PH525.1x: Week1

Shu Guo

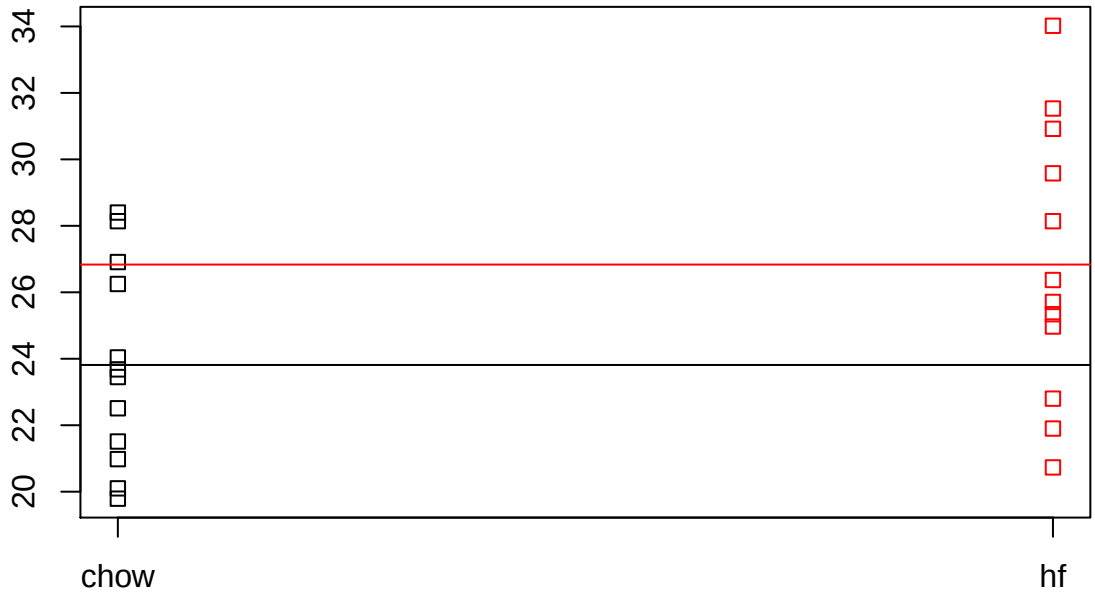
Sunday, February 01, 2015

Introduction to Random Varialbe I

```
## Load the data
dat = read.csv("femaleMiceWeights.csv")
## The observed difference between high fat diet and control was
mean(dat[13:24,2]) - mean(dat[1:12,2])
```

```
## [1] 3.020833
```

```
## A strip chart of the weights of these two groups
s = split(dat[,2], dat[,1])
stripchart(s, vertical=TRUE, col=1:2)
## Add the means to the plot
abline(h=sapply(s, mean), col=1:2)
```



```
## Question 1.1 How many of the high fat mice weigh less than the mean of the control mice (chow)?
sum(s$hf < mean(s$chow))
```

```
## [1] 3
```

```
## Question 1.2 How many of the control mice weigh more than the mean of the high fat mice?
sum(s$chow > mean(s$hf))
```

```
## [1] 3
```

```
## Question 1.3  What is the proportion of high fat diet mice over 30?
sum(s$hf > 30)/length(s$hf)
```

```
## [1] 0.25
```

Introduction to Random Variables II

```
## Course example
dat[1:12, 2]
```

```
##  [1] 21.51 28.14 24.04 23.45 23.68 19.79 28.40 20.98 22.51 20.10 26.91
## [12] 26.25
```

```
mean(dat[13:24, 2] - mean(dat[1:12, 2]))
```

```
## [1] 3.020833
```

```
population <- read.csv("femaleControlsPopulation.csv")

n <- 10000
null <- vector("numeric", n)
for (i in 1:n){
  control <- sample(population[, 1], 12)
  treatment <- sample(population[, 1], 12)
  null[i] <- mean(treatment) -mean(control)
}

diff <- mean(dat[13:24, 2]) - mean(dat[1:12, 2])
mean(null > diff)
```

```
## [1] 0.0107
```