exp-poi-gam_1_.R

jiayuan

Sun Oct 18 16:48:05 2015

```
library(ggplot2)
library(qualityTools)

#create a vector of w exponential waiting times with lambda = lam
wait <- function(w,lam){
  set.seed(50)
  a = NULL
  for(i in 1:w){
    a = c(a,rexp(1,rate = lam))
  }
  return(a)
}
wait(5,2) #a vector of 5 exponential waiting times with lambda = 2 with set.seed(50)
```

```
## [1] 0.20872710 0.41589001 0.01316189 1.60155638 0.19991279
```

```
#create a vector of exponential waiting times which total t <= Max with lambda = lam
wait.until <- function(Max,lam){
  set.seed(50)
  time = 0
  a = NULL
  while(time < Max){
    inter = rexp(1,lam)
    a = c(a,inter)
    time = time + inter
  }
  return(a[1:(length(a)-1)])
}
wait.until(3,2)
```

```
## [1] 0.20872710 0.41589001 0.01316189 1.60155638 0.19991279 0.14634362
```

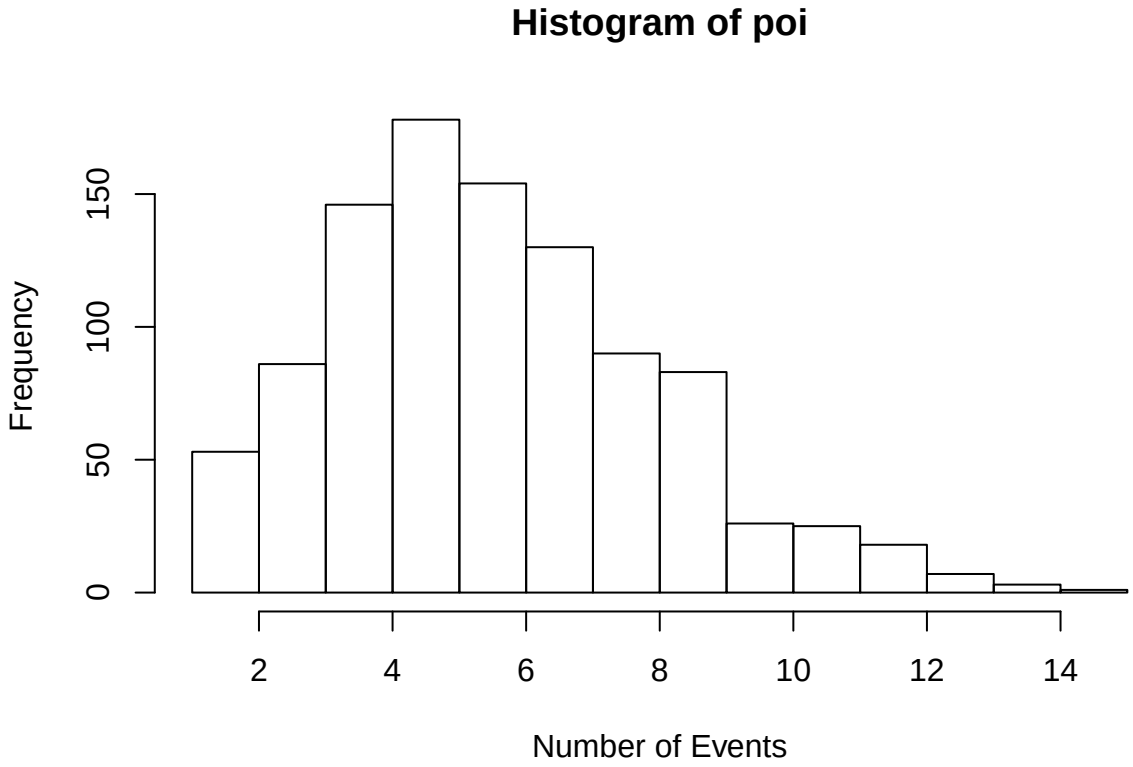
```
#a vector of exponential waiting times which total t <= 3 with lambda = 2 with set.seed(50)

#no seed
wait.until2 <- function(Max,lam){
  time = 0
  a = NULL
  while(time < Max){
    inter = rexp(1,lam)
    a = c(a,inter)
    time = time + inter
  }
  return(a[1:(length(a)-1)])
}
wait.until2(3,2)
```

```
## [1] 1.1332145 0.1408862 1.1601970 0.4011527
```

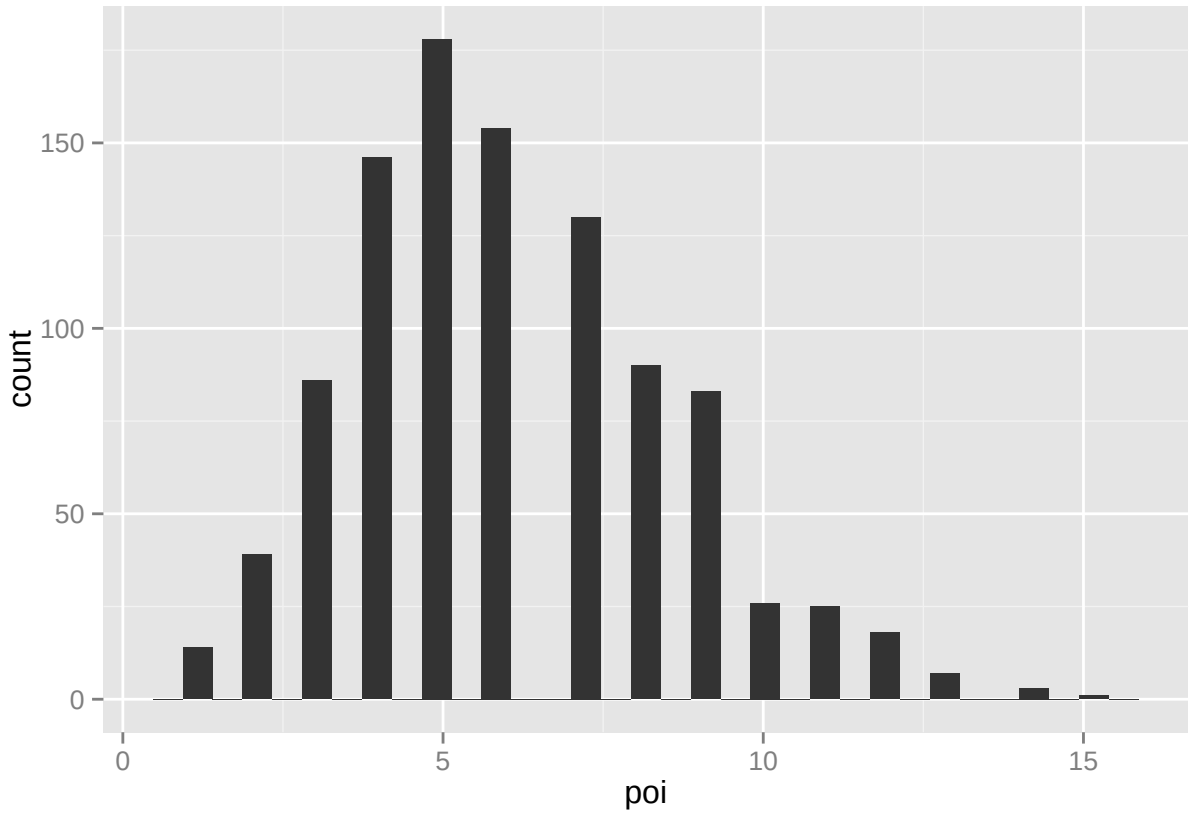
```r
#a vector of exponential waiting times which total t <= 3 with lambda = 2 without set.seed(50)

#now simulate the number of events to show that the number of events divided by
#exponential waiting times are Poisson distributed
poi.test <- function(rep, Max, lam){
  a = NULL
  for(i in 1:rep){
    q = wait.until2(Max,lam)
    #use the wait.until2() function which does not set seed
    #in order to get different waiting times until the max time with lambda = lam
    a = c(a,length(q))
  }
  return(a)
}
poi <- poi.test(1000,3,2)
#the number of events until the max time is 3 with lambda is 2 when repeating for 1000 times
hist(poi, xlab="Number of Events") #hisgram to see the distribution
```




```r
qplot(poi, geom = "histogram") #qplot the hisgram
```

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```



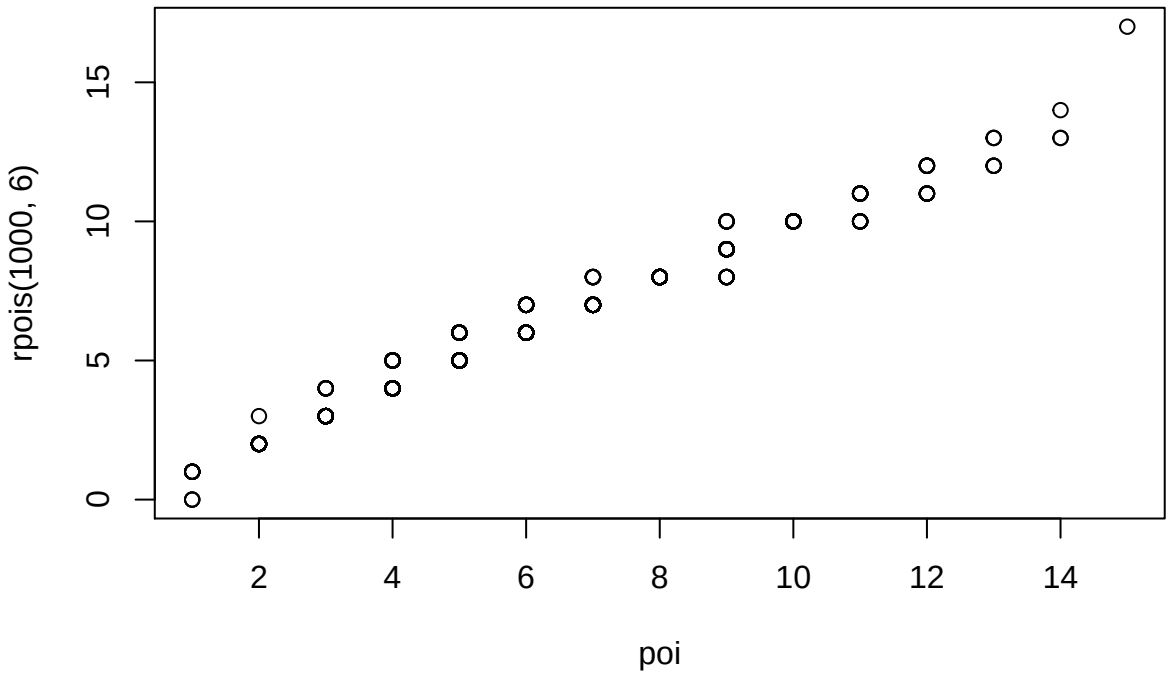

```r
mean(poi) #equals to 6=3*2=Max*lam
```

```
## [1] 6.024
```

```r
var(poi) #equals to 6=3*2=Max*lam
```

```
## [1] 6.057481
```

```r
qqplot(poi,rpois(1000,6)) #a straight line, same distribution the Poisson
```




```
mean(rpois(1000,6)) #check the mean and variance again
```

```
## [1] 6.002
```

```
var(rpois(1000,6)) #same as the mean and variance we get above
```

```
## [1] 6.177697
```

```
#number of events have a poisson distribution with parameter 6,
#which is the exponential max waiting times*lambda
#so the number of events divided by exponential waiting times are Poisson distributed also

# now simlate the total waiting time for k events to occur with lambda = lam
wait.for <- function(k, lam){
  set.seed(50)
  time = 0
  count = 0
  a = NULL
  while(count < k){
    inter=rexp(1,lam)
    count = count + 1
    time = time+inter
  }
  return(time)
}
wait.for(5,2) #the total waiting time for 5 events to occur with lambda = 2 with set.seed(50)
```

```
## [1] 2.439248
```

```r
sum(wait(5,2)) #same as the sum of interval waiting time, using the function wait() above
```

```
## [1] 2.439248
```

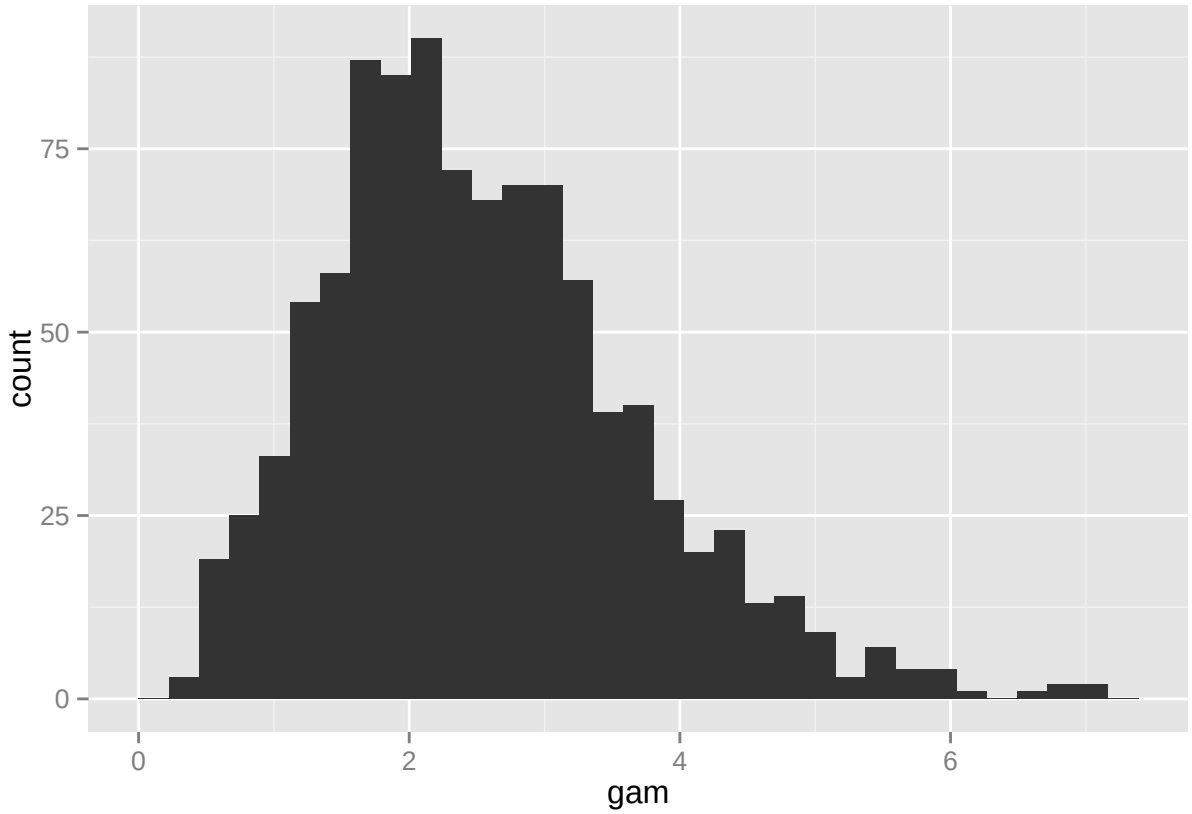
```r
#no seed
wait.for2 <- function(k, lam){
  time = 0
  count = 0
  a = NULL
  while(count < k){
    inter=rexp(1,lam)
    count = count + 1
    time = time+inter
  }
  return(time)
}
wait.for2(5,2) #the total waiting time for 5 events to occur with lambda = 2 without set.seed(50)
```

```
## [1] 4.13941
```

```r
#now simulate the total waiting time until the max.e events to occur
#to show that the total waiting time are Gamma distributed
gam.test <-function(rep, max.e, lam ){
  a=NULL
  for (i in 1:rep){
    t = wait.for2(max.e,lam)
    #use the wait.for2() function which does not set seed
    #in order to get different total waiting time until max.e events to occur with lambda = lam
    a = c(a,t)
  }
  return(a)
}
gam <- gam.test(1000,5,2)
#the total waiting time until 5 events to occur with lambda 2 when repeating for 1000 times
hist(gam, xlab="Total Waiting Time") #hisgram to see the distribution
```


```
qplot(gam, geom = "histogram") ##qplot the hisgram
```

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```



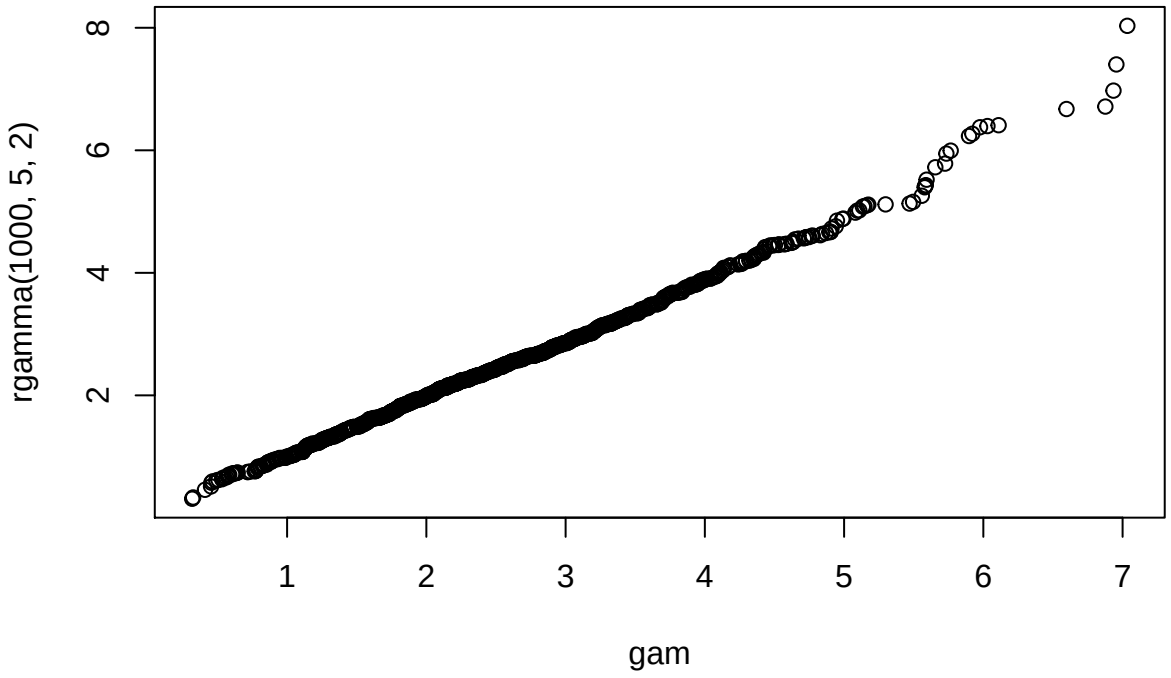
```r
mean(gam) #equals to 2.5 = 5/2 = max.e/lam
```

```
## [1] 2.537346
```

```r
var(gam) #equals to 1.25 = 5/2^2 = max.e/lam^2
```

```
## [1] 1.295177
```

```r
qqplot(gam,rgamma(1000,5,2)) #a straight line, same distribution the Gamma
```




```r
mean(rgamma(1000,5,2)) #check the mean and variance again
```

```
## [1] 2.504025
```

```r
var(rgamma(1000,5,2)) #same as the mean and variance we get above
```

```
## [1] 1.328073
```

```r
#so the total waiting time has a Gamma distribution with parameter 5,2,
#which is the number of events and lambda
```