# 6. Checking the Model Assumptions

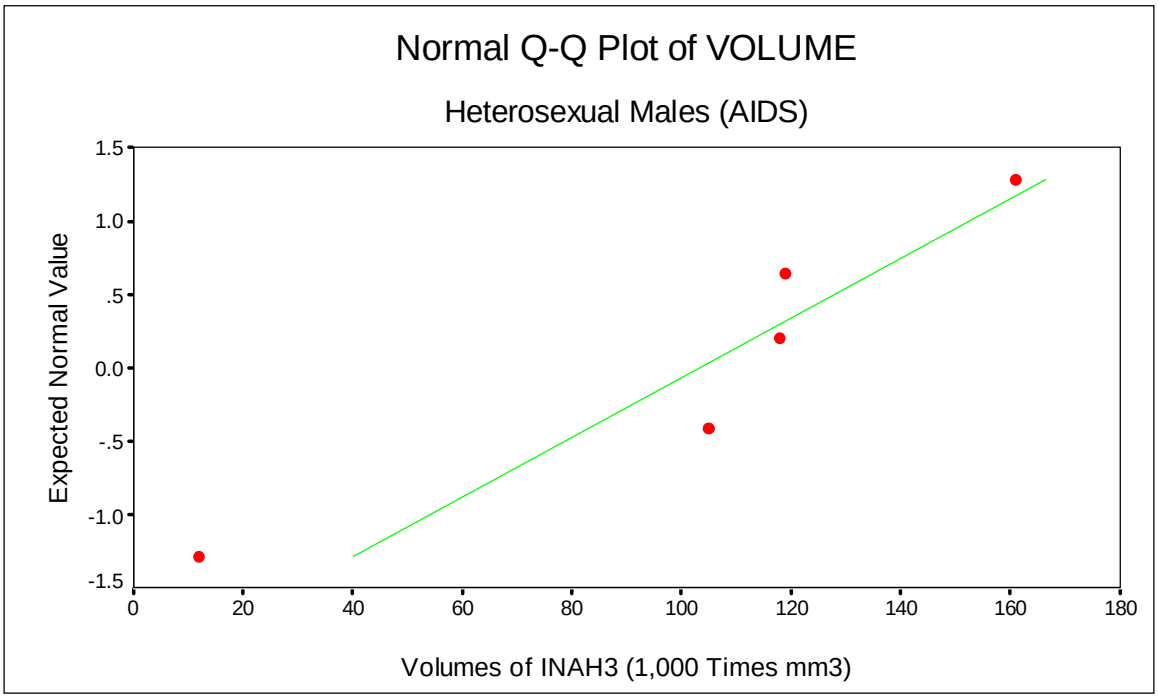
**6.1 Checking the Normality Assumption**
**6.2 Checking the Assumption of Equal Variances**
**6.3 Checking the Independence Assumption**

In order to determine whether there are significant differences among volumes of INAH3 for the five groups, we use one-way ANOVA model. However, the conclusions based on the model are valid only if the underlying assumptions are satisfied. Specifically we assume that:

1. The volumes have normal distributions for each of the five groups.
2. The standard deviations of volumes are all the same.
3. Observations within each group are independent of each other.
4. Observations in any one group are independent of observations in other groups.

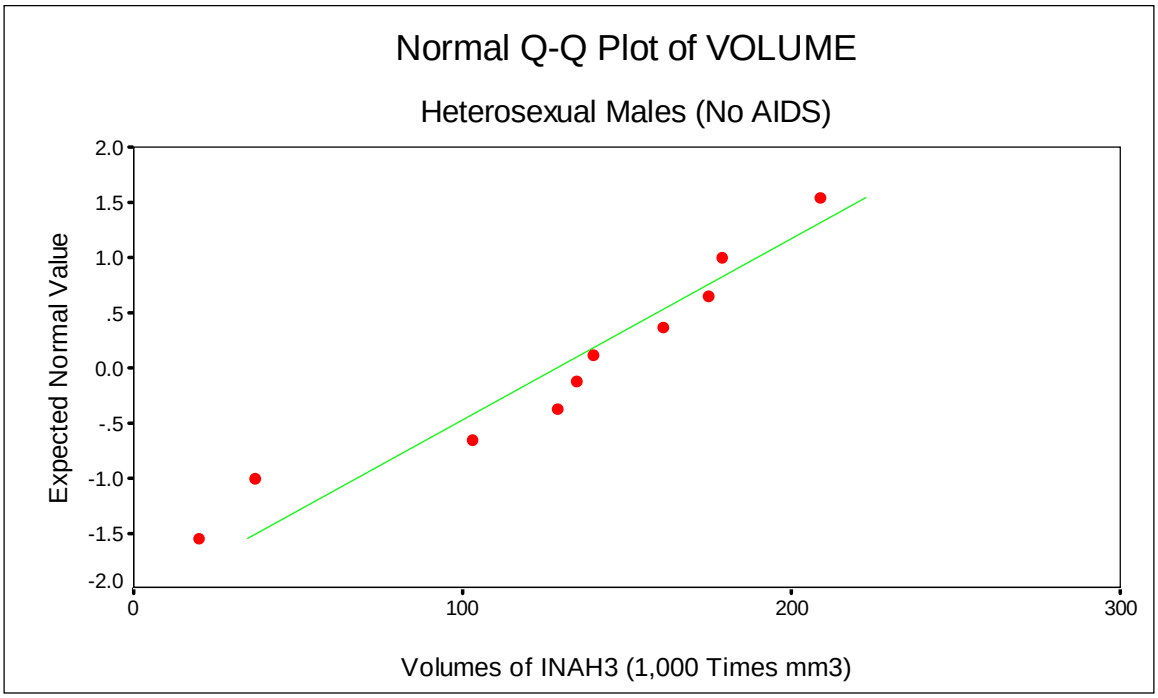
## 6.1 Checking the Normality Assumption

In order to verify whether the assumption of normality is plausible for our data, you can use either normal quantile plot of observations or the normal quantile plot of residuals, the differences between each observation and its group mean. As SPSS doesn't provide a normal quantile plot for the residuals, we will obtain a normal quantile plot for each of the five experimental groups.

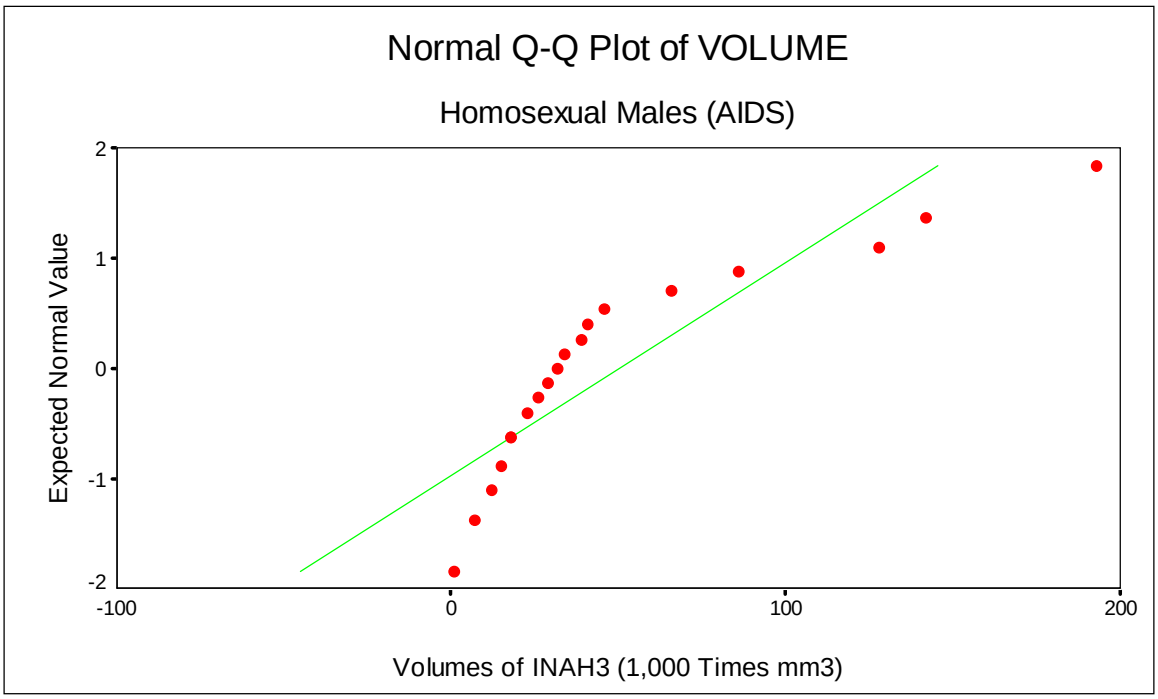



In the normal quantile plot of the volumes of INAH3 for the heterosexual males (AIDS) displayed above, all the points lie close to a straight line, indicating that the normality assumption is not violated. In other words, the data provided are consistent with the assumption of normality. However, notice that the sample size is relatively small which makes our conclusion about the general pattern weak.

Let us examine the plots for the remaining groups.

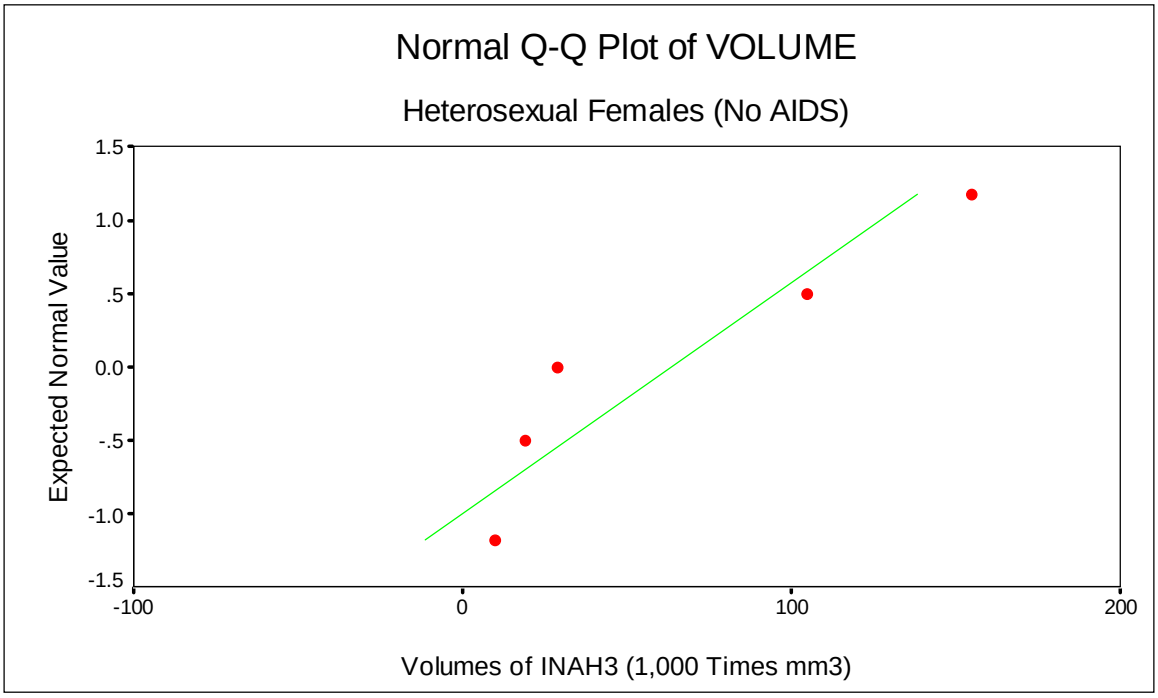



Also in this case there is no departure of the points in the plot from a straight line. Thus, it is very unlikely that the normality assumption is violated. However, the plot is obtained for just 10 observations, making it difficult to detect nonnormality.




The plot is based on 19 observations. It shows a serious departure of the points in the plot from a straight line. It is very likely that the normality assumption is violated. The point on the right lying far away from the line indicates a high outlier.

The fourth group (heterosexual females, AIDS) consists of just one observation. This is why we don't produce the normal quantile plot in this case.




There is some departure of the points in the above plot from a straight line. Thus, it is very likely that the normality assumption is violated. However, the plot is obtained for just five observations making the statement about nonnormality very weak.

## 6.2 Checking the Assumption of Equal Variances

Now we examine the assumption of equal variances. The assumption is crucial while making any inferences about the data.

The side-by-side boxplots obtained in Section 4 although obtained for relatively small numbers of observations indicate that the normality assumption might be seriously violated. Formal tests for the equality of standard deviation in several groups share lack of robustness against nonnormality. Because ANOVA procedures are not extremely sensitive to unequal standard deviations, it is not recommended to carry out a formal test of equality of standard deviations as a preliminary to the ANOVA. Instead, the following rule of thumb is used: If the ratio of the largest sample standard deviation to the smallest sample standard deviation is less than 2, the assumption of equal standard deviations is plausible.

In Section 5 we obtained the following summary statistics:

| Group | Count | Mean | Standard Deviation |
|---|---|---|---|
| 1 | 6 | 103.33 | 49.26 |
| 2 | 10 | 128.00 | 60.75 |
| 3 | 19 | 50.32 | 51.67 |
| 4 | 1 | 12.00 | 0 |
| 5 | 5 | 29.00 | 63.50 |

We will exclude the group 4 from the comparison, as this group consists of just one observation. A quick glance at the data ensures us that the assumption of equal variances is plausible in our case.

The Levene test for the homogeneity of variances for the data produces the p-value of 0.582, which is also consistent with the assumption of equal variances.

### 6.3 Checking the Independence Assumption

Of the all assumptions, independence is the most crucial. If this assumption is violated, the effect on inferences about the average volumes for the five groups can be severe. This assumption can be violated by the method in which the measurements were obtained. There is no sufficient information here to indicate that the assumption might be violated.