HW2\_pnegandh

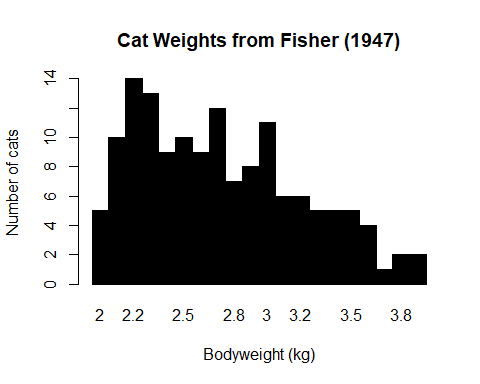
library("MASS")  
data(cats)  
str(cats)

## 'data.frame': 144 obs. of 3 variables:  
## $ Sex: Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 1 1 ...  
## $ Bwt: num 2 2 2 2.1 2.1 2.1 2.1 2.1 2.1 2.1 ...  
## $ Hwt: num 7 7.4 9.5 7.2 7.3 7.6 8.1 8.2 8.3 8.5 ...

head(cats)

## Sex Bwt Hwt  
## 1 F 2.0 7.0  
## 2 F 2.0 7.4  
## 3 F 2.0 9.5  
## 4 F 2.1 7.2  
## 5 F 2.1 7.3  
## 6 F 2.1 7.6

counts = table(cats$Bwt)  
barplot(counts, space = 0, col = "black", xlab = "Bodyweight (kg)", ylab = "Number of cats", main = "Cat Weights from Fisher (1947)")



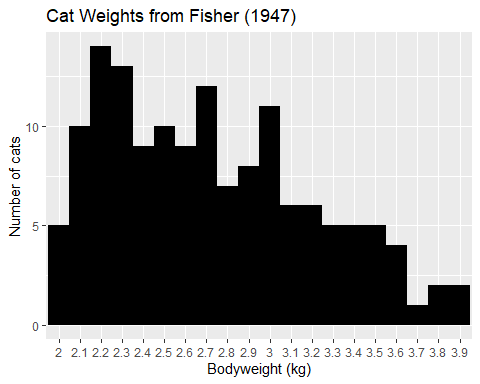
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.5.3

p<-ggplot(data=cats, aes(x=factor(Bwt))) + geom\_histogram(stat="count", width=1, fill="black") + xlab("Bodyweight (kg)") + ylab("Number of cats") + ggtitle("Cat Weights from Fisher (1947)")

## Warning: Ignoring unknown parameters: binwidth, bins, pad

p



Discussion:

This distribution at first glance looks like a trimodal distribution as it has three peaks. However after staring at it for some more time, it started looking like a Beta distribution with some outliers or some error terms added to it.

However, if we want we can still fit a mixtire of gaussians model to it with the number of gaussians equalling 3 and I am sure it will give good results to predict the proportion of cats equaling a certain weight in a population.