

# ExampleANOVA

2024-01-12

## Antibody Responses

Researchers compared antibody responses in normal and alloxan diabetic mice. Three groups of mice were studied: normal, alloxan diabetic, and alloxan diabetic treated with insulin. Several comparisons are of interest:

- Does the antibody response of alloxan diabetic mice differ from the antibody response of normal mice?
- Does treating alloxan diabetic mice with insulin affect their antibody response?
- Does the antibody response of alloxan diabetic mice treated with insulin differ from the antibody response of normal mice?

## Requirements of ANOVA

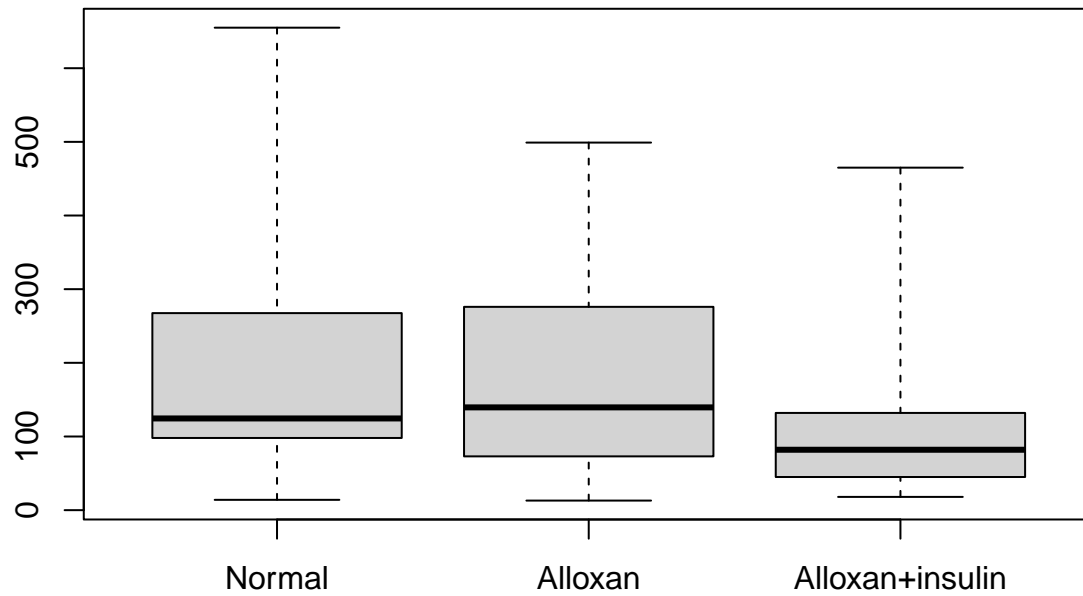
In order to use ANOVA, we need the ANOVA assumptions of normality and homoscedasticity to be met. To check, included below are the box plots and the QQ plots of all three sets of data, Normal, Alloxan, and Alloxan+insulin.

```
library(ggplot2)

normal=c(156, 282, 197, 297, 116, 127, 119, 29,
253, 122, 349, 110, 143, 64, 26, 86, 122, 455, 655, 14)
alloxan=c(391, 46, 469, 86, 174, 133, 13, 499,
168, 62, 127, 276, 176, 146, 108, 276, 50, 73)
insulin=c(82, 100, 98, 150, 243, 68, 228, 131,
73, 18, 20, 100, 72, 133, 465, 40, 46, 34, 44)

boxplot(normal, alloxan, insulin, range=0,
names=c("Normal", "Alloxan", "Alloxan+insulin"),
ylab="Nitrogen-bound bovine serum albumin produced")
```

Nitrogen-bound bovine serum albumin produced

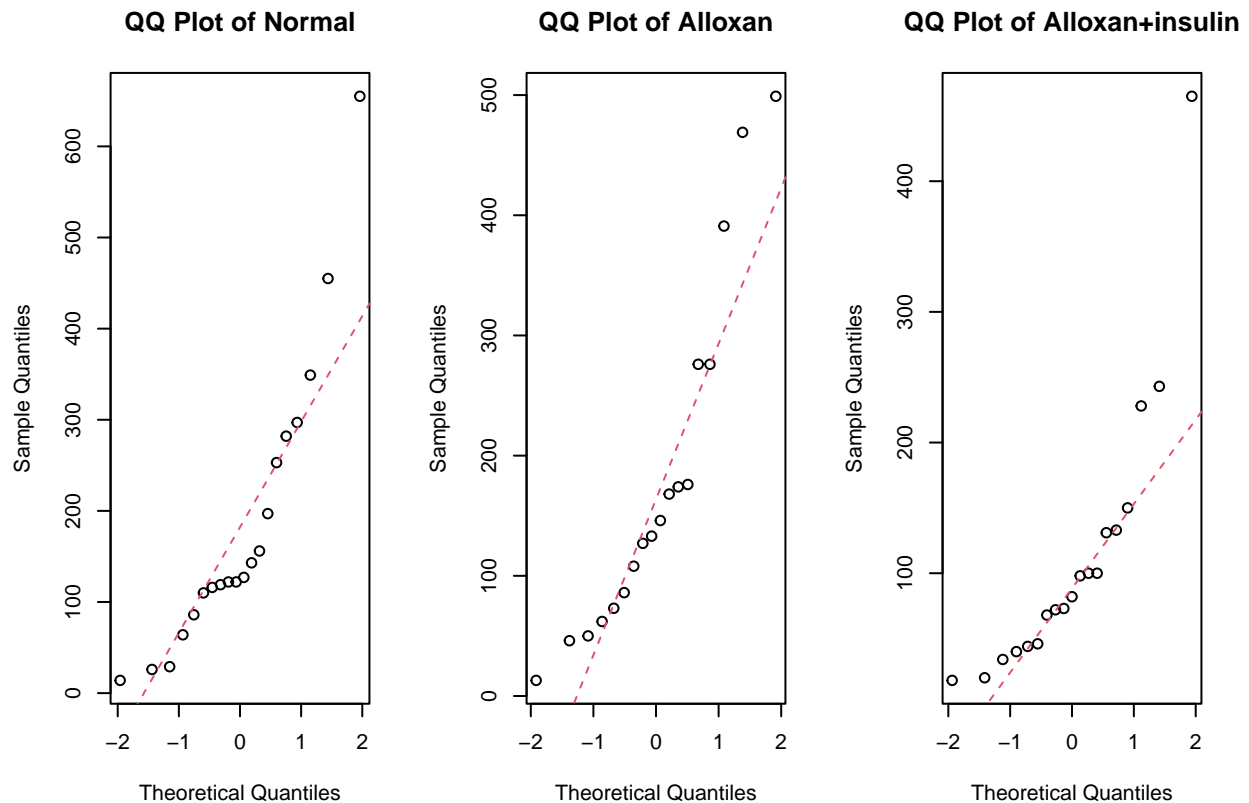


```
par(mfrow = c(1, 3))

qqnorm(normal, main = "QQ Plot of Normal", xlab = "Theoretical Quantiles", ylab = "Sample Quantiles")
qqline(normal, col = 2, lty = 2)

qqnorm(alloxan, main = "QQ Plot of Alloxan", xlab = "Theoretical Quantiles", ylab = "Sample Quantiles")
qqline(alloxan, col = 2, lty = 2)

qqnorm(insulin, main = "QQ Plot of Alloxan+insulin", xlab = "Theoretical Quantiles", ylab = "Sample Quantiles")
qqline(insulin, col = 2, lty = 2)
```



```
print(sd(normal))
```

```
## [1] 158.8349
```

```
print(sd(alloxan))
```

```
## [1] 144.8493
```

```
print(sd(insulin))
```

```
## [1] 105.7896
```

## Analyzing Plots

Looking at the boxplots and what the standard deviation of each sample is, we can see that the normal mice and the alloxan mice are closer in size and have a closer standard deviation, however, the mice treated with insulin has a much smaller standard deviation and it's width is much smaller than the other two. Even though the first two are closer in size, the standard deviation is quite different from each other and based on this, as well as the large difference from the last group of data, it seems less plausible that the assumptions for homoscedasticity are met. Then, looking at the QQ plots for all three, we can observe that all three plots are not straight lines. They all seem to curve upward. Therefore, it seems unlikely that the given data is normal. Hence, the assumptions of normality and homoscedasticity seem unlikely. Therefore, we next need to take the square root of the data to see if we can use ANOVA on the transformed data.

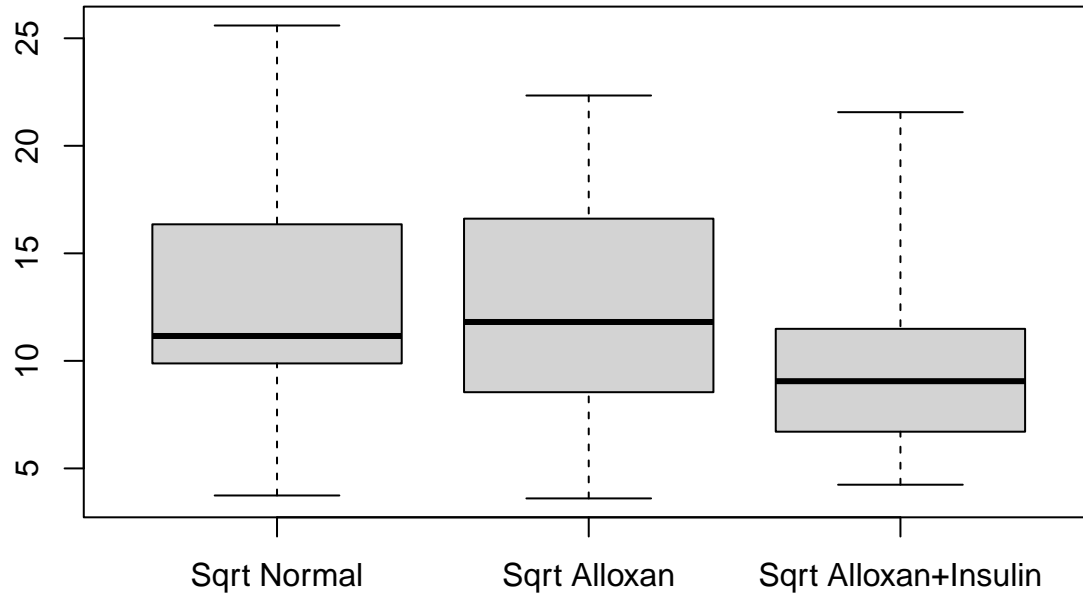
```
snormal=sqrt(normal)
```

```
salloxan=sqrt(alloxan)
```

```
sinsulin=sqrt(insulin)
```

```
boxplot(snormal, salloxan, sinsulin, range=0,
        names=c("Sqrt Normal", "Sqrt Alloxan", "Sqrt Alloxan+Insulin"),
        ylab="Square Root of Nitrogen-bound bovine serum albumin produced")
```

Square Root of Nitrogen-bound bovine serum albumin product

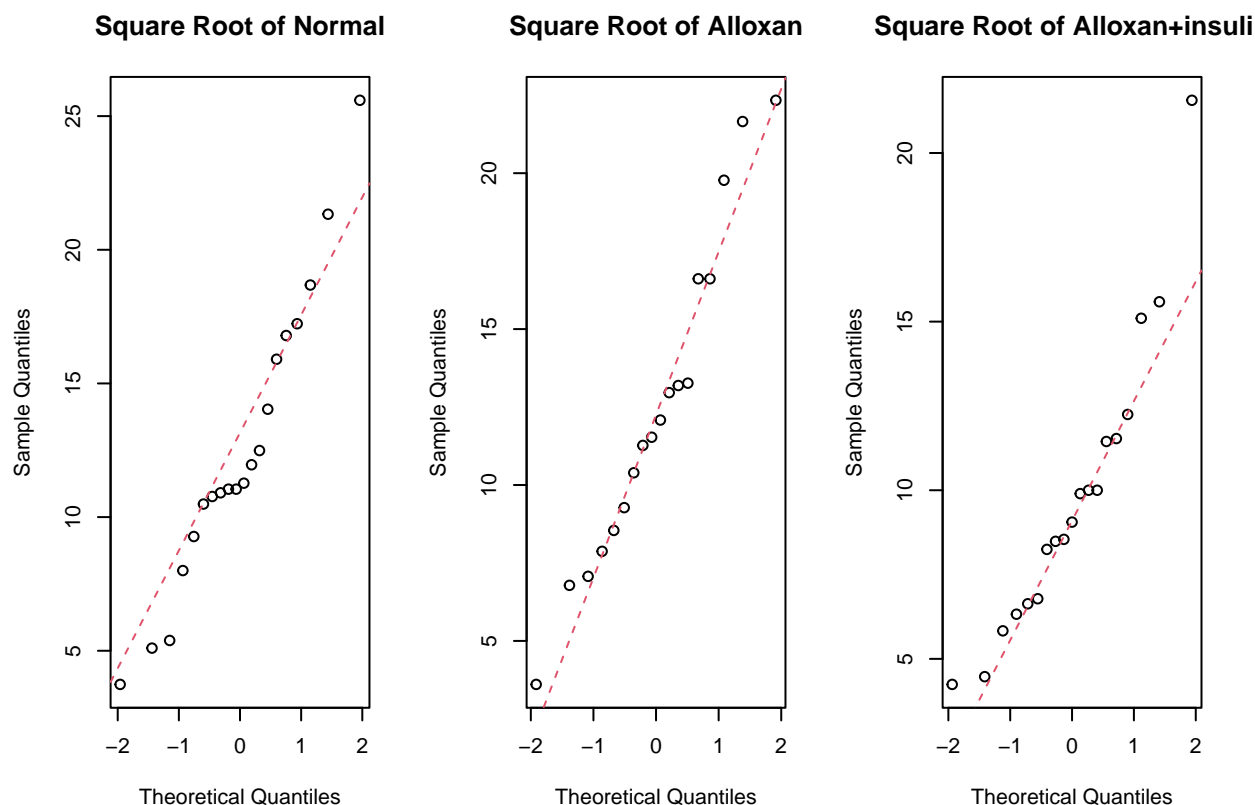


```
par(mfrow = c(1, 3))
```

```
qqnorm(snormal, main = "Square Root of Normal", xlab = "Theoretical Quantiles", ylab = "Sample Quantiles")
qqline(snormal, col = 2, lty = 2)
```

```
qqnorm(salloxan, main = "Square Root of Alloxan", xlab = "Theoretical Quantiles", ylab = "Sample Quantiles")
qqline(salloxan, col = 2, lty = 2)
```

```
qqnorm(sinsulin, main = "Square Root of Alloxan+insulin", xlab = "Theoretical Quantiles", ylab = "Sample Quantiles")
qqline(sinsulin, col = 2, lty = 2)
```



```
print(sd(snormal))
```

```
## [1] 5.480753
```

```
print(sd(salloxan))
```

```
## [1] 5.226939
```

```
print(sd(sinsulin))
```

```
## [1] 4.244474
```

We now can look at the boxplots and what the standard deviation of each sample is for the square root transformed data. We can see that the transformed normal mice data and the transformed alloxan mice data are even closer in size and have a much closer standard deviation. The transformed data for the mice treated with insulin still has a smaller standard deviation and it's width is still smaller than the other two. In this case however, they are all much closer in size than with the non-transformed data. The assumptions of homoscedasticity being met seem much more plausible with the transformed data. Then, looking at the Q-Q plots of the transformed data for all three, we can observe that all three plots are smoother than the non-transformed data. The plots for the alloxan and insulin treated mice data especially seem closer to straight line. The normal data however is still quite curved around 0. Therefore, it seems more likely that the square root of the data is normal. Hence, the assumptions of normality and homoscedasticity both seem more plausible in the square root transformed data.

## ANOVA

The null and alternative hypothesis tested by this method are:

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

$$H_1 : \text{at least one of the population means are not equal}$$

where  $\mu_1$  is the mean of Normal,  $\mu_2$  is the mean of Alloxan, and  $\mu_3$  is the mean of Alloxan+insulin.

Then,

```
sqrtall=c(snormal, salloxan, sinsulin)

n1=length(normal)
n2=length(alloxan)
n3=length(insulin)

N=length(sqrtall)

total_mean=mean(sqrtall)

snorm_mean=mean(snormal)
sall_mean=mean(salloxan)
sinsul_mean=mean(sinsulin)

SSB=n1*(snorm_mean-total_mean)^2+n2*(sall_mean-total_mean)^2+n3*(sinsul_mean-total_mean)^2

SSW=sum((snormal-snorm_mean)^2)+sum((salloxan-sall_mean)^2)+sum((sinsulin-sinsul_mean)^2)

x.all=c(snormal,salloxan,sinsulin)
group = factor(c(rep(1,20),rep(2,18),rep(3,19)))
print(summary(aov(lm(x.all ~ group))))

##           Df Sum Sq Mean Sq F value Pr(>F)
## group      2   94.7   47.37   1.882  0.162
## Residuals  54 1359.5   25.18
```

Since  $0.162 > \alpha = 0.05$  we fail to reject the null hypothesis.