Lab 3

AnyaConti28661255

February 13, 2017

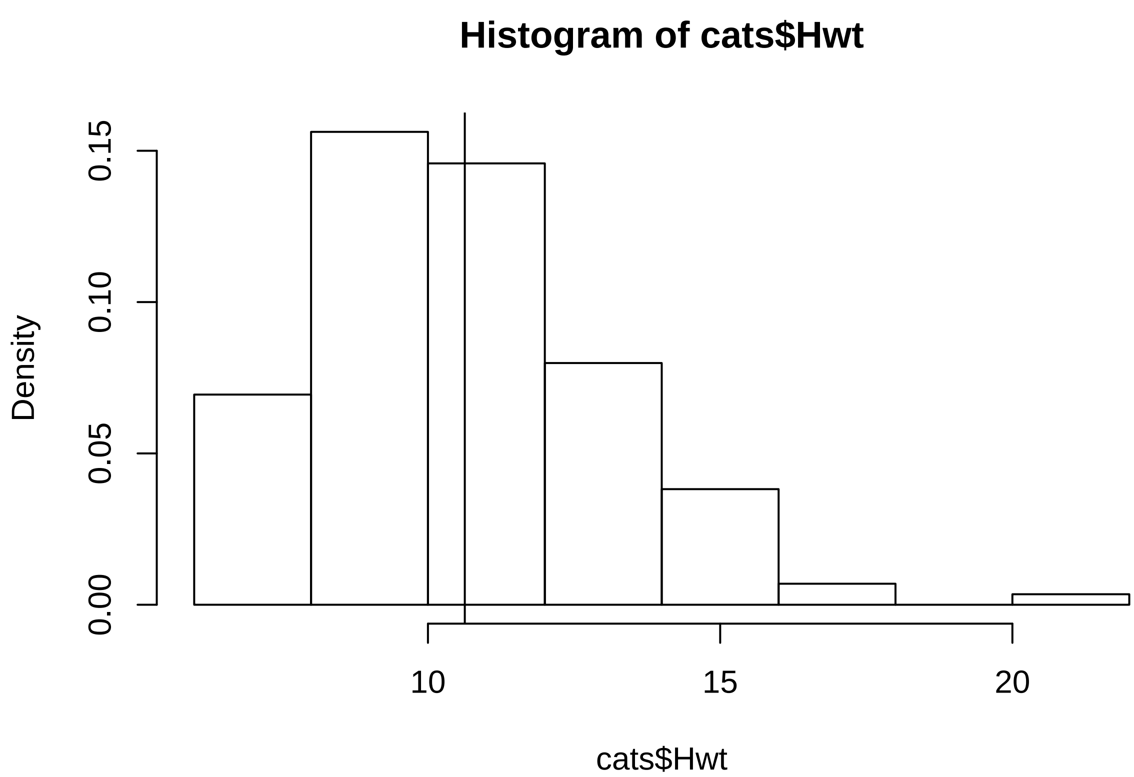
1. First, the R package MASS is loaded using library(), and then the data called cats is loaded using data(). This data is then summarized using summary(). The first section entitled sex shows how many females and how many males there are, then the second section called Bwt shows summary statistics for the body weight, and the third section called Hwt shows the summary statistics for heart weight (hwt). There is no cross-analysis between males vs females for these categories however.

```
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
```

2. A histogram of the heart weights is then plotted using the probability=TRUE so that the total area under them is equal to one, so it can be a probability function. Then abline() was used to add a line. The v means vertical, and it is set to be at the mean of the heart weights of cats.

```
hist(x=cats$Hwt, probability = TRUE)
  abline(v=mean(cats$Hwt))
```



3. Writes a function called gammafunction which takes two inputs (mean and variance) and calculates a and s from them, then creates a list called output that has both new values inside it, and returns that output.

```
gammafunction <- function(gamma.mean, gamma.var){
  a = (gamma.mean ^ 2) / (gamma.var)
  s = (gamma.var) / (gamma.mean)
  output <- c("a"=a,"s"=s)
  return(output)
}
```

4. Calculates the mean of the heart weights of cats which is 10.631. Then calculates the standard deviation of the heart weights of cats which is 2.435. Then calculates the variance of the heart weights of cats which is 5.927. Then sses the gamma function from question 3 with the cat heart weight mean and variance as the inputs to calculate a and s for the data set. The calculated value for a is 19.065 and the calculated value for s is 0.558.

```
mean(cats$Hwt)
```

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

```
sd(cats$Hwt)
```

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

```
var(cats$Hwt)
```

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

```
gammafunction(mean(cats$Hwt),var(cats$Hwt))
```

```
##          a          s
## 19.0653121  0.5575862
```

5. Creates a function called cat.stats with 1 input called data. The function then does the following. It calculates the mean of the data vector input, and assigns it to a new variable called cats.mean. It calculates the variance of the data vector input and assigns it to a new variable called cats.var. It calculates an estimated a value for the data vector from the mean and the variance, and assigns it to a new variable calledd cats.a. It calculates an estimated s value for the data vector from the mean and the variance, and assigns it to a new variable calledd cats.s. It creates a list called output, in which it stores all of these values, and then returns the output. An example of it running with the cats heart weights as the input data vector is shown below

```
cat.stats <- function(data){
  cats.mean = mean(data)
  cats.var = var(data)
  cats.a = (cats.mean ^ 2) / (cats.var)
  cats.s = (cats.var) / (cats.mean)
  output <- c("mean"=cats.mean, "var"=cats.var, "a"=cats.a, "s"=cats.s)
  return(output)
}

cat.stats(cats$Hwt)
```

```
##       mean        var          a          s
## 10.6305556  5.9274514 19.0653121  0.5575862
```

6. First the cats.stats function is run with the cat heart weight data, but only for the entries where the sex is female. The results of this were saved as FHwt. This is then repeated, but only for the entries where the sex is male, and the results were saved as MHwt.

```
FHwt <- cat.stats(cats$Hwt[cats$Sex=="F"])
FHwt
```

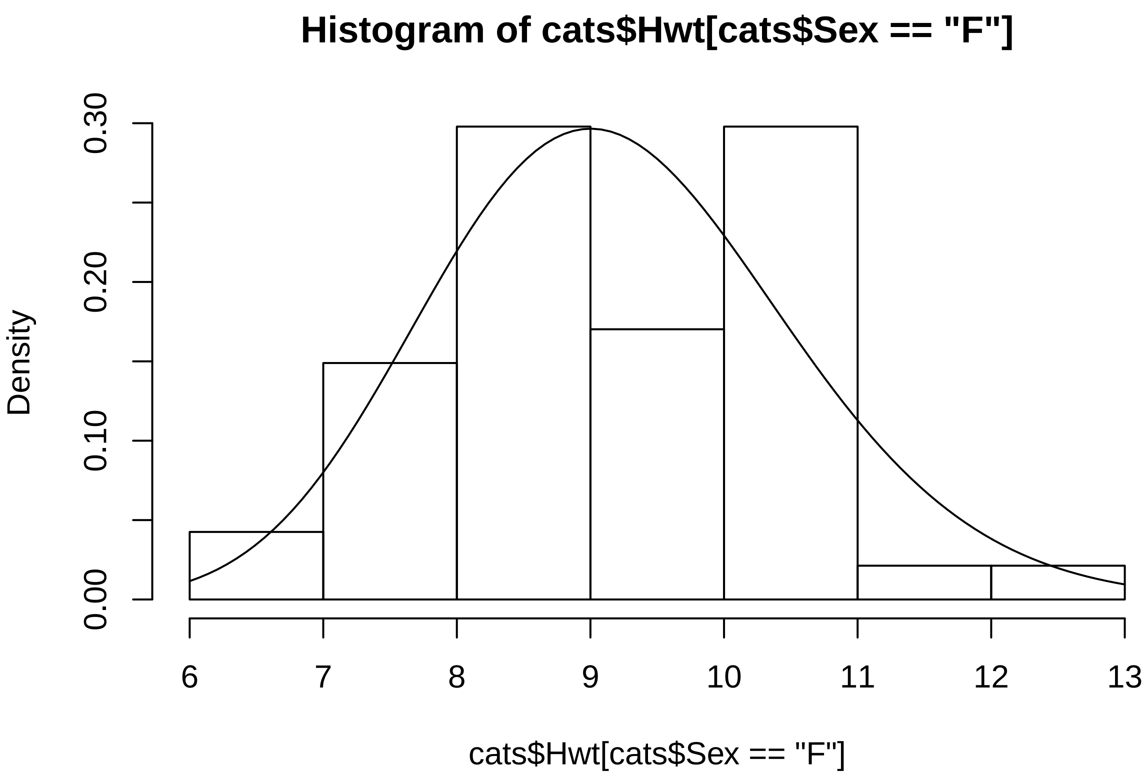
```
##        mean        var          a          s
##  9.2021277  1.8432562 45.9399792  0.2003076
```

```
MHwt <- cat.stats(cats$Hwt[cats$Sex=="M"])
MHwt
```

```
##        mean        var          a          s
## 11.3226804  6.4632302 19.8357612  0.5708216
```

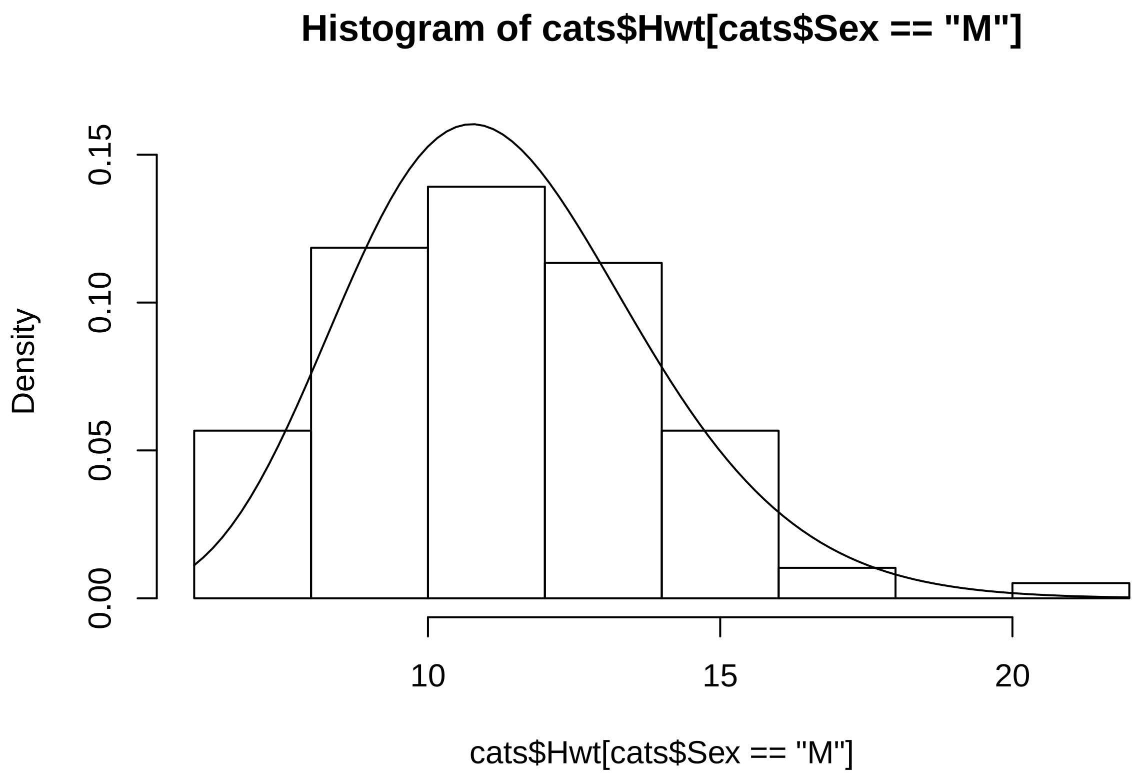
7. A histogram was created using the heart weight data, but only where the Sex is equal to "F" for female. Then the probability option was used, so that it would be a probability distribution funtion. Then a curve was added to this, where the expression was a gamma funtion with shape equal to the a value from the Female heart weight calculations from the previous problem (the 3rd entry of FHwt) and scale equal to the s value from the Female heart weight calculations (the 4th entry of FHwt).

```
hist(x=cats$Hwt[cats$Sex=="F"], probability = TRUE)
 curve(dgamma(x, shape = FHwt[[3]], scale = FHwt[[4]]), add=TRUE)
```



8. A histogram was created using the heart weight data, but only where the Sex is equal to "M" for male. Then the probability option was used, so that it would be a probability distribution funtion. (The y axis was adjust to have a slightly larger range so that the curve would fit on the graph.) Then a curve was added to this, where the expression was a gamma funtion with shape equal to the a value from the Male heart weight calculations from the previous problem (the 3rd entry of MHwt) and scale equal to the s value from the Female heart weight calculations (the 4th entry of MHwt).

```
hist(x=cats$Hwt[cats$Sex=="M"], probability = TRUE, ylim = c(0,0.16))
  curve(dgamma(x, shape = MHwt[[3]], scale = MHwt[[4]]), add=TRUE)
```


Histogram of cats$Hwt[cats$Sex == "M"]
Density
0.00
0.05
0.10
0.15
10
15
20
cats$Hwt[cats$Sex == "M"]