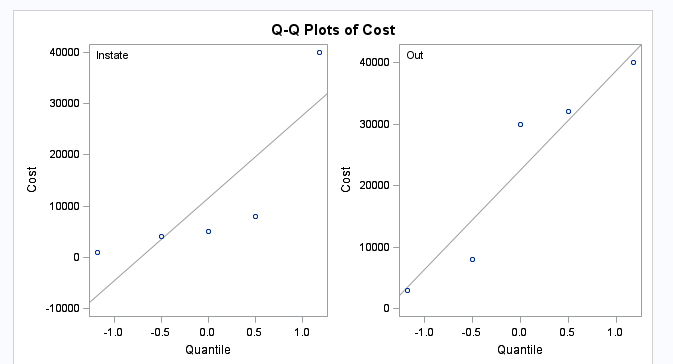
Instate

|  |  |  |  |
| --- | --- | --- | --- |
| Original Data | Percentage for percentiles given number of values | Z-score of original data | Z-score percentiles assuming normal distribution given the values in column 2. |
| 1,000 | .1 | -0.659 | -1.281 |
| 4,000 | .3 | -0.472 | -0.524 |
| 5,000 | .5 | -0.411 | 0 |
| 8,000 | .7 | -0.224 | 0.524 |
| 40,000 | .9 | 1.176 | 1.281 |

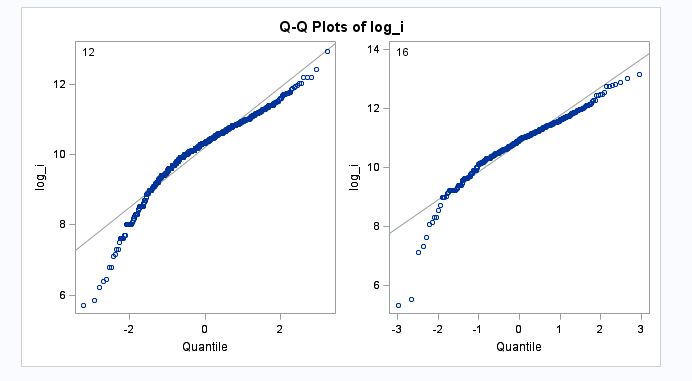
Out of State

|  |  |  |  |
| --- | --- | --- | --- |
| Original Data | Percentage for percentiles given number of values | Z-score of original data | Z-score percentiles assuming normal distribution given the values in column 2. |
| 3,000 | .1 | -1.21 | -1.281 |
| 8,000 | .3 | -0.904 | -0.524 |
| 30,000 | .5 | 0.548 | 0 |
| 32,000 | .7 | 0.582 | 0.524 |
| 40,000 | .9 | 1.077 | 1.281 |

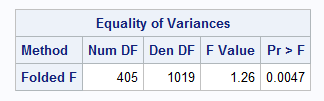
This data does not have a normal distribution. The instate cost sees a large jump from the 4th to 5th ranked data on the plot indicating an outlier. The out of state is perhaps closer to a normal distribution with a big jump between the 2nd and 3rd ranked data on the Q-Q plot.



2. We want to see if college (for extra years of education) makes a difference in a person’s income later in life (ie their 40s). After performing a log transformation the data becomes near enough normal.



However, according to the F-test the standard deviations are not equal. With a value of 0.0047 we must reject the null hypothesis that the variances are equal.



Proceeding with the test anyway we can see that there is in fact a difference in income. People who were in their 40s in 2005 and had gone to college made more on average than those of a similar age who hadn’t.

