rain_gauge2.R

Qian

Sun Oct 18 21:45:59 2015

```
library(stringr)
theFiles<-dir("/Users/Qian/Desktop/rain gauge/",pattern="\\.txt")
theFiles
```

```
##  [1] "L.00.01.txt" "L.00.02.txt" "L.00.03.txt" "L.00.04.txt" "L.00.05.txt"
##  [6] "L.00.06.txt" "L.00.07.txt" "L.00.08.txt" "L.00.09.txt" "L.00.10.txt"
## [11] "L.00.11.txt" "L.00.12.txt" "L.01.01.txt" "L.01.02.txt" "L.01.03.txt"
## [16] "L.01.04.txt" "L.01.05.txt" "L.01.06.txt" "L.01.07.txt" "L.01.08.txt"
## [21] "L.01.09.txt" "L.01.10.txt" "L.01.11.txt" "L.01.12.txt" "L.02.01.txt"
## [26] "L.02.02.txt" "L.02.03.txt" "L.02.04.txt" "L.02.05.txt" "L.02.06.txt"
## [31] "L.02.07.txt" "L.02.08.txt" "L.02.09.txt" "L.02.10.txt" "L.02.11.txt"
## [36] "L.02.12.txt" "L.03.01.txt" "L.03.02.txt" "L.03.03.txt" "L.03.04.txt"
## [41] "L.03.05.txt" "L.03.06.txt" "L.03.07.txt" "L.03.08.txt" "L.03.09.txt"
## [46] "L.03.10.txt" "L.03.11.txt" "L.03.12.txt" "L.04.01.txt" "L.04.02.txt"
## [51] "L.04.03.txt" "L.04.04.txt" "L.04.05.txt" "L.04.06.txt" "L.04.07.txt"
## [56] "L.04.08.txt" "L.04.09.txt" "L.04.10.txt" "L.04.11.txt" "L.04.12.txt"
```

```
for (a in theFiles){
  nameToUse<-str_sub(string=a,start=1,end=7)
  temp<-read.csv(file=file.path("/Users/Qian/Desktop/rain gauge/",a),skip=2,stringsAsFactors = F)
  assign(x=nameToUse,value=temp)
}

rain <- rbind(L.00.01,L.00.02,L.00.03,L.00.04,L.00.05,L.00.06,L.00.07,L.00.08,L.00.09,L.00.10,L.00.11,L
            L.01.01,L.01.02,L.01.03,L.01.04,L.01.05,L.01.06,L.01.07,L.01.08,L.01.09,L.01.10,L.01.11,L.0
            L.02.01,L.02.02,L.02.03,L.02.04,L.02.05,L.02.06,L.02.07,L.02.08,L.02.09,L.02.10,L.02.11,L.0
            L.03.01,L.03.02,L.03.03,L.03.04,L.03.05,L.03.06,L.03.07,L.03.08,L.03.09,L.03.10,L.03.11,L.0
            L.04.01,L.04.02,L.04.03,L.04.04,L.04.05,L.04.06,L.04.07,L.04.08,L.04.09,L.04.10,L.04.11,L.0

dim(rain)
```

```
## [1] 1827   25
```

```
colnames(rain) <- 0:24
head(rain)
```

```
##   0    1    2    3    4    5    6    7    8    9   10   11   12   13   14
## 1 1 ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
## 2 2 ---- ---- ---- ---- ---- ---- ---- ---- ---- ----  .03 T    ---- ----
## 3 3 ---- ---- ---- T    ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
## 4 4 ----  .01 T    T    T    T    ---- ---- ----  .01 T    T    T    T
## 5 5  .13  .07  .03 ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
## 6 6 ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
##     15   16   17   18   19   20   21   22   23   24
## 1 ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
```

```
## 2 ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
## 3 ---- ---- T    T     .02  .01 T    T    T    ----
## 4 ---- ---- T    T     .02  .03  .12  .21  .16   .2
## 5 ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
## 6 ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
```

```
rain[rain=="----"] <- 0
rain[rain=="M   "] <- 0
rain[rain=="M"] <- 0
rain[rain=="T   "] <- 10^(-8)
head(rain)
```

```
##   0   1   2     3     4     5     6 7 8 9  10    11    12    13    14 15
## 1 1   0   0     0     0     0     0 0 0 0   0     0     0     0     0  0
## 2 2   0   0     0     0     0     0 0 0 0   0   .03 1e-08     0     0  0
## 3 3   0   0     0 1e-08     0     0 0 0 0   0     0     0     0     0  0
## 4 4   0 .01 1e-08 1e-08 1e-08 1e-08 0 0 0 .01 1e-08 1e-08 1e-08 1e-08  0
## 5 5 .13 .07   .03     0     0     0 0 0 0   0     0     0     0     0  0
## 6 6   0   0     0     0     0     0 0 0 0   0     0     0     0     0  0
##   16    17    18  19  20    21    22    23 24
## 1  0     0     0   0   0     0     0     0  0
## 2  0     0     0   0   0     0     0     0  0
## 3  0 1e-08 1e-08 .02 .01 1e-08 1e-08 1e-08  0
## 4  0 1e-08 1e-08 .02 .03   .12   .21   .16 .2
## 5  0     0     0   0   0     0     0     0  0
## 6  0     0     0   0   0     0     0     0  0
```

```
r01<-rain[,(2:25)]
head(r01)
```

```
##     1   2     3     4     5     6 7 8 9  10    11    12    13    14 15 16
## 1   0   0     0     0     0     0 0 0 0   0     0     0     0     0  0  0
## 2   0   0     0     0     0     0 0 0 0   0   .03 1e-08     0     0  0  0
## 3   0   0     0 1e-08     0     0 0 0 0   0     0     0     0     0  0  0
## 4   0 .01 1e-08 1e-08 1e-08 1e-08 0 0 0 .01 1e-08 1e-08 1e-08 1e-08  0  0
## 5 .13 .07   .03     0     0     0 0 0 0   0     0     0     0     0  0  0
## 6   0   0     0     0     0     0 0 0 0   0     0     0     0     0  0  0
##      17    18  19  20    21    22    23 24
## 1     0     0   0   0     0     0     0  0
## 2     0     0   0   0     0     0     0  0
## 3 1e-08 1e-08 .02 .01 1e-08 1e-08 1e-08  0
## 4 1e-08 1e-08 .02 .03   .12   .21   .16 .2
## 5     0     0   0   0     0     0     0  0
## 6     0     0   0   0     0     0     0  0
```

```
bos <- as.data.frame(sapply(r01, as.numeric))
bosrain<-bos[complete.cases(bos), ]



brain <- as.vector(t(bosrain))
```

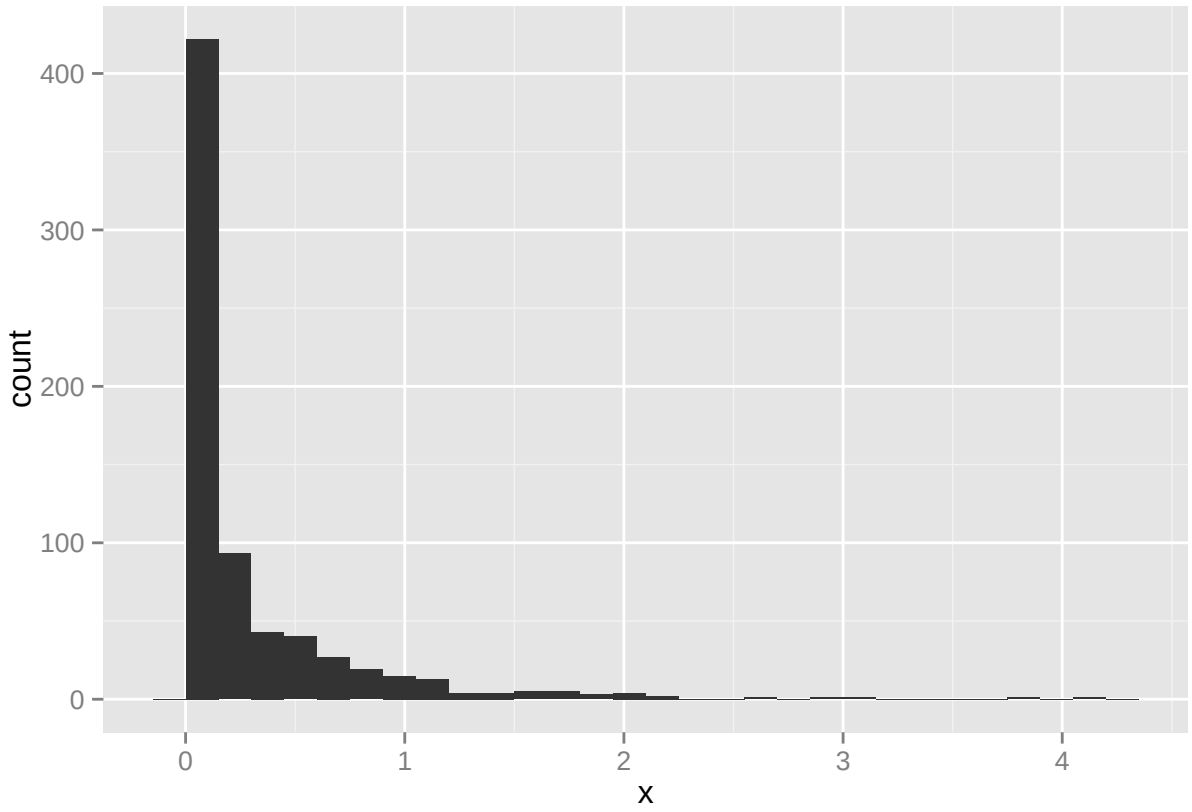
```
sum <- 0
j=1
vector<-0
for(i in 1:length(brain))
{
  if(brain[i] != 0)
  {
    sum=sum+brain[i]
  }
  if(brain[i]==0 && sum!=0)
  {
    vector[j]=sum
    j=j+1
    sum=0
  }
  if(brain[i]!=0 & i==length(brain))
  {
    vector[j]=sum
  }
}

## in order to delete those T without surrounding by numbers, we choose to keep only two digits parts
vector1<-round(vector, 2)
v2<-vector1[vector1 != 0.00]

class(v2)
```

```
## [1] "numeric"
```

```
logan <- data.frame(v2)
colnames(logan) <- "x"
library(ggplot2)
qplot(x, data=logan, geom = "histogram",binwidth=.15)
```




```
## it looks like gamma distribution

mean(logan$x)
```

```
## [1] 0.2831108
```

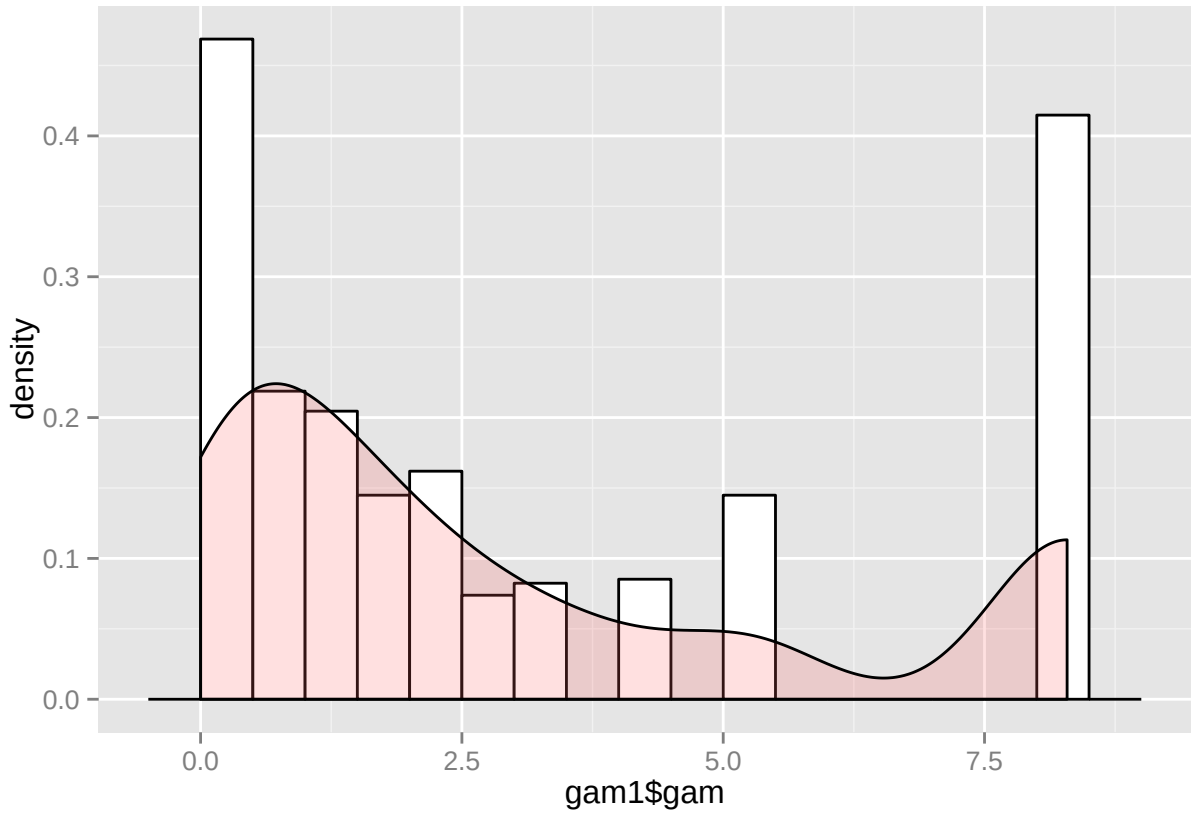
```
var(logan$x)
```

```
## [1] 0.2218382
```

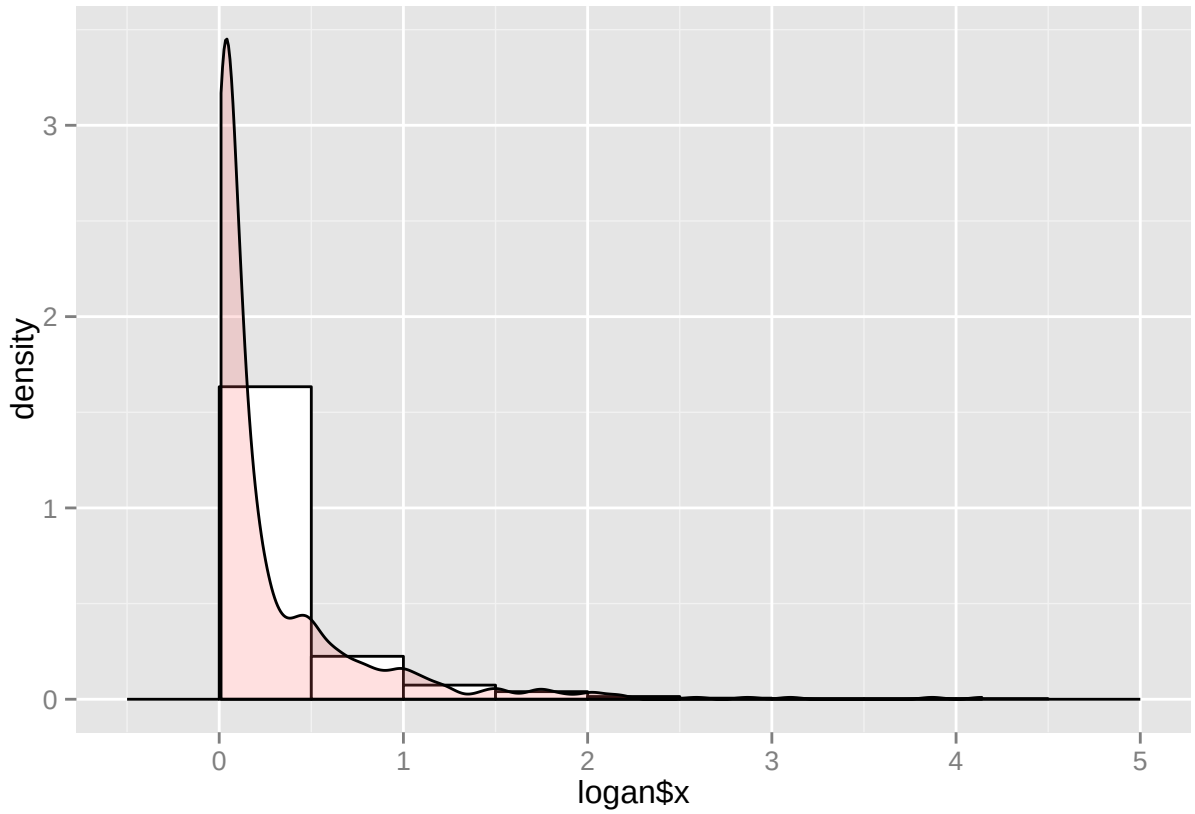
```
alpha <- mean(logan$x)^2/var(logan$x)  # alpha = 0.36
lambda <- mean(logan$x)/var(logan$x)   # lambda = 1.28


gam<-(lambda^(alpha)/gamma(alpha))*(logan$x^(alpha-1))*exp(-lambda*logan$x)
gam1<-data.frame(gam)


## gamma distribution density plot
library(ggplot2)
ggplot(gam1, aes(x=gam1$gam)) +
  geom_histogram(aes(y=..density..),      # Histogram with density instead of count on y-axis
                 binwidth=.5,
                 colour="black", fill="white") +
  geom_density(alpha=.2, fill="#FF6666")  # Overlay with transparent density plot
```




```
## data distribution density plot
ggplot(logan, aes(x=logan$x)) +
  geom_histogram(aes(y=..density..),      # Histogram with density instead of count on y-axis
                 binwidth=.5,
                 colour="black", fill="white") +
  geom_density(alpha=.2, fill="#FF6666")
```




```
# for Variance & confidence interval

## using MEM
lam<-mean(logan$x)/(sd(logan$x)^2)
alp<-(mean(logan$x))^2/(sd(logan$x)^2)

B<-1000
Tboot1<-rep(0,B)
Tboot2<-rep(0,B)
for(i in 1:B){
  x <- sample(logan$x,1000,replace=TRUE)
  Tboot1[i] <- mean(x)/(sd(x)^2)
  Tboot2[i] <- (mean(x))^2/(sd(x)^2)
}

Percentile1 <- c(quantile(Tboot1,.025),quantile(Tboot1,.975))
pivotal1 <- c((2*lam - quantile(Tboot1, .975)),(2*lam - quantile(Tboot1, .025)))

cat("Method       95% Interval\n")
```

```
## Method       95% Interval
```

```
cat("Pivotal1     (", pivotal1[1], ",     ", pivotal1[2], ") \n")
```

```
## Pivotal1     ( 0.9837294 ,      1.478039 )
```

```r
cat("Percentile1  (", Percentile1[1], ",    ", Percentile1[2], ") \n")
```

```
## Percentile1  ( 1.074369 ,     1.568678 )
```

```r
Percentile2 <- c(quantile(Tboot2,.025),quantile(Tboot2,.975))
pivotal2 <- c((2*alp - quantile(Tboot2, .975)),(2*alp - quantile(Tboot2, .025)))

cat("Method       95% Interval\n")
```

```
## Method       95% Interval
```

```r
cat("Pivotal2     (", pivotal2[1], ",     ", pivotal2[2], ") \n")
```

```
## Pivotal2     ( 0.2907355 ,      0.4094917 )
```

```r
cat("Percentile2  (", Percentile2[1], ",    ", Percentile2[2], ") \n")
```

```
## Percentile2  ( 0.3131224 ,     0.4318786 )
```

```r
##  for MLE method

mle.x <- logan$x
n <- length(logan$x)
# first we need to have alpha and lambda from MEM
mem.alp <- mean(mle.x)^2/var(mle.x)
mem.lam <- (mean(mle.x))/var(mle.x)
mem.alp
```

```
## [1] 0.3613071
```

```r
mem.lam
```

```
## [1] 1.276204
```

```r
# second we use MLE to get parameter value
minus.likelihood <- function(theta) {-(n*theta[1]*log(theta[2])-n*lgamma(theta[1])+(theta[1]-1)*sum(log

max.likelihood <- nlminb(start=c(mem.alp, mem.lam), obj = minus.likelihood)

max.likelihood$par
```

```
## [1] 0.5461541 1.9291179
```