HW5

Calvin Makelky, A01872013

September 12, 2016

1

a

```
a=10:1
append(a,1:10)
```

```
## [1] 10  9  8  7  6  5  4  3  2  1  1  2  3  4  5  6  7  8  9 10
```

b

```
b=c(0,25,50,75,100)
rep(b, c(5,4,3,2,1))
```

```
## [1]   0   0   0   0   0  25  25  25  25  50  50  50  75  75 100
```

c

```
seq(1,8,len=7)
```

```
## [1] 1.000000 2.166667 3.333333 4.500000 5.666667 6.833333 8.000000
```

d

```
rep(c(0,5,10), len = 15)
```

```
## [1]  0  5 10  0  5 10  0  5 10  0  5 10  0  5 10
```

e

```
probv = seq((.5/207), (1-(.5/207)), len=207)
head(probv)
```

```
## [1] 0.002415459 0.007246377 0.012077295 0.016908213 0.021739130 0.026570048
```

```
tail(probv)
```

```
## [1] 0.9734300 0.9782609 0.9830918 0.9879227 0.9927536 0.9975845
```

2

a

```
olympics = read.table("athletes2016.txt")
cnames = c("Country", "Athletes", "Golds", "Silvers", "Bronzes", "Medals")
colnames(olympics) = cnames
country = olympics$Country
row.names(olympics) = country
olympics$Country = NULL
head(olympics)
```

```
##                 Athletes Golds Silvers Bronzes Medals
## Afghanistan            3     0       0       0      0
## Albania                6     0       0       0      0
## Algeria               67     0       2       0      2
## American_Samoa         4     0       0       0      0
## Andorra                5     0       0       0      0
## Angola                25     0       0       0      0
```
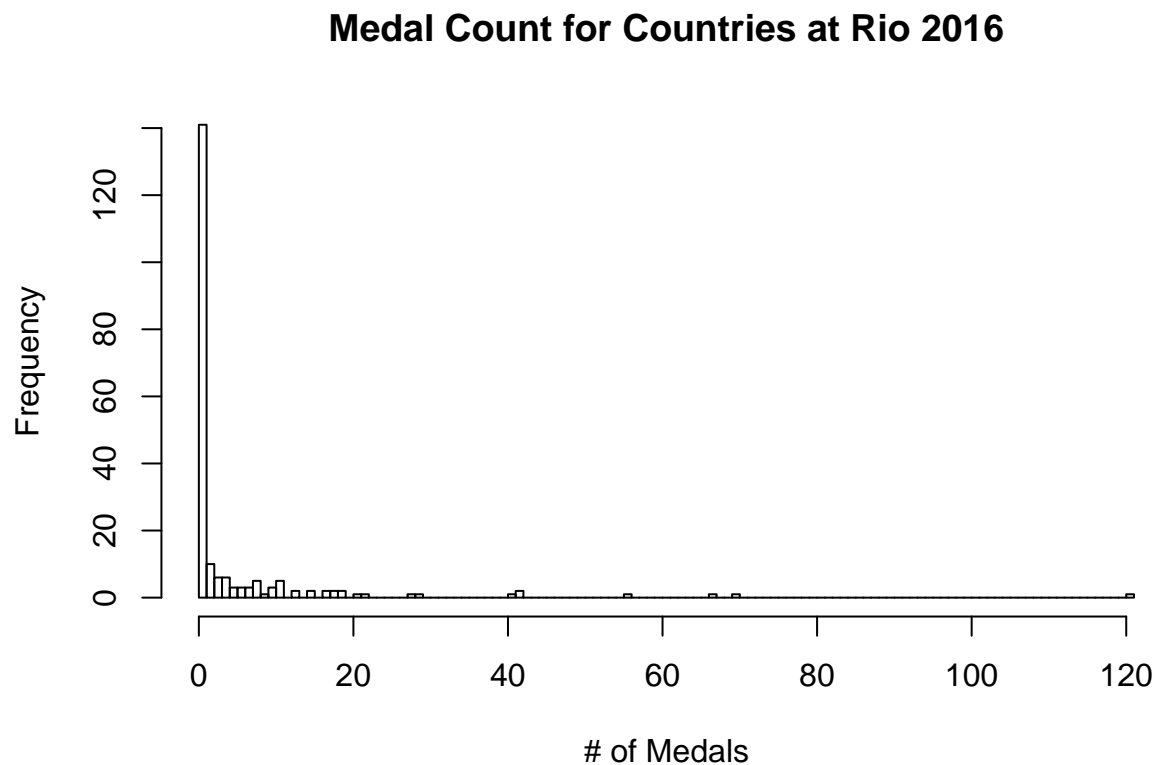
b

```
medalcount = olympics[, -(1:4), drop=FALSE]
head(medalcount)
```

```
##                 Medals
## Afghanistan          0
## Albania              0
## Algeria              2
## American_Samoa       0
## Andorra              0
## Angola               0
```

c

```
hist(medalcount$Medals, breaks=100, main="", xlab="")
title(main= "Medal Count for Countries at Rio 2016", xlab= "# of Medals")
```

Medal Count for Countries at Rio 2016

d

```
lhat = mean(medalcount$Medals)
lhat
```

```
## [1] 4.705314
```

e

```
qvect = qpois(probv, lhat)
qvect
```

```
##    [1] 0 0 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2
##   [24] 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3
##   [47] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4
```

```
## [70]    4  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4
## [93]    4  4  4  4  4  4  4  4  4  4  5  5  5  5  5  5  5  5  5  5  5  5  5
## [116]   5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## [139]   6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [162]   6  6  6  6  6  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [185]   7  8  8  8  8  8  8  8  8  8  8  8  8  9  9  9  9  9 10 10 10 11 12
```
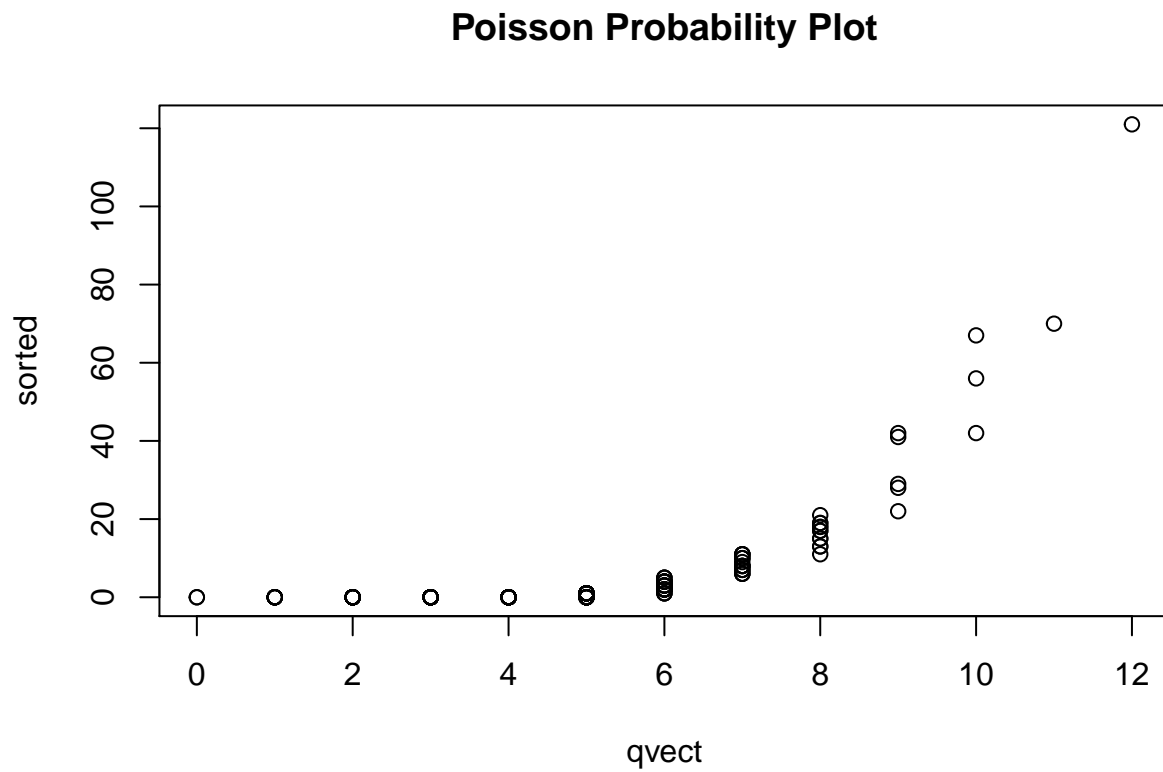
f

```
sorted = sort(medalcount$Medals)
head(sorted)
```

```
## [1] 0 0 0 0 0 0
```

```
tail(sorted)
```

```
## [1]   42   42   56   67   70  121
```

g

```
plot(qvect, sorted, main="Poisson Probability Plot")
```

Poisson Probability Plot

4

h

The medal count data does not seem to follow a Poisson distribution because the graph is not close to being a straight line. It is highly right skewed. No I'm not surprised because there are a ton of zero counts in the data and only two in the quantile vector. A better model would be the zero inflated Poisson model.