Stat 430 - HW 4
October 14, 2010

1.

- The distributions of samples from both clinics have long tails as indicated by the distal departure from linearity on the normal probability plots.

```
> # Load and sort the data
> hiv <- read.table(file("hiv_status_data.Rtxt"))
> X <- subset(hiv, Clinic.ID == 1)
> Y <- subset(hiv, Clinic.ID == 2)
> X.data <- sort(X$Y)
> Y.data <- sort(Y$Y)
>
> # Create norm prob plots manually
> X.n <- length(X.data)
> X.n <- length(Y.data)
> X.seq <- seq(from=1, to=X.n, by=1)/(X.n+1)
> Y.seq <- seq(from=1, to=Y.n, by=1)/(Y.n+1)
> X.q <- qnorm(X.seq)
> Y.q <- qnorm(Y.seq)
> plot(X.q, X.data)
> plot(Y.q, Y.data)
>
> # Create plots with built-in functions
> qqnorm(X.data)
> qqline(X.data, col="red")
> qqnorm(Y.data)
> qqline(Y.data, col="red")
```

- Since we do not have a theoretical distribution for X and Y, we cannot show the relationship between X and $\text{VAR}(X)$. However, we can partition each sample set at the median and show that the variance of the larger samples is smaller than the variance of the larger samples.

- If the order of the transformed samples is not the same as the order of the original samples, it may not be easy to interpret the results of the test.