

# lab17

Meha Thakur A16020450

## Table of contents

Population analysis. Data is as follows - c1: sample, c2: genotype, c3: ORMDL3 expression (I'm assuming) . . . . . 1

### Population analysis. Data is as follows - c1: sample, c2: genotype, c3: ORMDL3 expression (I'm assuming)

```
data<-read.delim("https://bioboot.github.io/bggm213_W19/class-material/rs8067378_ENSG0000017")
```

summary stats determine sample size for each genotype and their corresponding median expression levels for each of these genotypes.

```
geno1<-data[data$geno=="A/A",]  
geno2<-data[data$geno=="A/G",]  
geno3<-data[data$geno=="G/G",]  
  
nrow(geno1)
```

```
[1] 108
```

```
median(geno1$exp)
```

```
[1] 31.24847
```

```
nrow(geno2)
```

```
[1] 233
```

```
median(geno2$exp)
```

```
[1] 25.06486
```

```
nrow(geno3)
```

```
[1] 121
```

```
median(geno3$exp)
```

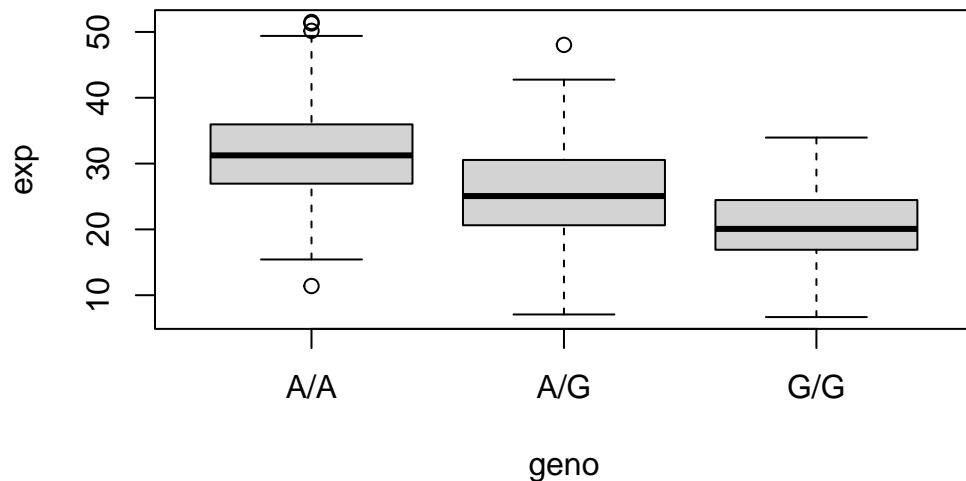
```
[1] 20.07363
```

Answers:

A/A: sample size = 108, median expression = 31.24847 A/G: sample size = 233, median expression = 25.06486 G/G: sample size = 121, median expression = 20.07363

boxplot

```
boxplot(exp~geno,data)
```



calculating significance of difference between A/A and G/G

```
t<-t.test(geno1$exp,geno3$exp)
t$p.value
```

```
[1] 9.48963e-26
```