Statistical Computing 1, Stat 590 | Homework 1 | Name: Answer Key
Fall 2015 | | Prof. Erik B. Erhardt

Part I. (60 points) Do all calculations in LaTeX + R + knitr. Insert computer text output and graphics to support what you are saying. For this assignment, all R code should well commented and be visible (`echo=TRUE`) in the document where you have written it.

(30pts) **1. stats package:** Go through the help to find out what is contained in the `stats` package. Learn how to use, and then Demonstrate, 15 functions (graphical or computational) of your choice, not including functions included in problem 2.
Start by getting help by typing `?stats` and clicking on "Index" at the bottom of the help page.

Solution: Anything sensible is fine. For example:
`mean()`:

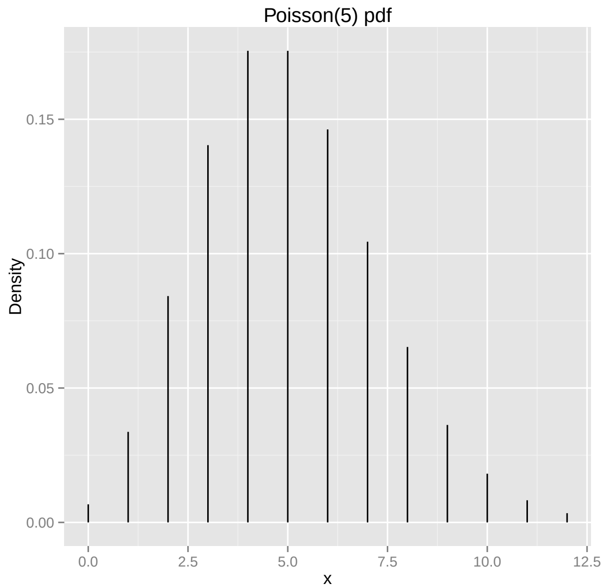
```
z <- c(rnorm(10), NA)
mean(z)

## [1] NA

mean(z, na.rm=TRUE)

## [1] 0.07705762

mean(z, na.rm=TRUE, trim=.2)

## [1] -0.0463304
```

(30pts) **2. Selected stats functions:** Demonstrate functions for probability distributions in the `stats` package. Show results numerically and graphically. For simplicity, you can choose two of your favorite probability distributions to discuss.
To get started, type `?Distributions`.

(a) (10 pts) Compute density or probability functions.

Solution: Poisson probability mass function.

```
x <- 0:12
d <- dpois(x, lambda=5)
df <- data.frame(x, d)
head(df)

##   x           d
## 1 0 0.006737947
## 2 1 0.033689735
## 3 2 0.084224337
## 4 3 0.140373896
## 5 4 0.175467370
## 6 5 0.175467370

library(ggplot2)
p <- ggplot(df, aes(x=x, y=d, ymin=0, ymax=d))
p <- p + geom_linerange()
p <- p + labs(title = "Poisson(5) pdf")
p <- p + ylab("Density")
p
```



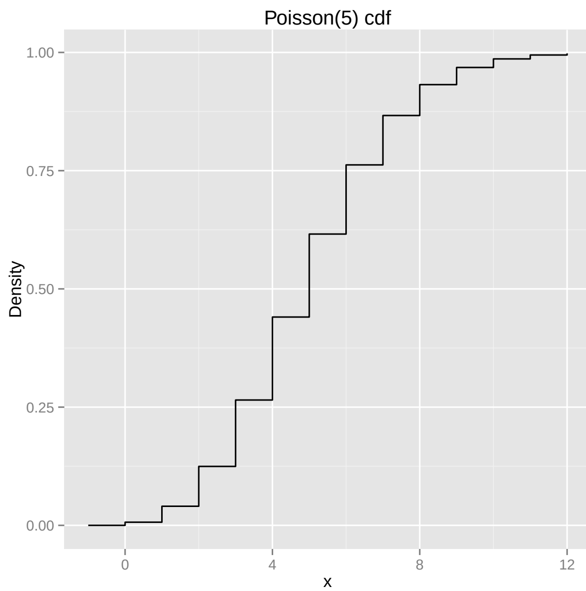
(b) (10 pts) Compute cumulative distribution functions and percentile functions.

Solution: Poisson cumulative distribution function.

```
x <- -1:12
d <- ppois(x, lambda=5)
df <- data.frame(x, d)
head(df)
```

```
##    x           d
## 1 -1 0.000000000
## 2  0 0.006737947
## 3  1 0.040427682
## 4  2 0.124652019
## 5  3 0.265025915
## 6  4 0.440493285
```

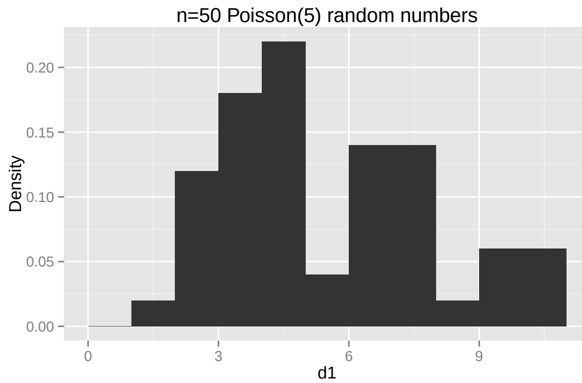
```
library(ggplot2)
p <- ggplot(df, aes(x=x, y=d))
p <- p + geom_step()
p <- p + labs(title = "Poisson(5) cdf")
p <- p + ylab("Density")
p
```

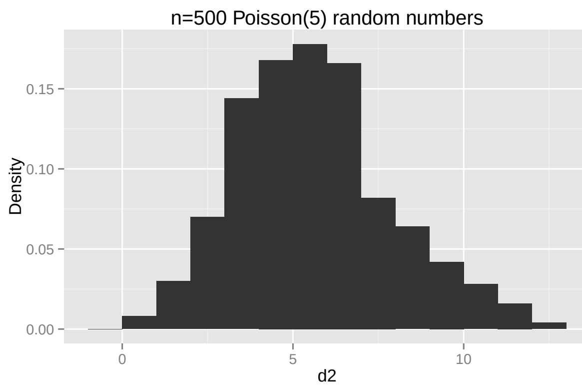


(c) (10 pts) Generate random numbers.

Solution: Poisson random numbers.

```
d1 <- rpois( 50, lambda=5)
d2 <- rpois(500, lambda=5)
df <- data.frame(d1, d2)
head(df)

##   d1 d2
## 1  2  5
## 2  3  5
## 3  4  5
## 4  4  5
## 5  8  6
## 6  1  2

library(ggplot2)
p1 <- ggplot(df, aes(x=d1))
p1 <- p1 + geom_histogram(aes(y=..density..), binwidth=1, position="identity")
p1 <- p1 + labs(title = "n=50 Poisson(5) random numbers")
p1 <- p1 + ylab("Density")

p2 <- ggplot(df, aes(x=d2))
p2 <- p2 + geom_histogram(aes(y=..density..), binwidth=1, position="identity")
p2 <- p2 + labs(title = "n=500 Poisson(5) random numbers")
p2 <- p2 + ylab("Density")

# plot all three together
library(gridExtra)
grid.arrange(p1, p2, ncol=1)
```


n=50 Poisson(5) random numbers
Density
0.20
0.15
0.10
0.05
0.00
0
3
6
9
d1


n=500 Poisson(5) random numbers
Density
0.15
0.10
0.05
0.00
0
5
10
d2