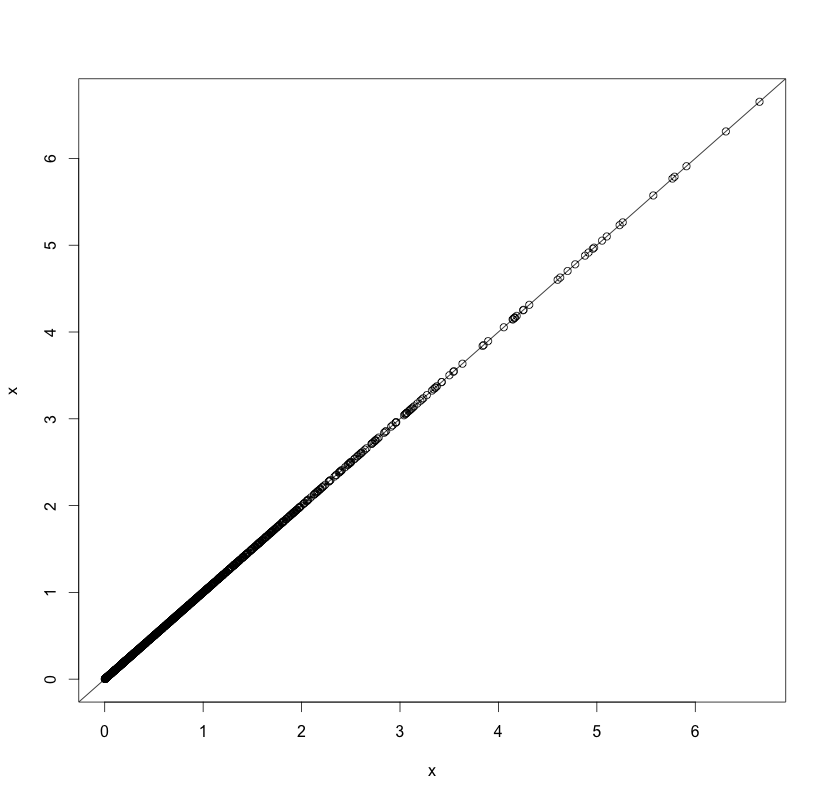


x <- rnorm(500,mean = 0,sd = 1)

y <- rexp(500,rate = 1)

qqplot(x,y, plot.it = TRUE)

abline(0,1)

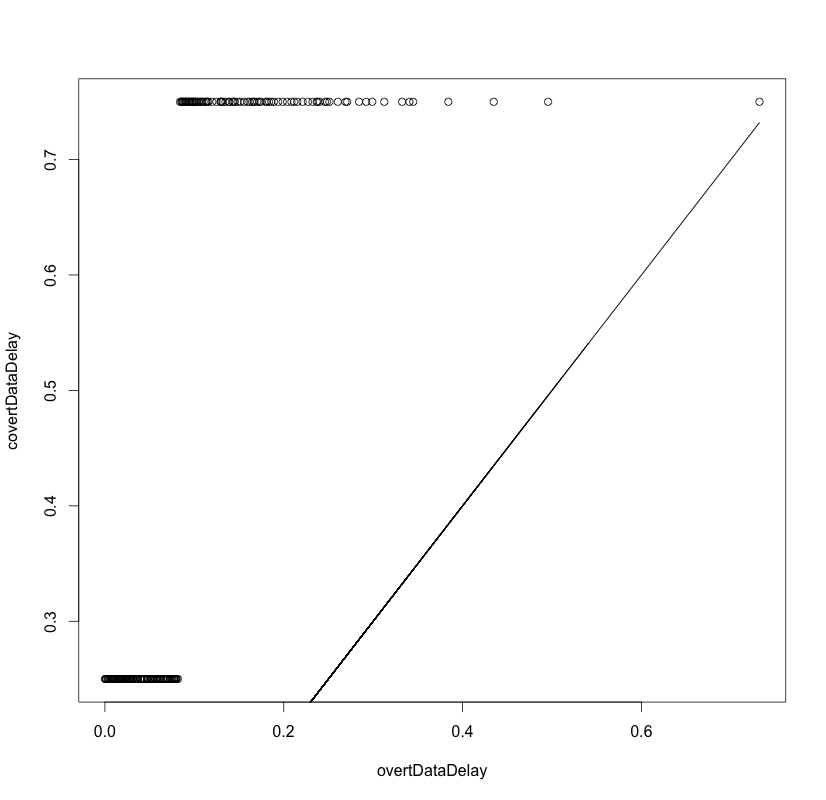


Step #6

*Code*

*qqplot(overtDataDelay, covertDataDelay, plot.it = TRUE)*

*lines(overtDataDelay, overtDataDelay)*

**

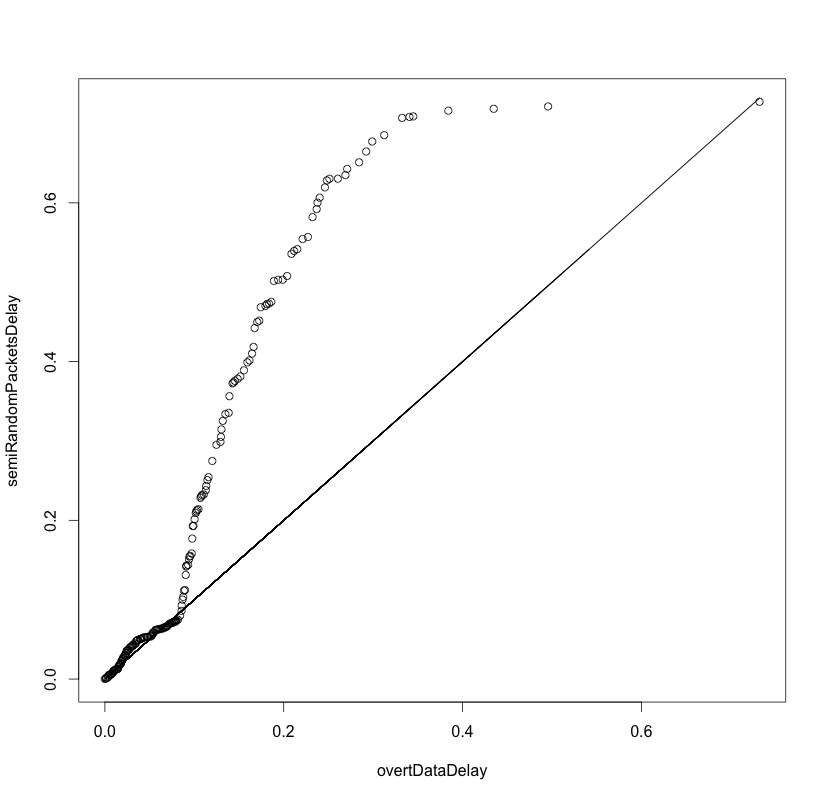
*Observation*

Step #7

*Code*

*qqplot(overtDataDelay, semiRandomPacketsDelay, plot.it = TRUE)*

*lines(overtDataDelay,overtDataDelay)*

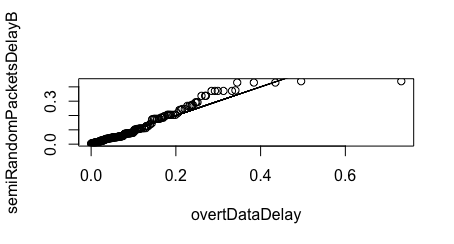


*Observation*

Step #8

*Code*

qqplot(overtDataDelay, semiRandomPacketsDelayB, plot.it = TRUE)



**Note: Code is built up on code from question 1**

**Implementation**

1. IPD is exponential with λ = 1

|  |  |  |  |
| --- | --- | --- | --- |
| m | i | P(overflow) | P(underflow) |
| 16 | 2 | 0 | 0.561 |
|  | 6 | 0 | 0.43 |
|  | 10 | 0 | 0.325 |
|  | 14 | 0 | 0.199 |
|  | 18 | 0 | 0 |
| 16 | 2 | 0.023 | 0.462 |
|  | 6 | 0.017 | 0.397 |
|  | 10 | 0.022 | 0.32 |
|  | 14 | 0.019 | 0.216 |
|  | 18 | 0.02 | 0.203 |

1. IPD is uniform

|  |  |  |  |
| --- | --- | --- | --- |
| m | i | P(overflow) | P(underflow) |
| 16 | 2 | 0 | 0.591 |
|  | 6 | 0 | 0.565 |
|  | 10 | 0 | 0.514 |
|  | 14 | 0 | 0.37 |
|  | 18 | 0 | 0 |
| 16 | 2 | 0 | 0.76 |
|  | 6 | 0 | 0.726 |
|  | 10 | 0 | 0.657 |
|  | 14 | 0.019 | 0.631 |
|  | 18 | 0.02 | 0.584 |