

Class 11 Homework (Q13 & 14)

Lilith Sadil

Class 11 Homework

Q13. Read this file into R and determine the sample size for each genotype and their corresponding median expression levels for each of these genotypes.

```
expr = read.table("class_11_data.txt")
head(expr)
```

	sample	geno	exp
1	HG00367	A/G	28.96038
2	NA20768	A/G	20.24449
3	HG00361	A/A	31.32628
4	HG00135	A/A	34.11169
5	NA18870	G/G	18.25141
6	NA11993	A/A	32.89721

```
nrow(expr)
```

```
[1] 462
```

There are 462 individuals in the dataset.

```
table(expr$geno)
```

A/A	A/G	G/G
108	233	121

There are 108 individuals with A/A genotype, 233 individuals with an A/G genotype, and 121 individuals with a G/G genotype.

We can now make a box plot to visualize the relative expressions of each genotype. Using summary statistics, we can see that the median expression levels for each genotype are as follows:

- A/A = 31.250
- A/G = 25.065
- G/G = 20.074

```
library(ggplot2)
x = ggplot(expr) + aes(x=geno, y=exp, fill=geno) + geom_boxplot(notch=TRUE)
summary_stats <- by(expr$exp, expr$geno, summary)
summary_stats
```

```
expr$geno: A/A
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
11.40  27.02   31.25   31.82  35.92   51.52
```

```
-----
expr$geno: A/G
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 7.075 20.626  25.065  25.397 30.552  48.034
```

```
-----
expr$geno: G/G
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 6.675 16.903  20.074  20.594 24.457  33.956
```

Q14: Generate a boxplot with a box per genotype, what could you infer from the relative expression value between A/A and G/G displayed in this plot? Does the SNP effect the expression of ORMDL3?

From the boxplot generated, the relative expression of ORMDL3 is highest with the A/A genotype, reduced with the A/G genotype, and the lowest among G/G genotypes. Since the notches assigned to each box exhibit no overlap, it is statistically sound to conclude that different genotypes exhibit different levels of ORMDL3 expression. Therefore, this SNP does most likely effect the expression of ORMDL3.

```
library(ggplot2)
ggplot(expr) + aes(x=geno, y=exp, fill=geno) + geom_boxplot(notch=TRUE)
```

