Shahab Geravesh Lab 3

Problem 1: The data is contained in a data frame called cats, in the R package MASS. (This package is part of the standard R installation.) This records the sex of each cat, its weight in kilograms, and the weight of its heart in grams. Load the data as follows: library(MASS) data(cats) Run summary(cats) and explain the results.

```
library(MASS)
data(cats)
```

```
summary(cats)
```

```
##  Sex         Bwt             Hwt       
##  F:47   Min.   :2.000   Min.   : 6.30  
##  M:97   1st Qu.:2.300   1st Qu.: 8.95  
##         Median :2.700   Median :10.10  
##         Mean   :2.724   Mean   :10.63  
##         3rd Qu.:3.025   3rd Qu.:12.12  
##         Max.   :3.900   Max.   :20.50
```

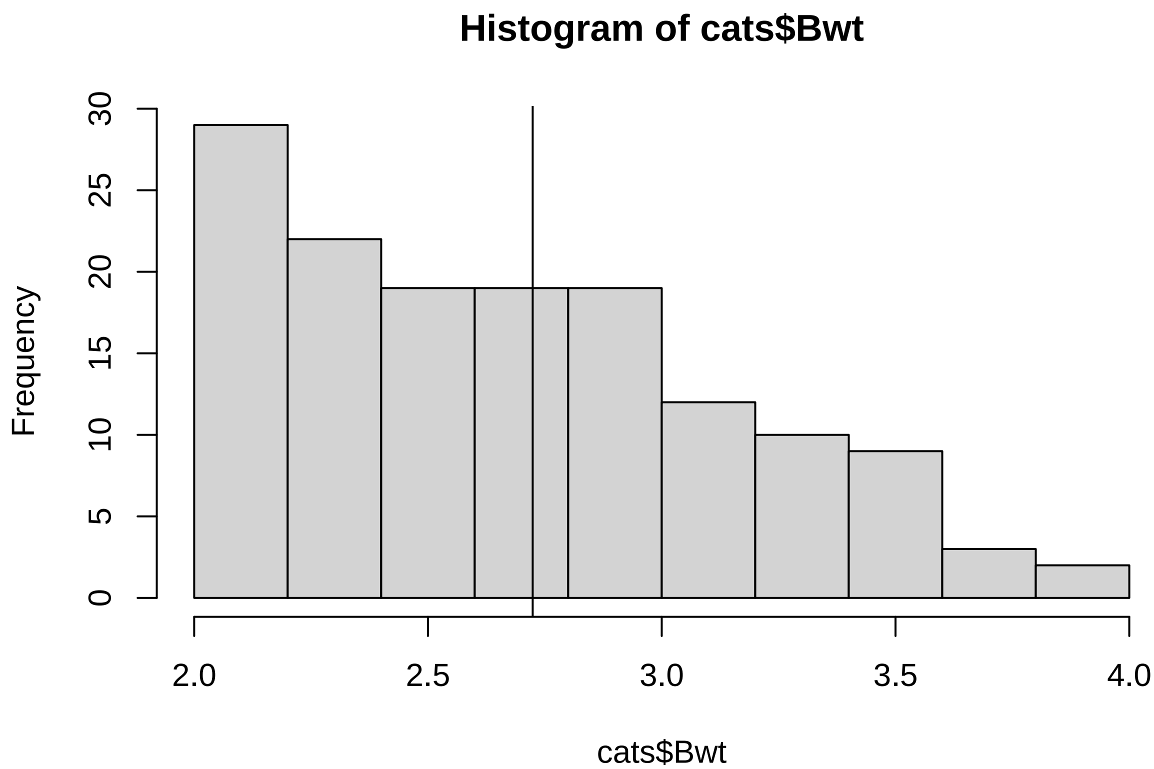
Female cats are 47 and male cats are 97. maximum weight is 3.9 kg an minimum weight is 2 kg. The maximum heart weight is 20.50 g and the minimum heart weight is 6.30 g. The average weight is 2.724 kg and the average heart weight is 10.63 g.

Problem 2: Plot a histogram of these weights using the probability=TRUE option. Add a vertical line with your calculated mean using abline(v=yourmeanvaluehere). Does this calculated mean look correct?

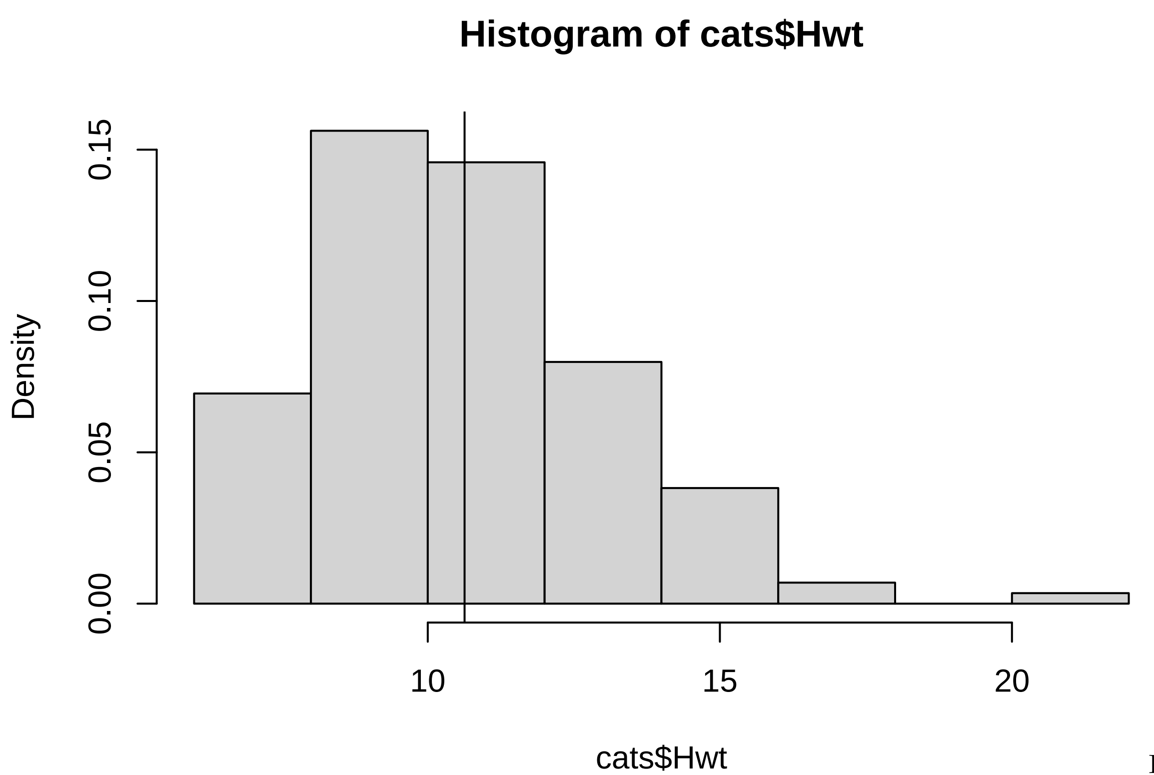
```
hist(cats$Bwt,propability=TRUE)
```

```
## Warning in plot.window(xlim, ylim, "", ...): "propability" is not a graphical
## parameter

## Warning in title(main = main, sub = sub, xlab = xlab, ylab = ylab, ...):
## "propability" is not a graphical parameter

## Warning in axis(1, ...): "propability" is not a graphical parameter

## Warning in axis(2, ...): "propability" is not a graphical parameter
```

```
abline(v=2.724)
```




```r
hist(cats$Hwt, probability=TRUE)

abline(v=10.63)
```



Here the calculated mean seems to be correct because data points are equally distributed on both sides of the mean in this histogram.

Problem 3) Define two variables, fake.mean <- 10 and fake.var <- 8. Write an expression for a using these placeholder values. Does it equal what you expected given the solutions above? Once it does, write another such expression for s and confirm.

```
fake.mean<-10

fake.var<-8
```

```
aa <- 10* (fake.mean) + 5*(fake.var)
aa
```

```
## [1] 140
```

```
ss <- 10* (fake.mean) - 5*(fake.var)
ss
```

```
## [1] 60
```

Problem 4) Calculate the mean, standard deviation, and variance of the heart weights using R's existing functions for these tasks. Plug the mean and variance of the cats' hearts into your formulas from the previous question and get estimates of a and s. What are they? Do not report them to more significant digits than is reasonable.

```
mean_hwt <- mean(cats$Hwt)
mean_hwt
```

```
## [1] 10.63056
```

```
variance_hwt <- var(cats$Hwt)
variance_hwt
```

```
## [1] 5.927451
```

```
stand_dev_hwt <- sd(cats$Hwt)
stand_dev_hwt
```

```
## [1] 2.434636
```

```
aa <- 10* (mean_hwt) + 5*(variance_hwt)
aa
```

```
## [1] 135.9428
```

```
ss <- 10* (mean_hwt) - 5*(variance_hwt)
ss
```

```
## [1] 76.6683
```

Problem 5) Write a function, cat.stats(), which takes as input a vector of numbers and returns the mean and variances of these cat hearts. (You can use the existing mean and variance functions within this function.) Confirm that you are returning the values from above.

```
cat.stats <- function(a)
{
mean <- mean(a);
variance <- var(a);
return(c(mean,variance))
}

temp <- cat.stats(cats$Hwt)

temp
```

```
## [1] 10.630556  5.927451
```

Part 2

Problem 6) Now, use your existing function as a template for a new function, gamma.cat(), that calculates the mean and variances and returns the estimate of a and s. What estimates does it give on the cats' hearts weight? Should it agree with your previous calculation?

```r
gamma.cat <- function(a)

{

mean_Hwt <- mean(a);

variance_hwt <- var(a);

a = (mean_Hwt*mean_Hwt)/variance_hwt;

s = variance_hwt/mean_Hwt;

   return(c(a,s))

}

temp <- gamma.cat(cats$Hwt)
temp
```

```
## [1] 19.0653121  0.5575862
```

Problem 7) Estimate the a and s separately for all the male cats and all the female cats, using gamma.cat(). Give the commands you used and the results.

```r
require(dplyr)
```

```
## Loading required package: dplyr

##
## Attaching package: 'dplyr'

## The following object is masked from 'package:MASS':
##
##     select

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
Female_Cats <- gamma.cat(filter(cats, Sex==("F"))$Hwt)
Female_Cats
```

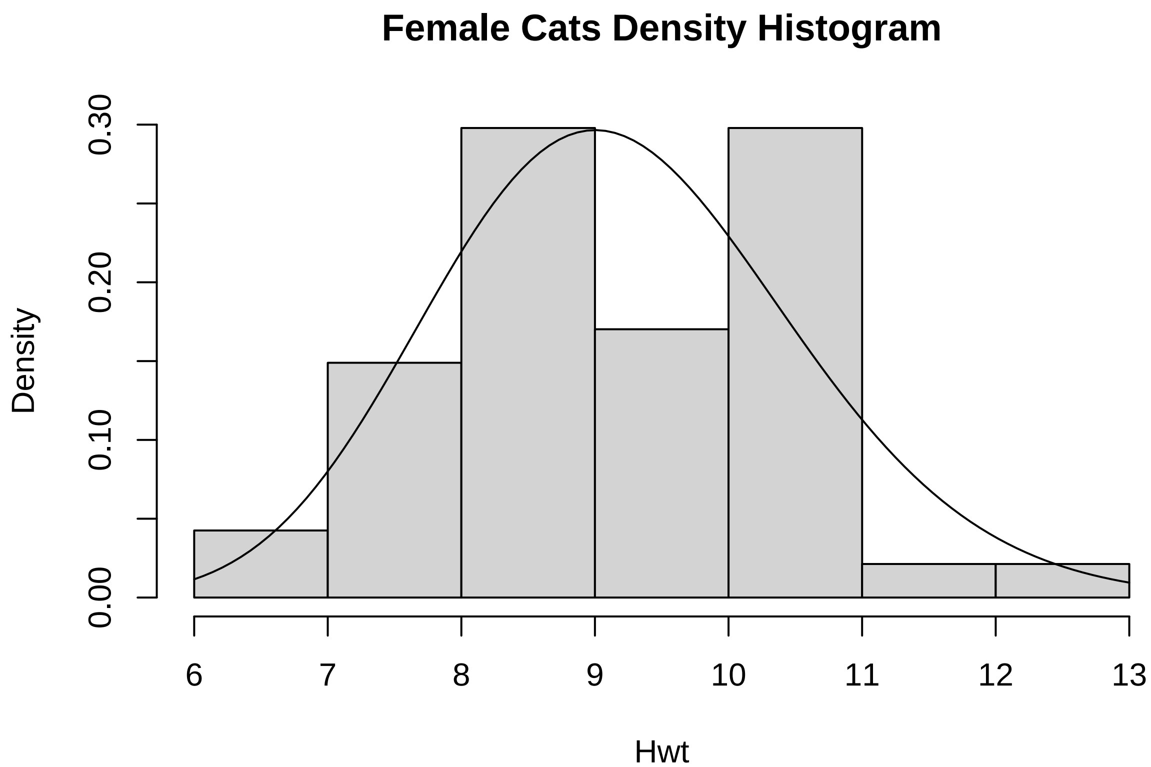
```
## [1] 45.9399792  0.2003076
```

```r
Male_Cats <- gamma.cat(filter(cats, Sex==("M"))$Hwt)
Male_Cats
```

```
## [1] 19.8357612  0.5708216
```

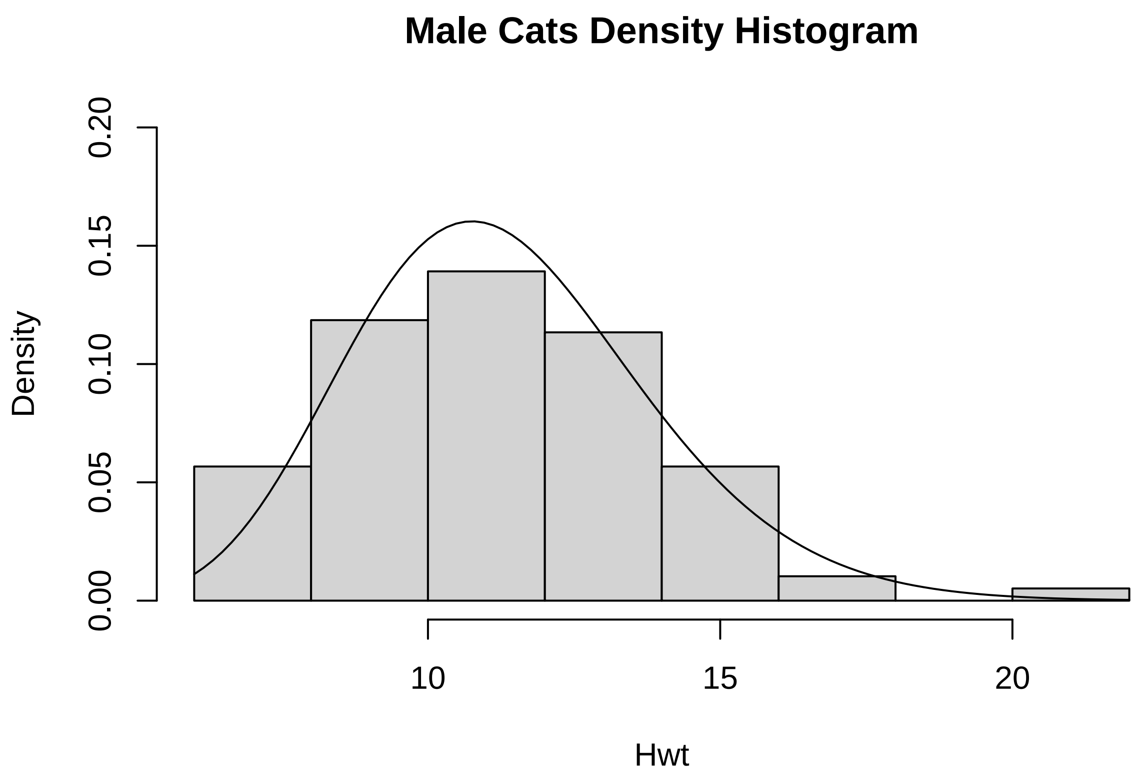
Problem 8) Now, produce a histogram for the female cats. On top of this, add the shape of the gamma PDF using curve() with its first argument as dgamma(), the known PDF for the Gamma distribution. Is this distribution consistent with the empirical probability density of the histogram?

```
hist(filter(cats,Sex==("F"))$Hwt,main = "Female Cats Density Histogram",xlab= 'Hwt',prob = T,
     ylim = c(0, 0.30))

curve(dgamma(x, shape = Female_Cats[1], scale = Female_Cats[2]), add = T)
```



Problem 9) Repeat the previous step for male cats. How do the distributions compare?

```
hist(filter(cats,Sex==("M"))$Hwt,main = "Male Cats Density Histogram",xlab= 'Hwt',prob = T,ylim = c(0, 

curve(dgamma(x, shape = Male_Cats[1], scale = Male_Cats[2]), add = T)
```



Male Cats Density Histogram
Density
0.00
0.05
0.10
0.15
0.20
10
15
20
Hwt