

Assessing Normality of a Sample

- normal quantile plots can be used to assess whether the data could have come from a normal distribution
- these plots are also called QQ or normal scores plots
- the sorted values are plotted against the values we would expect to get if the sample came from a normal distribution
- a straight line in this plot indicates that the data are normally distributed
- outliers show up as values distant from the overall pattern
- curvature indicates departure from normality e.g. skewness
- the NSCORES command in MINITAB can be used to produce the values to be plotted against the data

For example

```
MTB > nscores c1 c2
```

- Manually you can create the normal probability scores using the following commands

First rank the data from C1 to C3.

Then create the quantile of each observation in column C4

and finally find the point of the normal distribution which gives you the corresponding quantile.

```
MTB > rank c1 c3
```

```
MTB > let c4=(c3-.375)/(10+.25)
```

```
MTB > invcdf c4 c5;
```

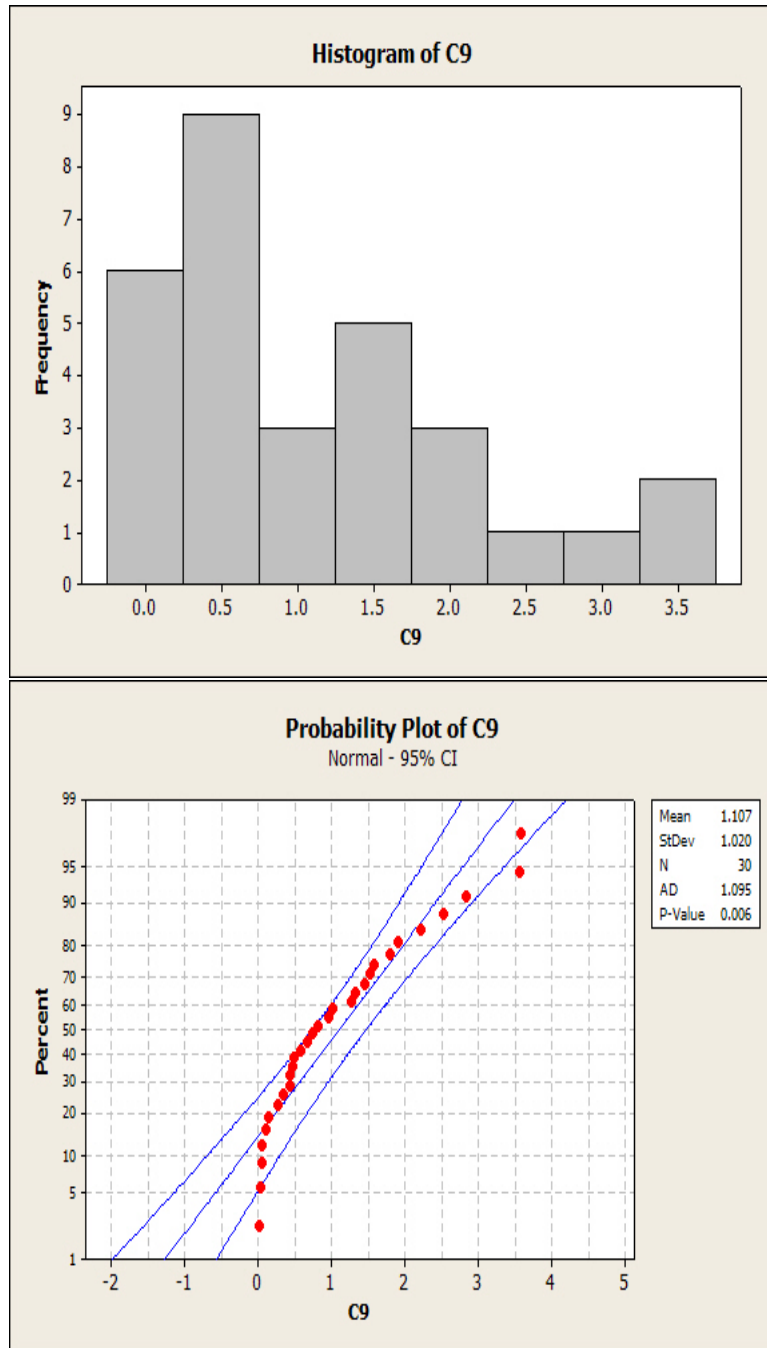
```
SUBC> normal 0 1.
```

Comment: 10 is the sample size,

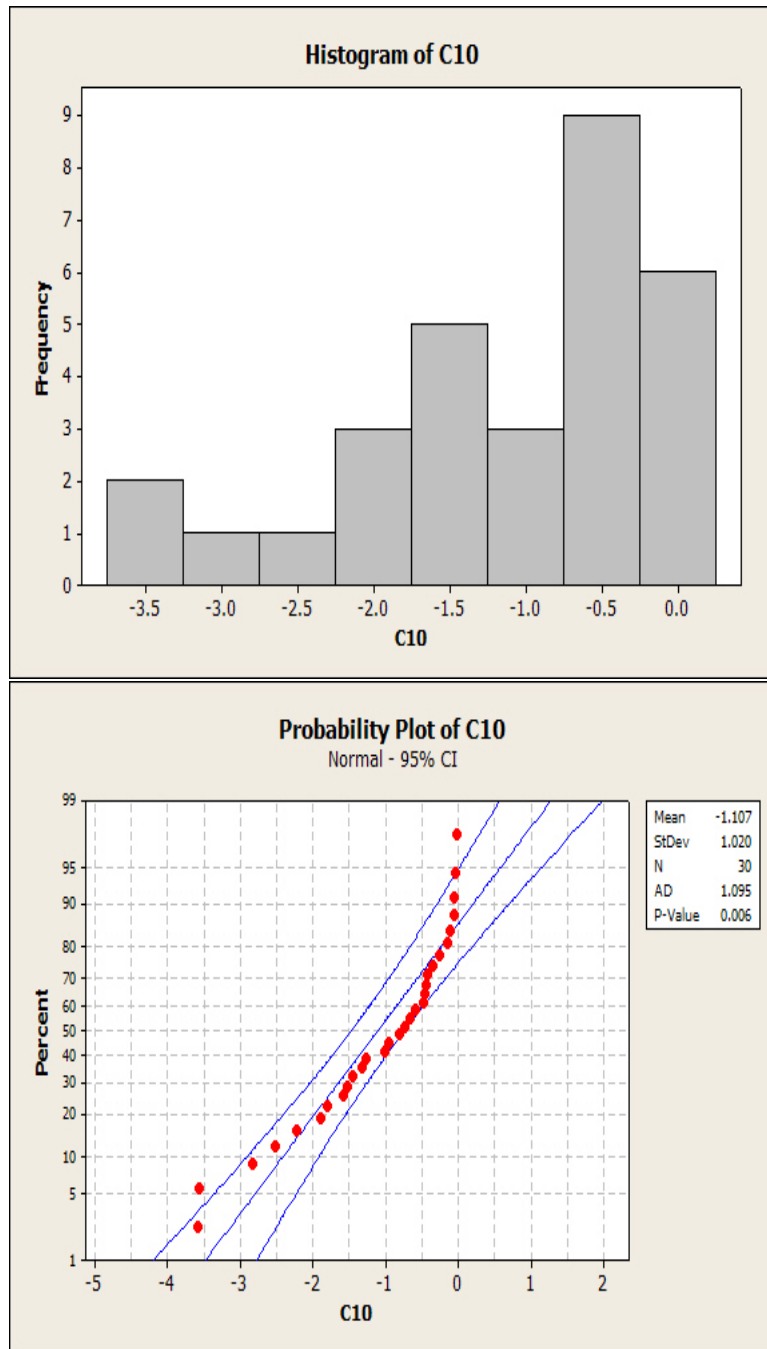
.375 and .25 are constants needed for the estimation of the quantiles.

- the curvature in the probability plot can reveal the shape of distribution

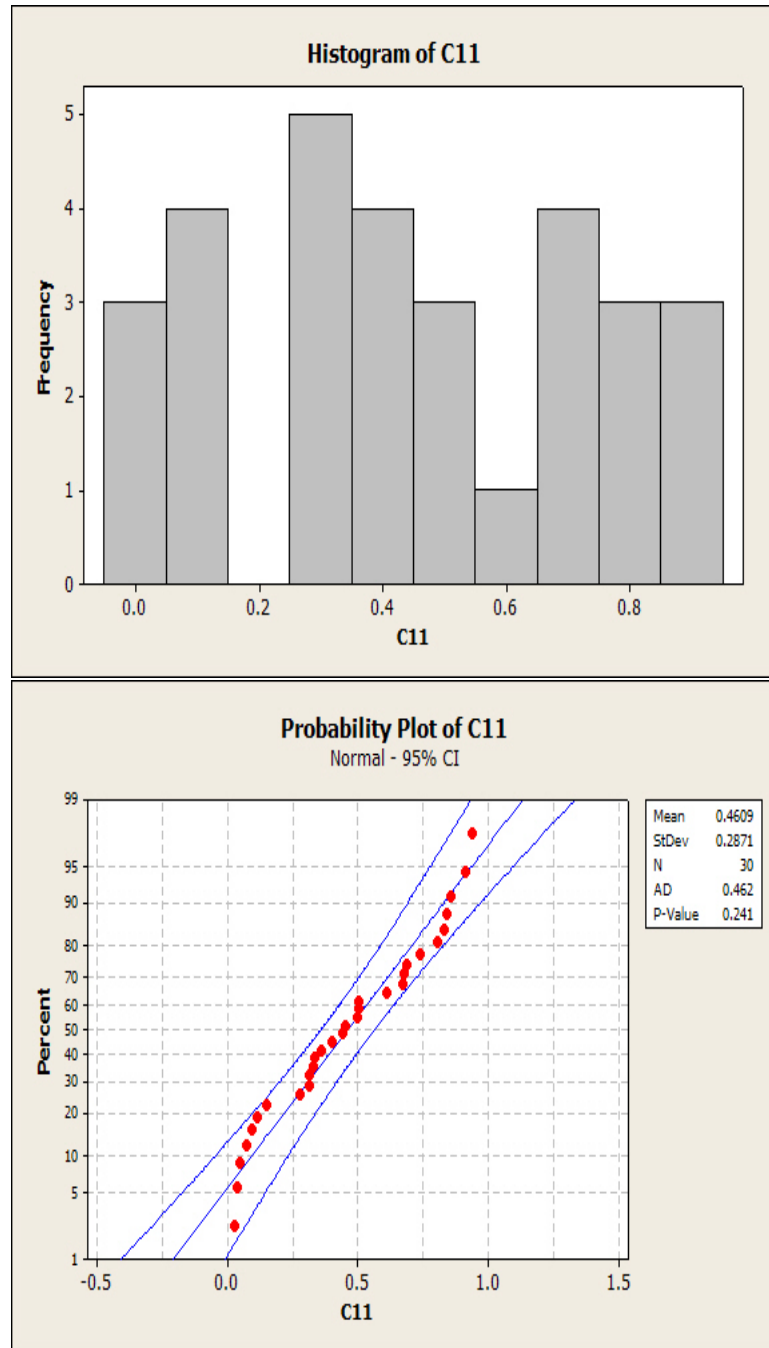
- if the distribution is skewed to the right, the probability plot curves down at both the left and the right



- if the distribution is skewed to the left, the probability plot curves up at each end



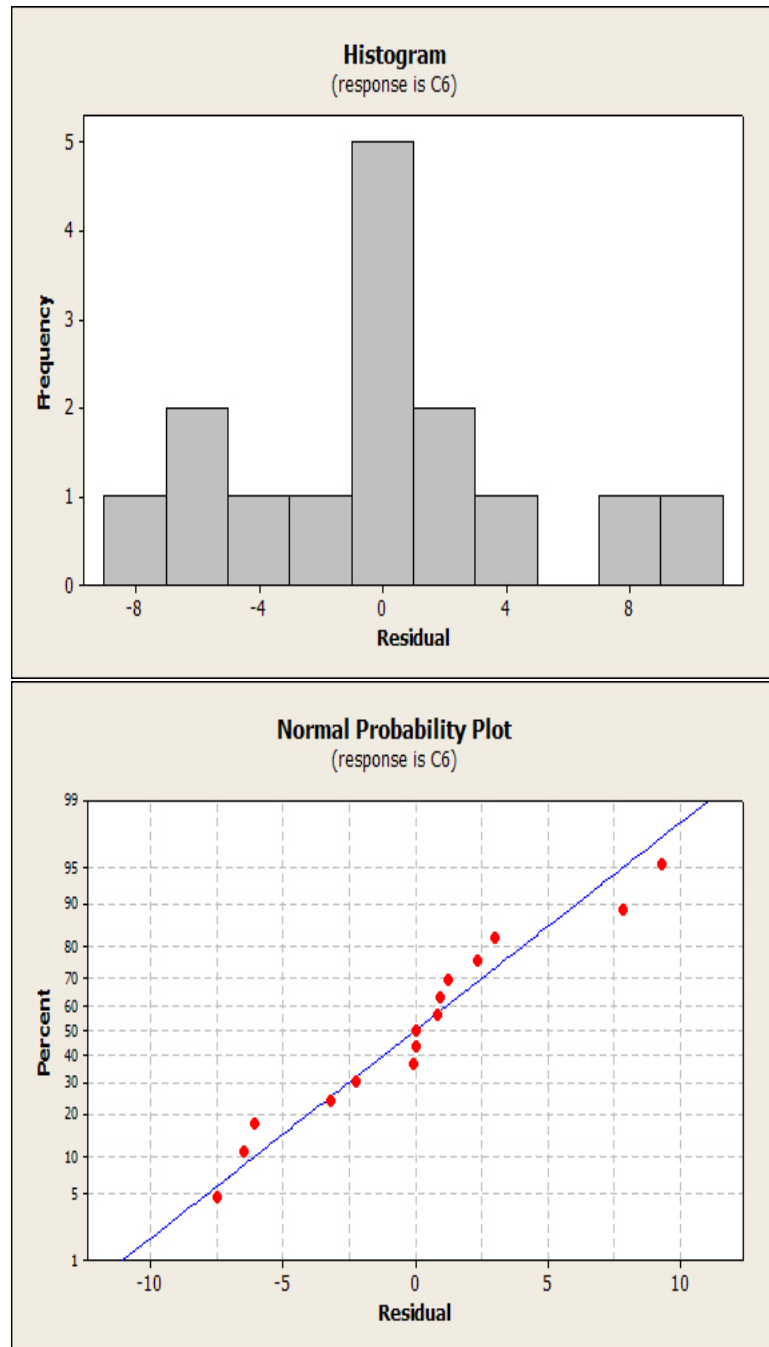
- if the distribution has a flatter peak than the normal, the normal scores plot curves down at the left and up at the right



- Example: The probability plot for residuals from the data on golf balls is shown below. Alternatively the residuals can be saved (in C3 for example) and the plot obtained using the commands

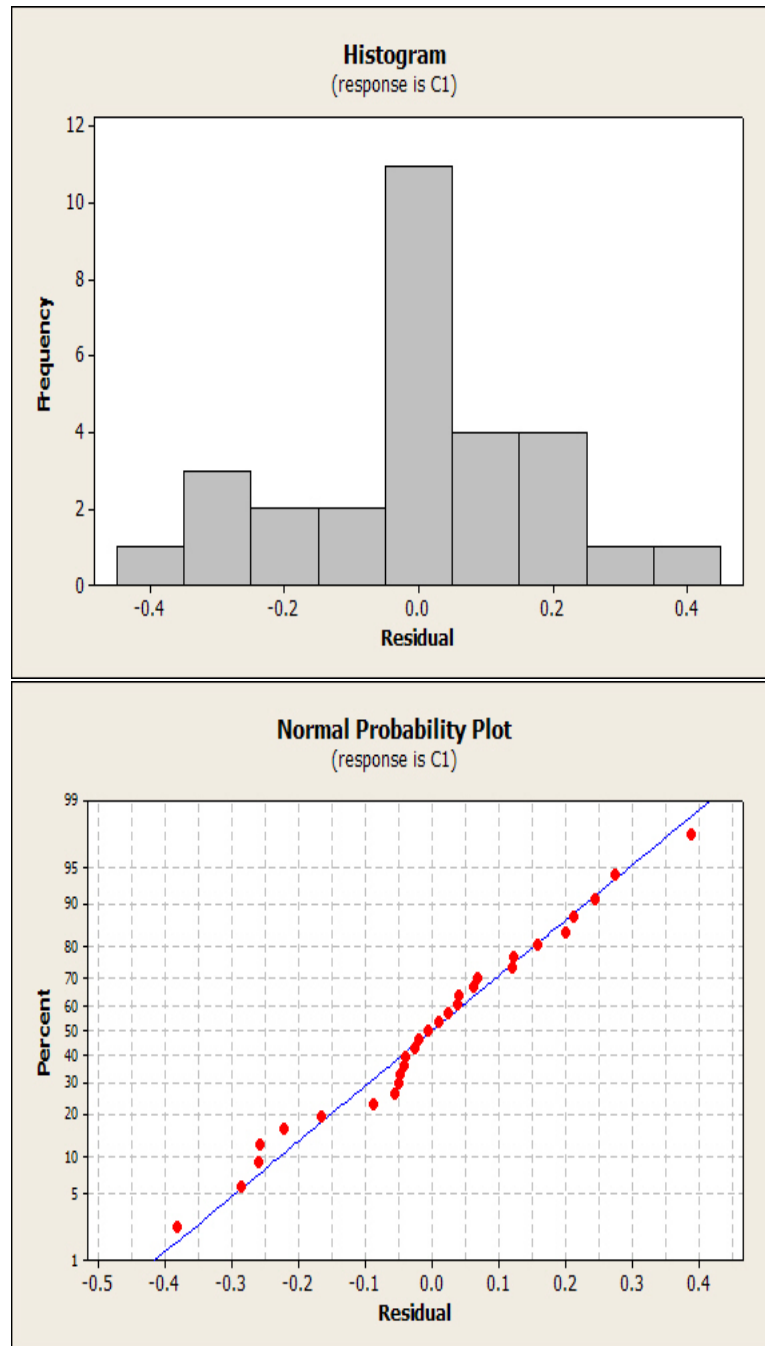
```
MTB > NSCORE C3 C4
```

```
MTB > PLOT C3*C4
```



- the points are fairly close to a straight line
- there is no evidence of a departure from normality

- a plot like this can be plotted directly from the pull-down menu for the One-Way command under the Graphs button.
- Example: For the liver weight data, the probability plot of the residuals is shown below.



- the normal scores are quite close to the straight line and again there is no evidence of a departure from normality