**1. Code**

# set working directory to Lab 04

setwd("C:/Users/jordan/Google Drive/Courses Spring 2018/STAT 350/STAT 350 Labs/Lab 04")

# set up ggplot2 for plotting

library(ggplot2)

###

# PART D: Gamma distribution with parameters alpha=3, beta=2

#

# For a gamma distribution with alpha = 3 and beta = 2,

# generate 1000 random samples of size n = 1, 5, 10, 20, 40, 60, 80, ...

#

# Then for each sample:

# find sample average and

# create histogram (with two colored curves) and

# normal probability plot of sample mean

# for each graph pair, indicate whether they appear sufficiently normal.

# NOTE: do not need to explain judgment of normality.

#

# Do this until n is large enough that distribution of sample mean

# appears normal.

###

samples <- 1000 # number of samples to collect

sizes <- c(1, 5, 10, 20, 40, 60, 80, 100) # sample sizes

# create graph pair for each sample size

for (n in sizes) {

title <- paste("Gamma Distribution: Averaged Over", n) # title for graphs

# generate data and calculate means

gamma.vec <- rgamma(samples\*n, 3, rate=2) # random gamma data

gamma.mat <- matrix(gamma.vec, nrow=samples) # separate data into rows

gamma.means <- apply(gamma.mat, 1, mean) # calculate means

# create histogram

hist <- ggplot(data.frame(gamma.means=gamma.means),aes(gamma.means))+

geom\_histogram(aes(y=..density..),bins=sqrt(samples)+2,

fill="grey",col="black")+

geom\_density(col="red",lwd=1)+

stat\_function(fun=dnorm,args=list(mean=mean(gamma.means),

sd=sd(gamma.means)),

col="blue",lwd=1)+

ggtitle(title)+

xlab("Data")+

ylab("Proportion")

ggsave(hist, filename=paste("gammaHist",n,".png",sep=""))

# create normal probability plot

qq <- ggplot(data.frame(gamma.means=gamma.means),aes(sample=gamma.means))+

stat\_qq()+

geom\_abline(slope=sd(gamma.means),intercept=mean(gamma.means))+

ggtitle(title)+

xlab("Theoretical")+

ylab("Sample")

ggsave(qq, filename=paste("gammaQQ",n,".png",sep=""))

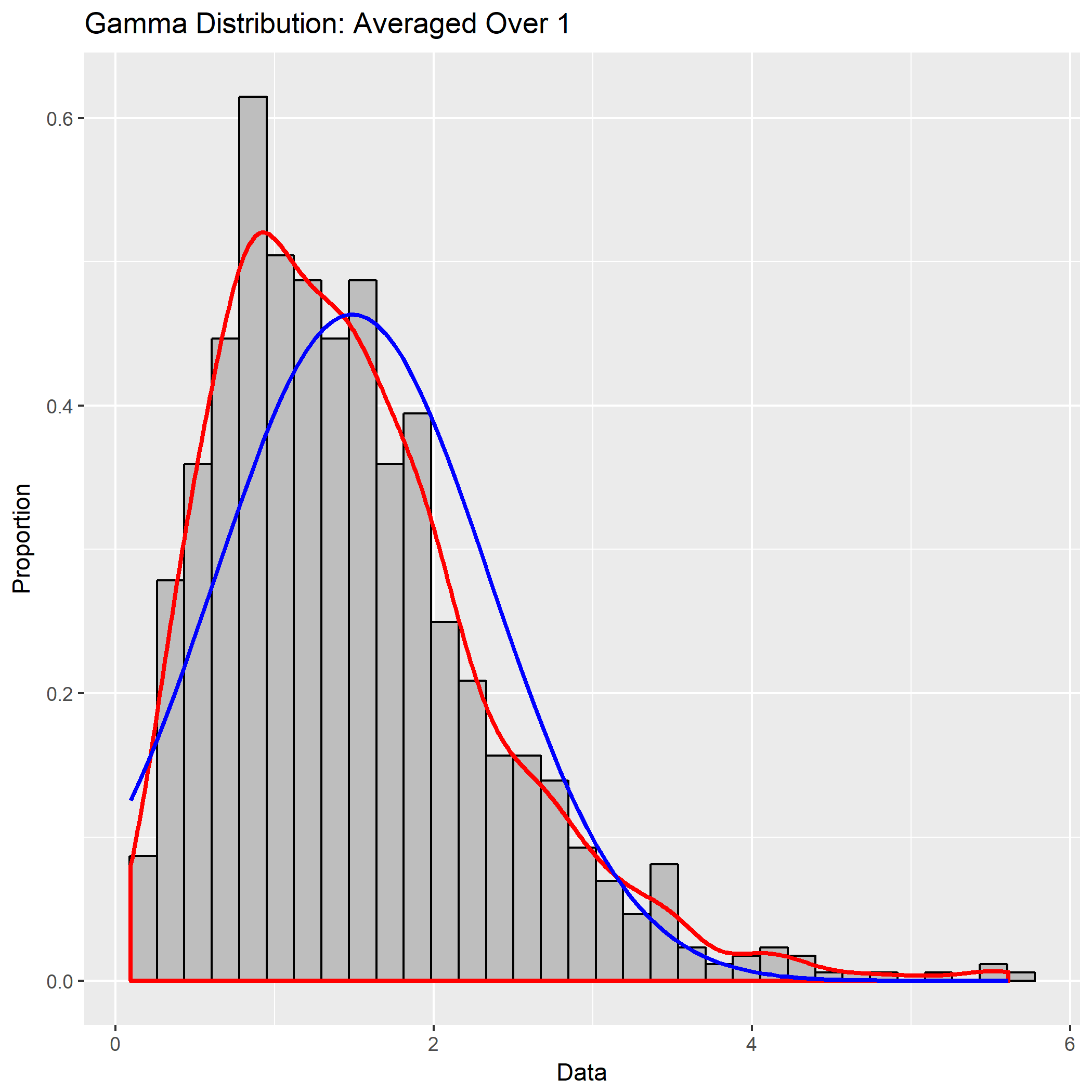
print(paste("n = ", n))

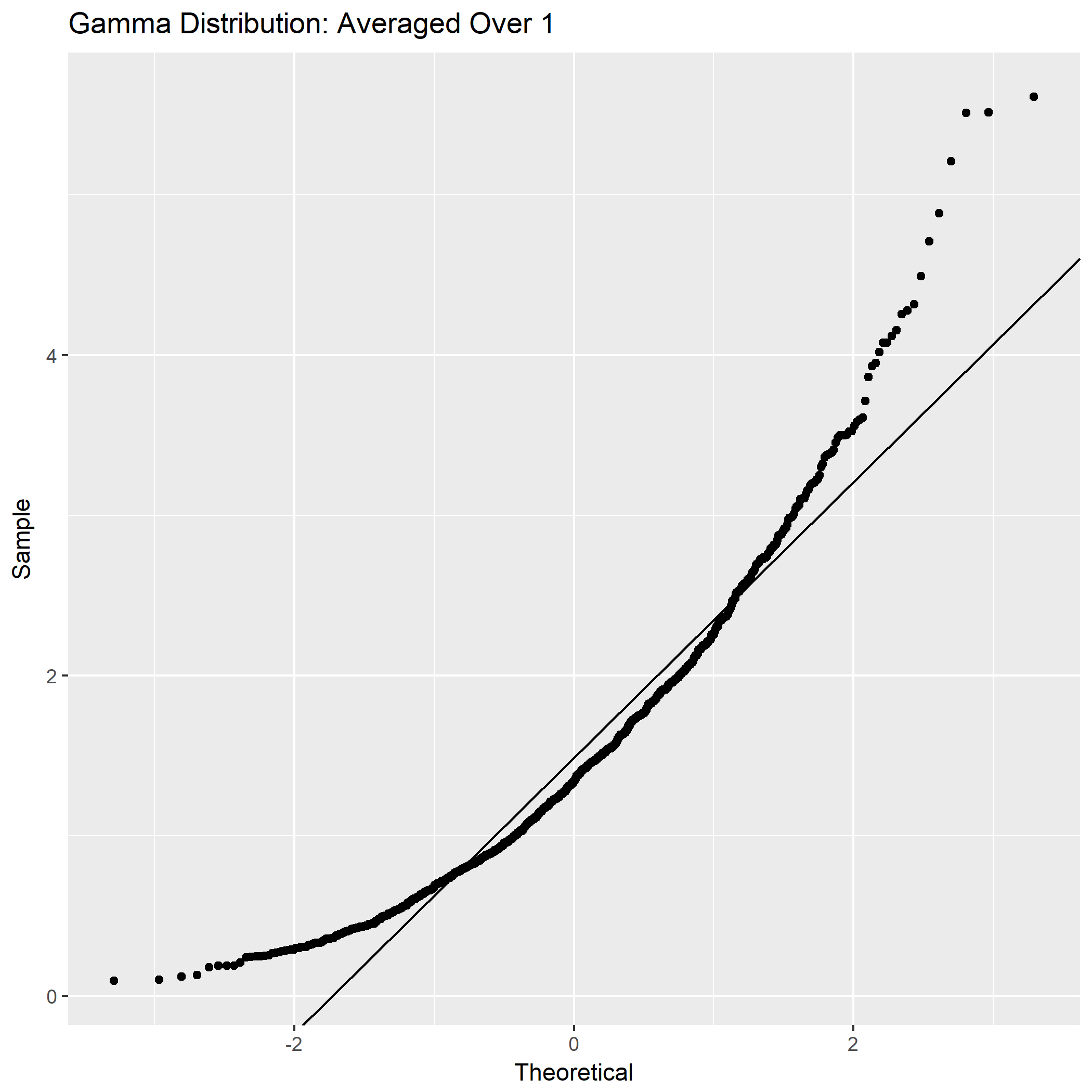
print(paste("mean = ", mean(gamma.means)))

print(paste("sd = ", sd(gamma.means)))

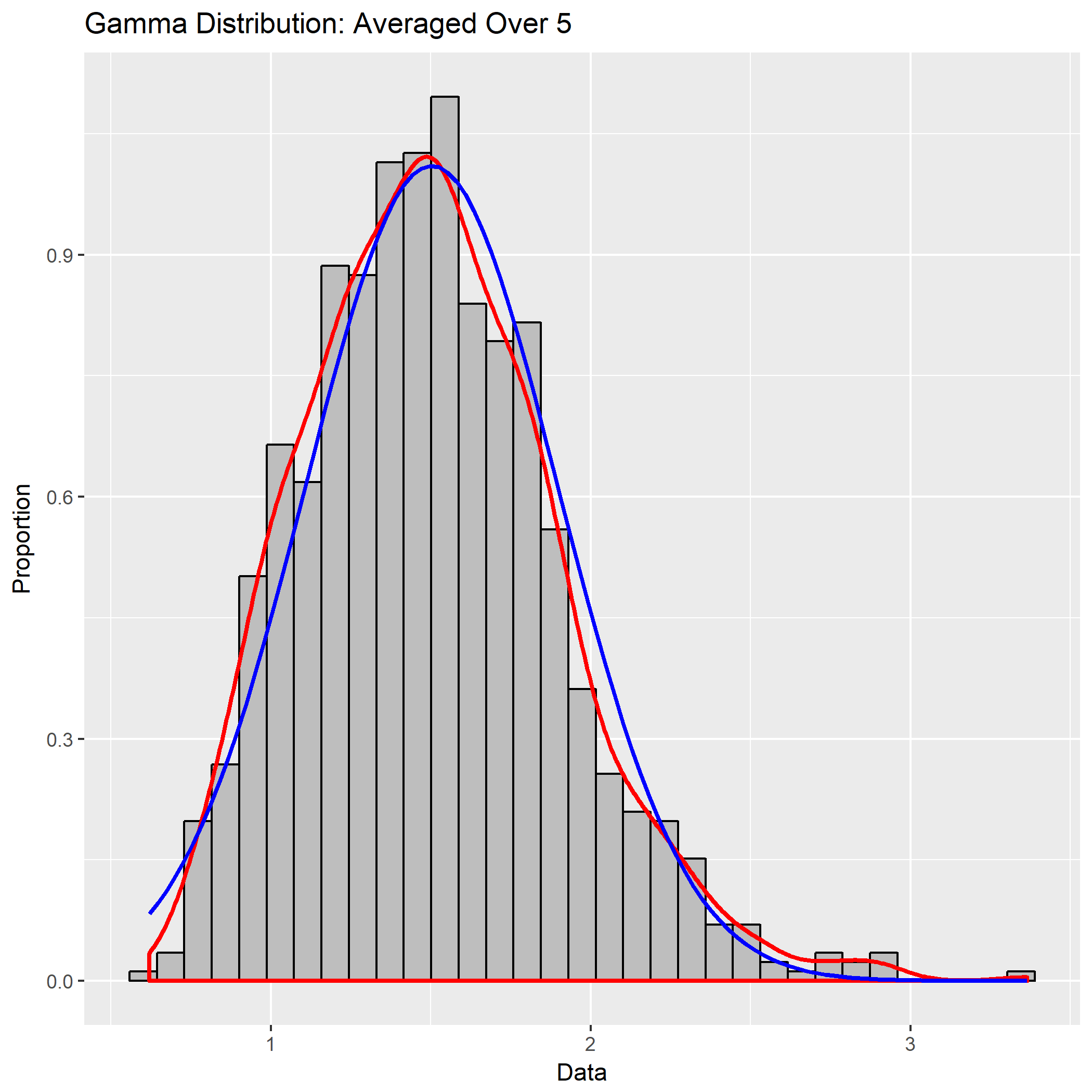
}

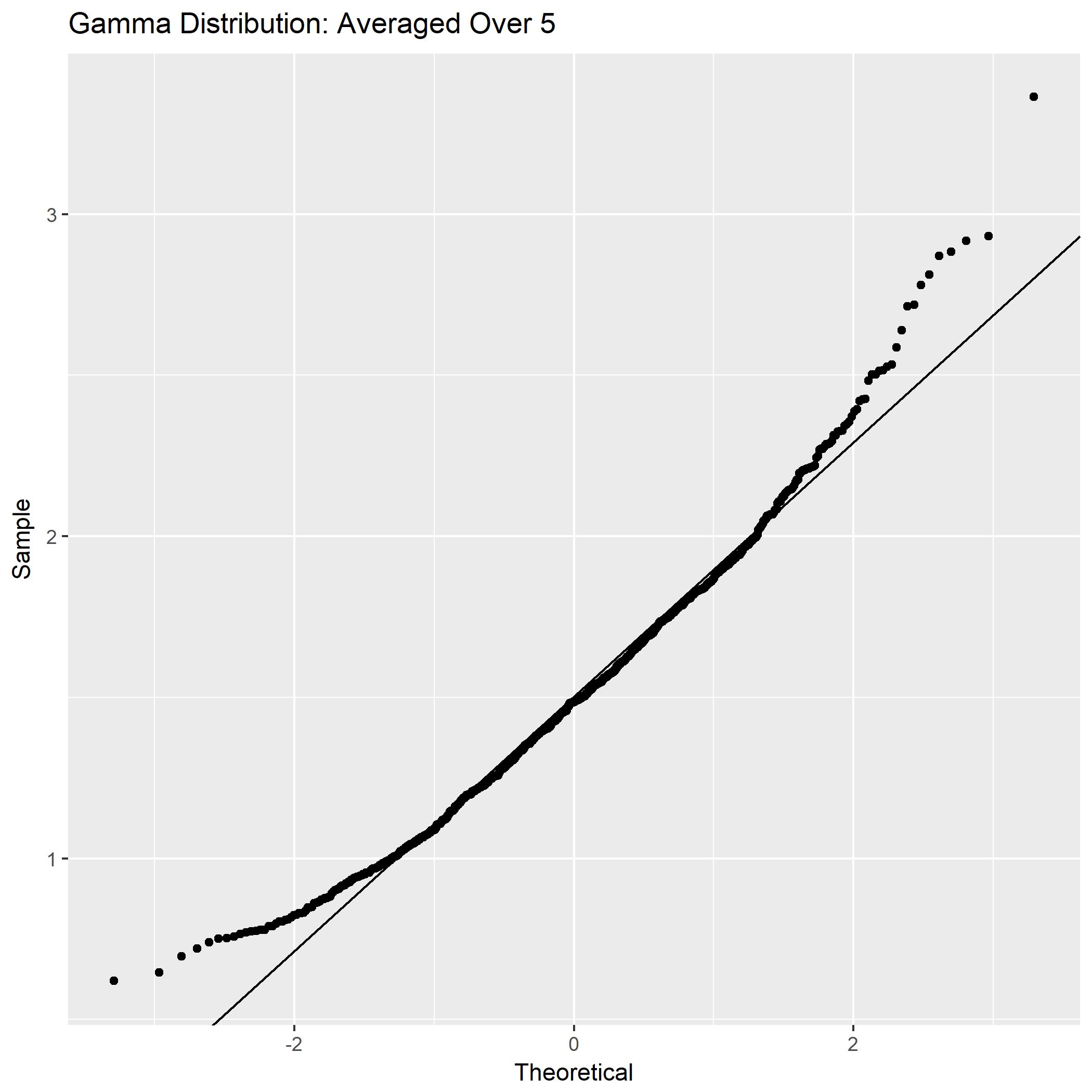
**2. Histogram/normal probability plots**



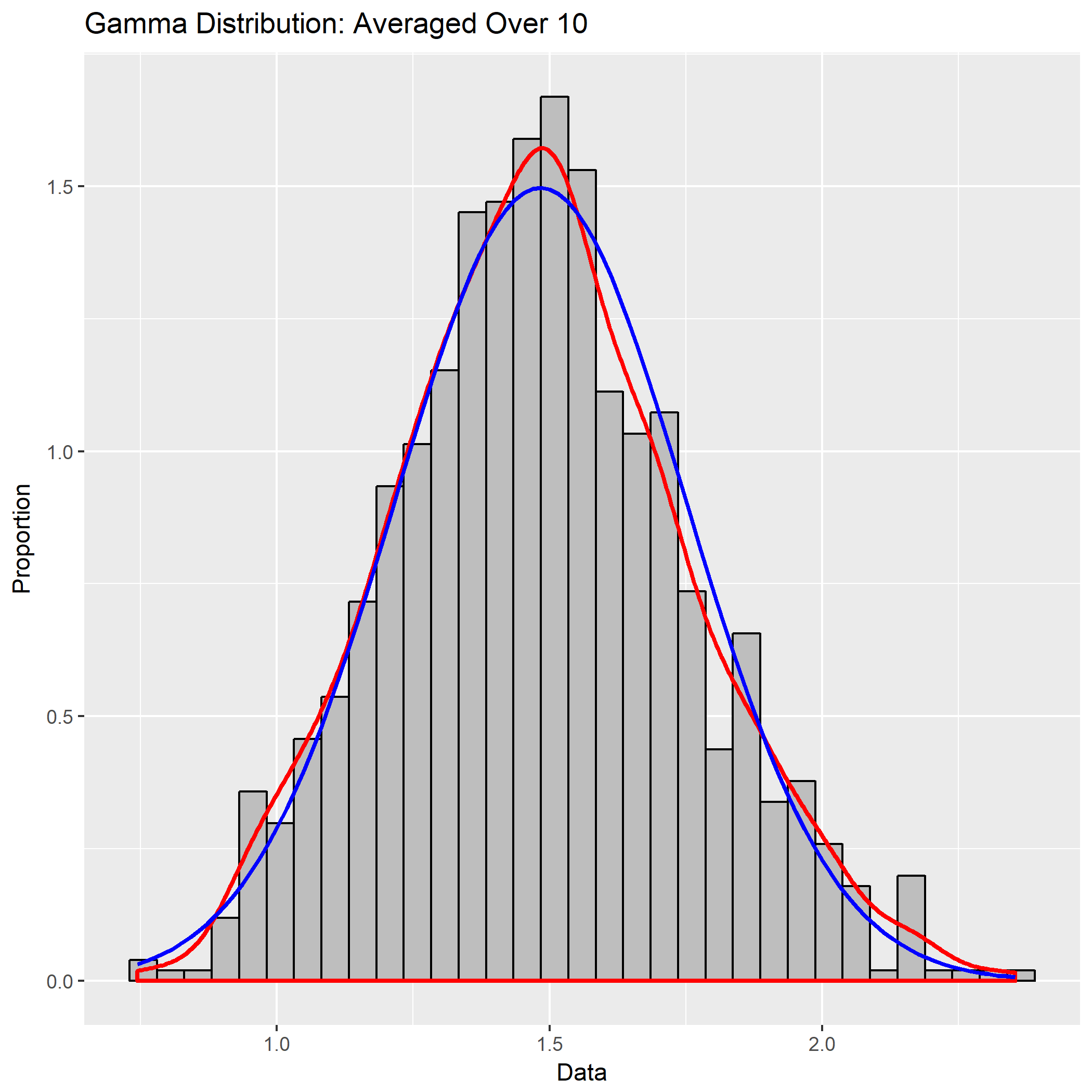


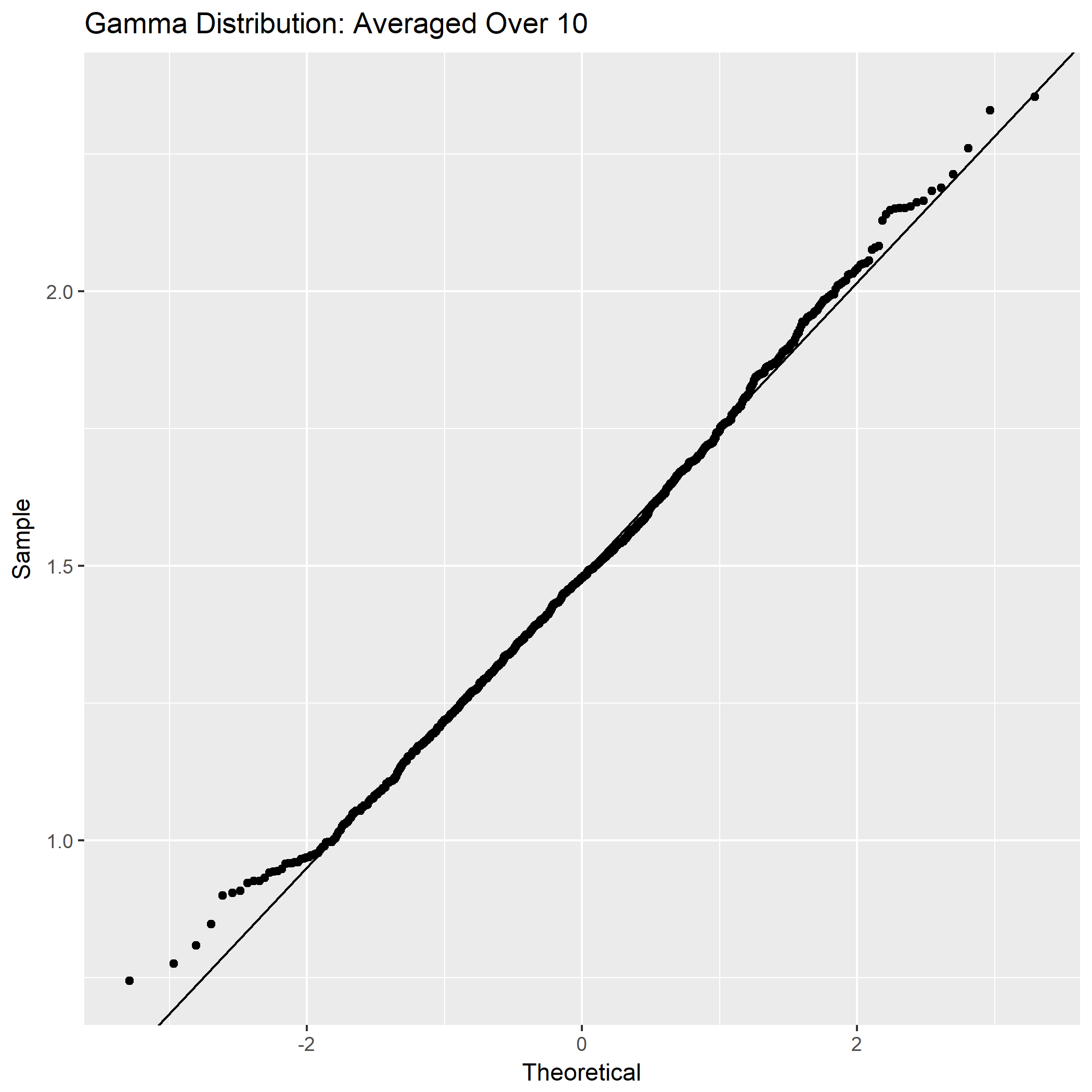
Not sufficiently normal



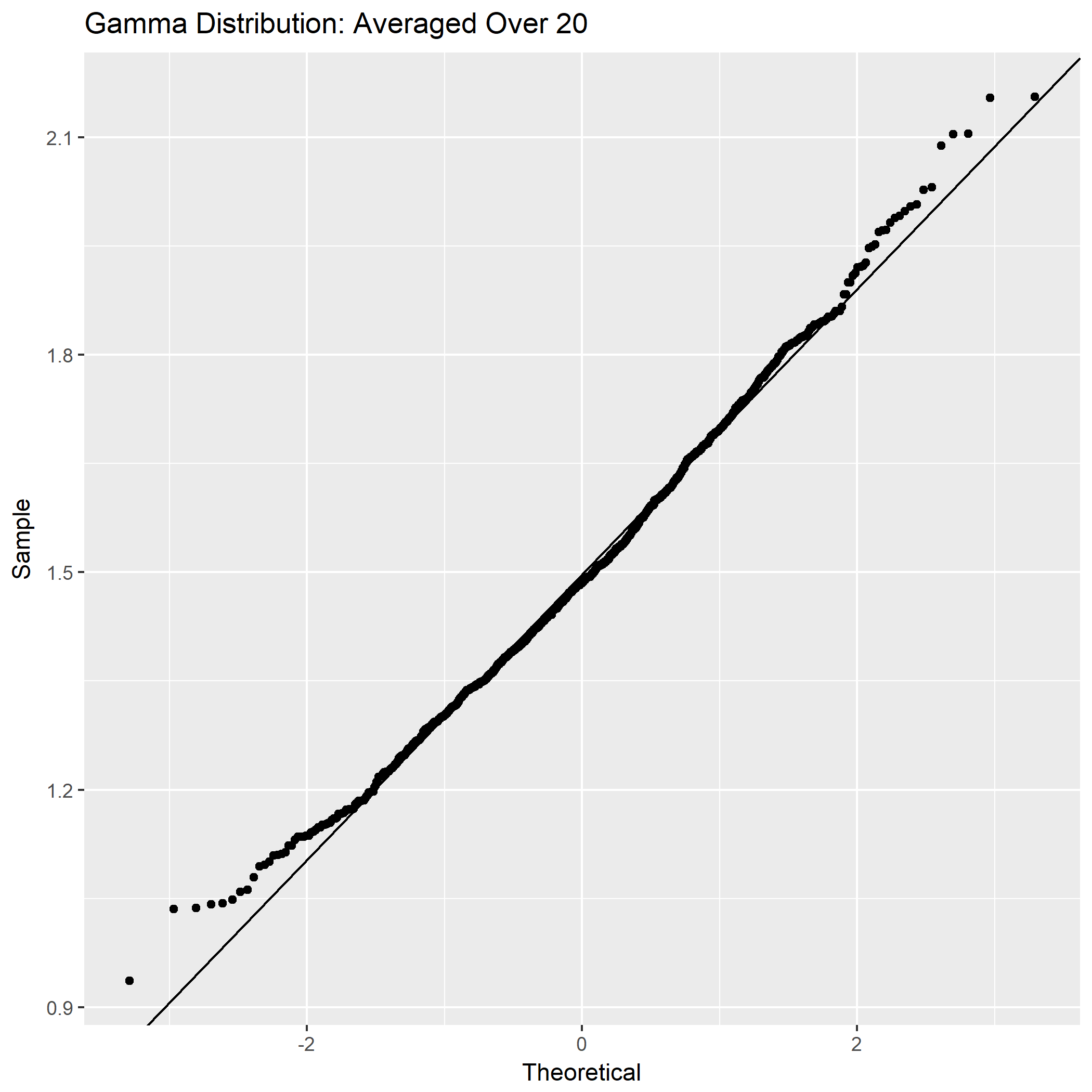
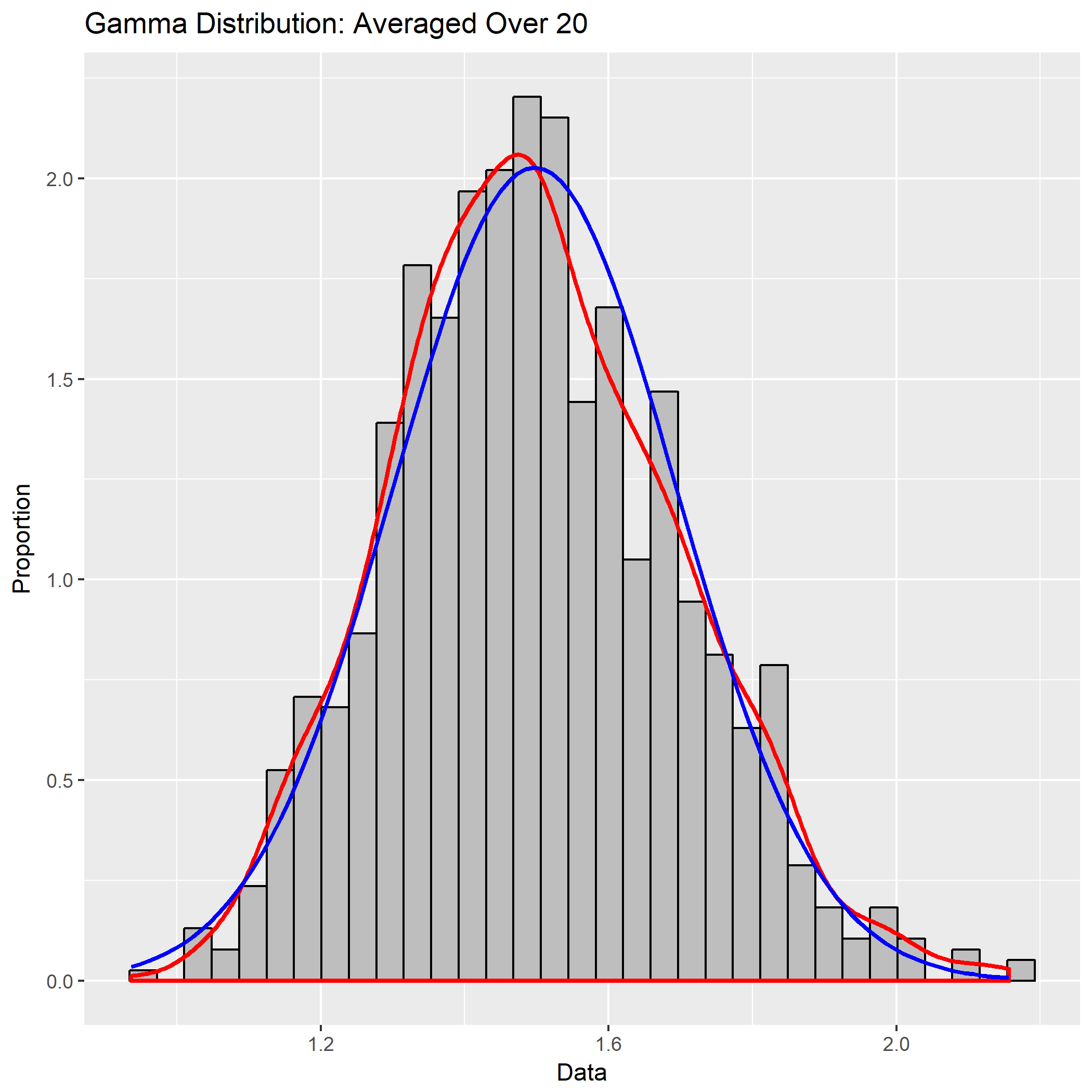


Not sufficiently normal

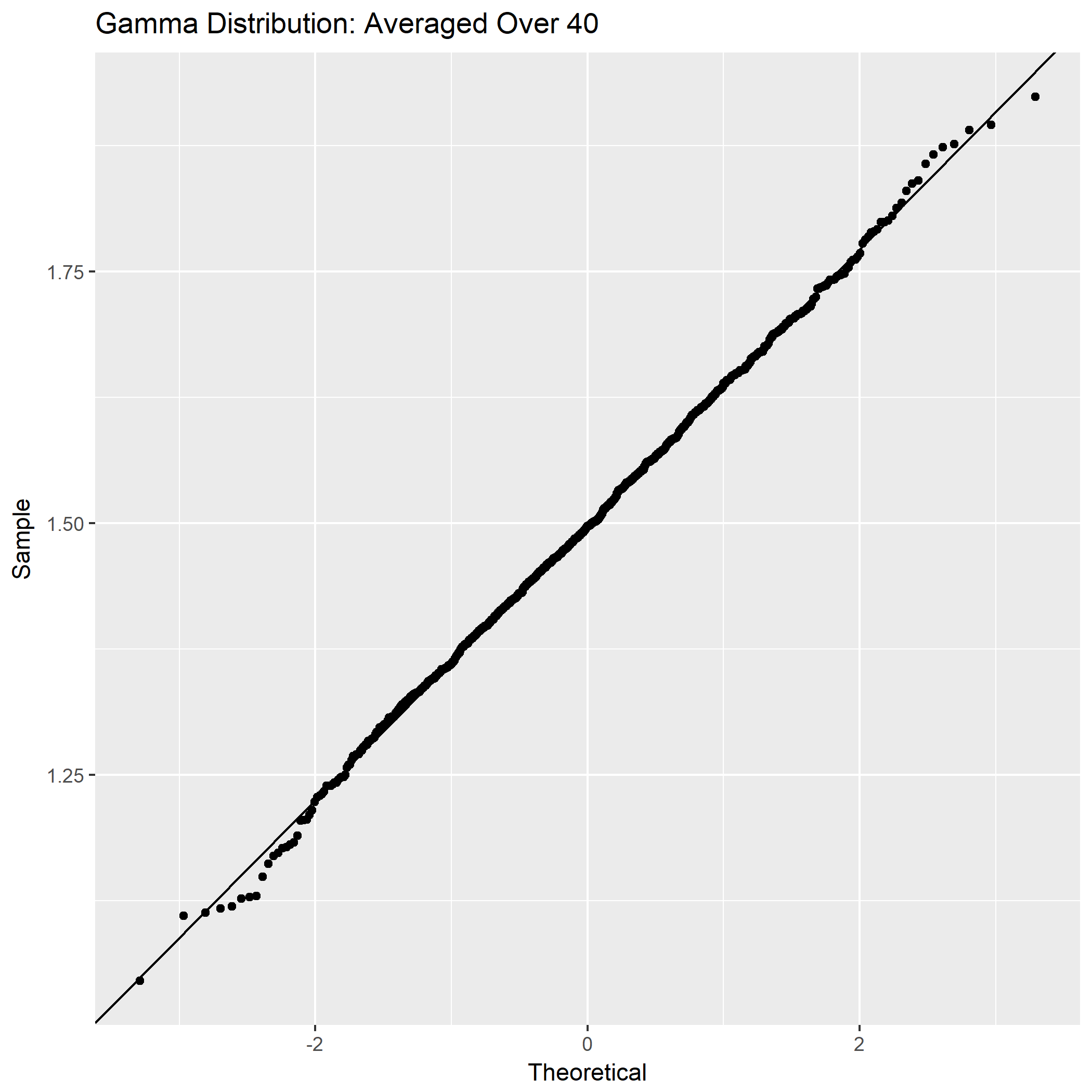
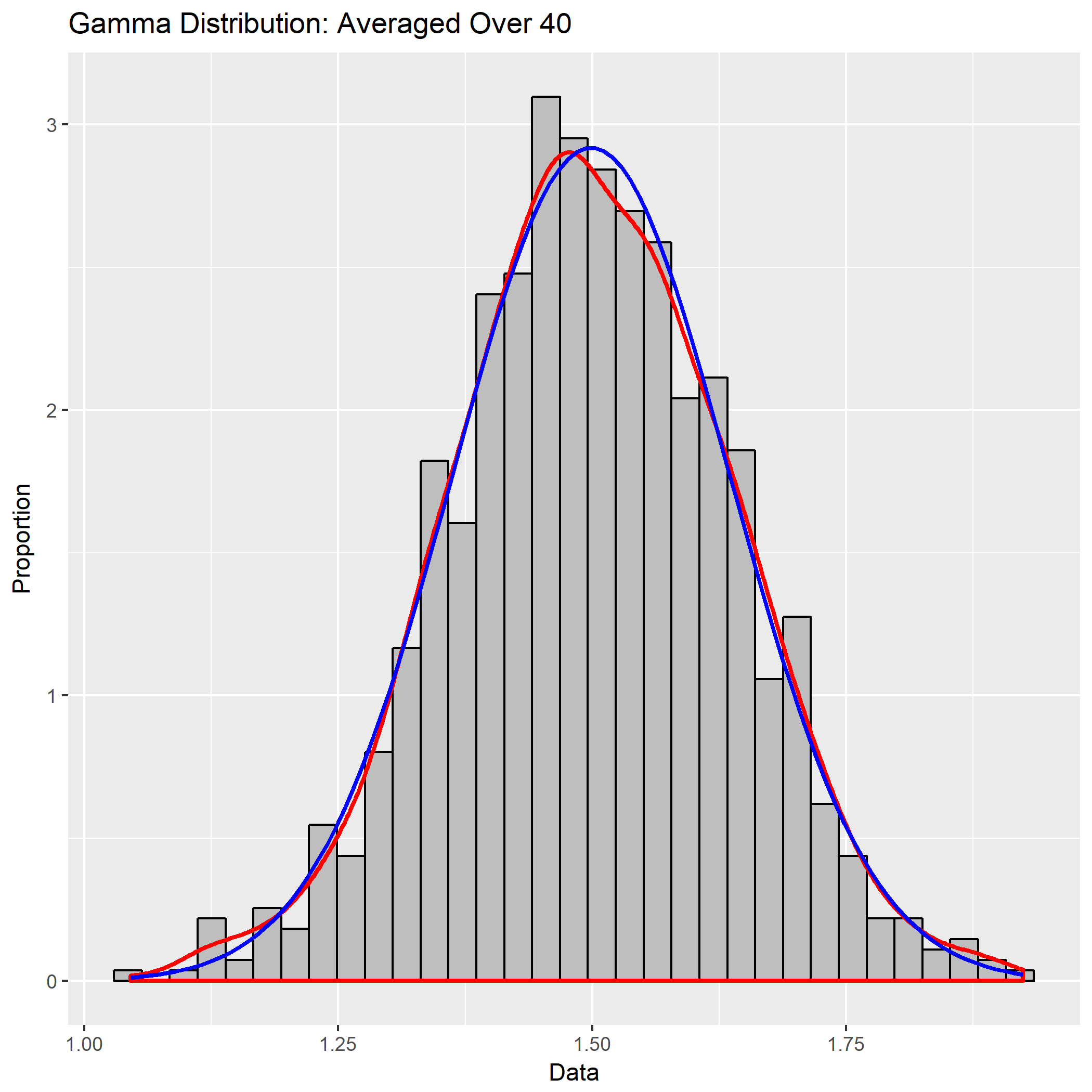




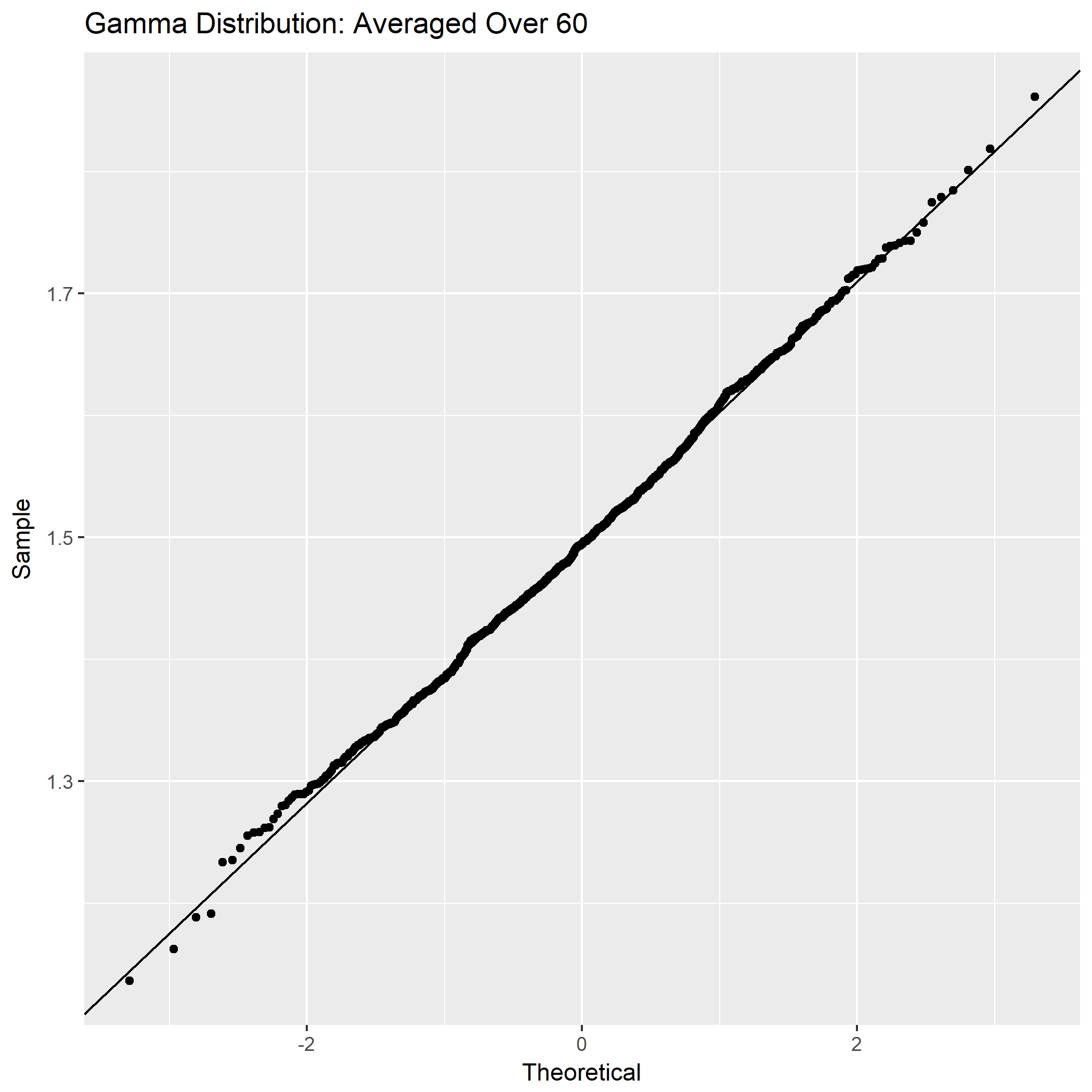
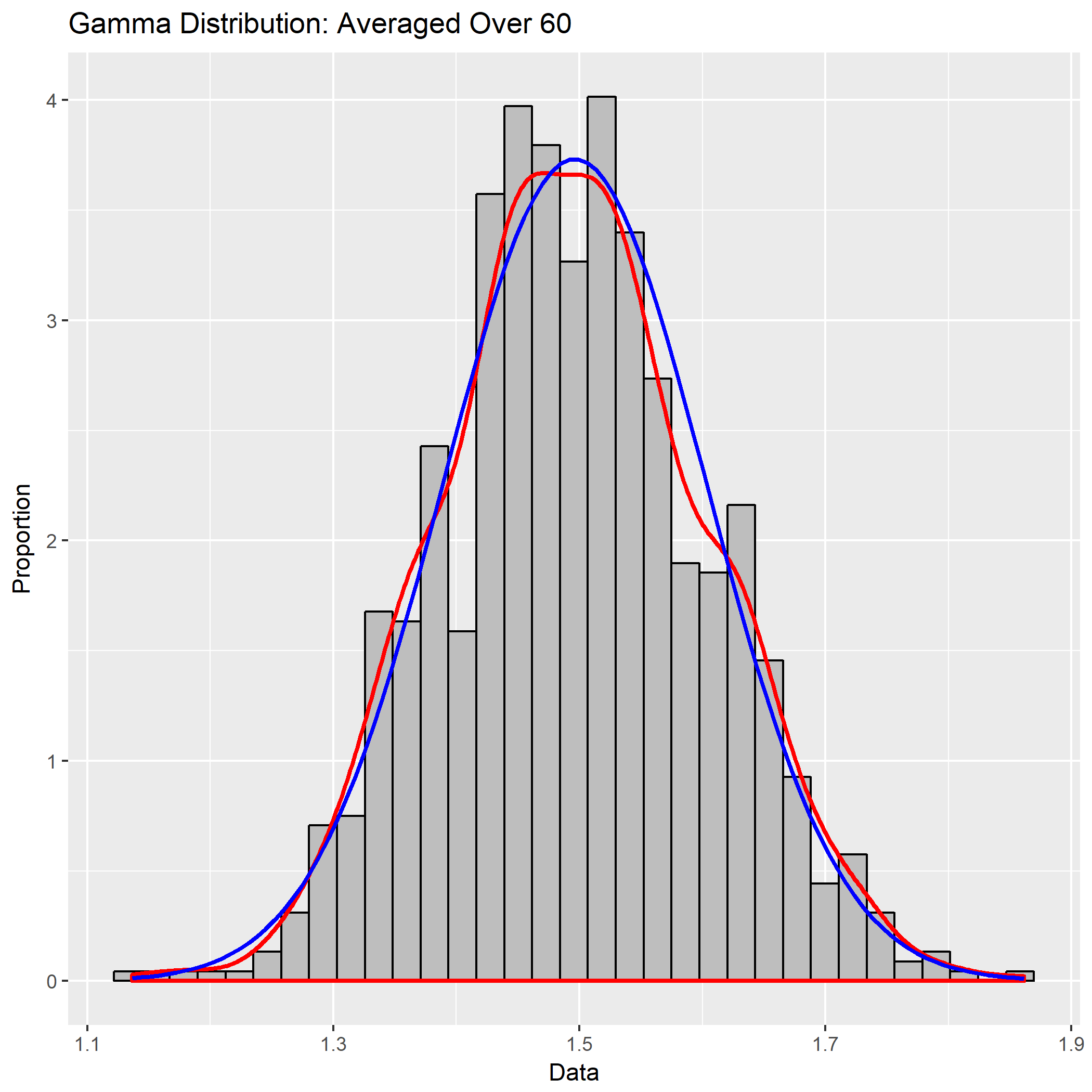
Not sufficiently normal



Not sufficiently normal



Not sufficiently normal



Approximately normal

**3. Summary table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***n*** | **experimental mean of 1000 (from output)** | **theoretical mean (given)** | **experimental standard deviation of 1000 (from output)** | **theoretical standard deviation (given and equation)** |
| 1 | 1.4934 | 1.5 | 0.863 | 0.8660 |
| 5 | 1.4946 | 1.5 | 0.383 | 0.3873 |
| 10 | 1.5076 | 1.5 | 0.2721 | 0.2739 |
| 20 | 1.4951 | 1.5 | 0.1899 | 0.1936 |
| 40 | 1.5027 | 1.5 | 0.1354 | 0.1369 |
| 60 | 1.5018 | 1.5 | 0.1169 | 0.1118 |