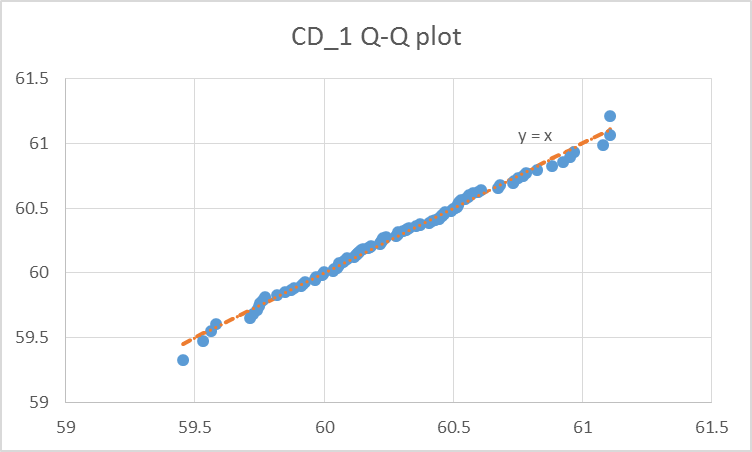
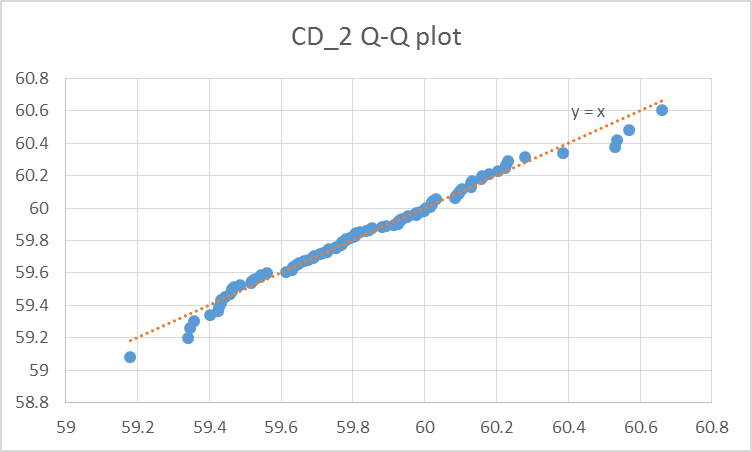
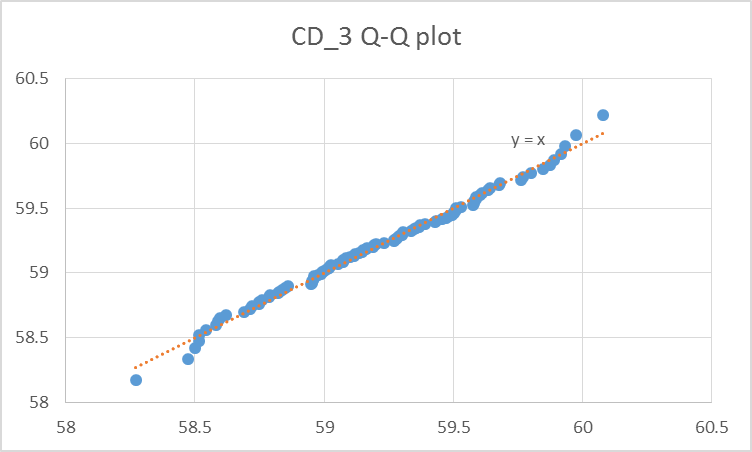
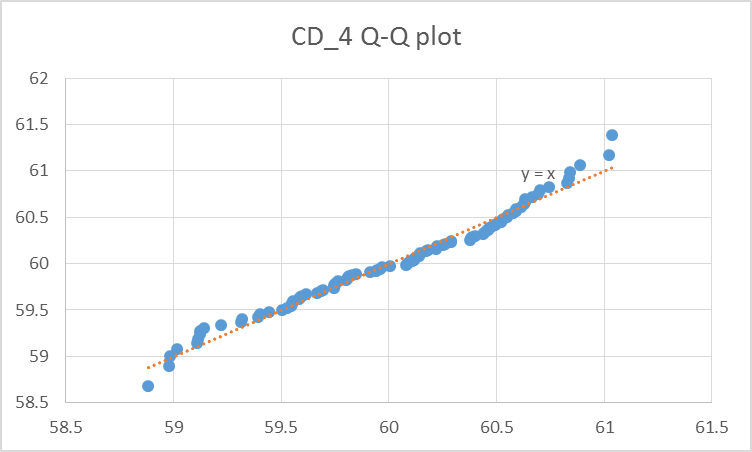
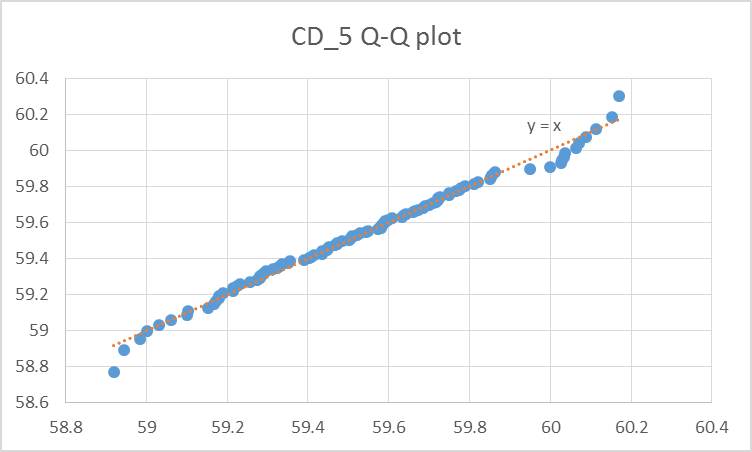
1. Q-Q plots by using normal distribution for CD\_1 ~ CD\_5 is shown as below:







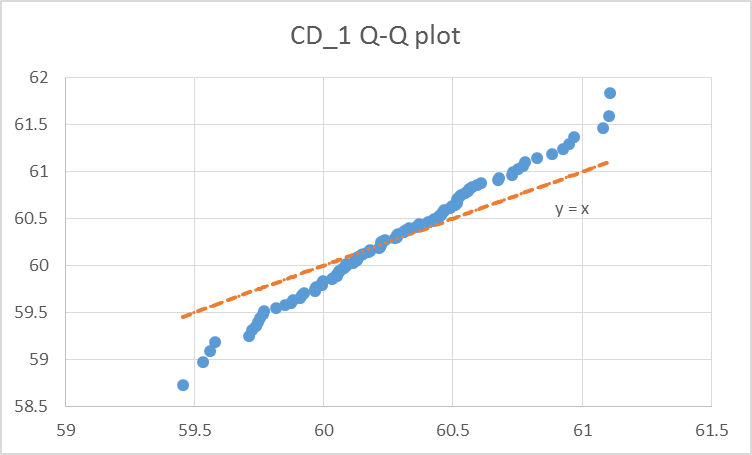


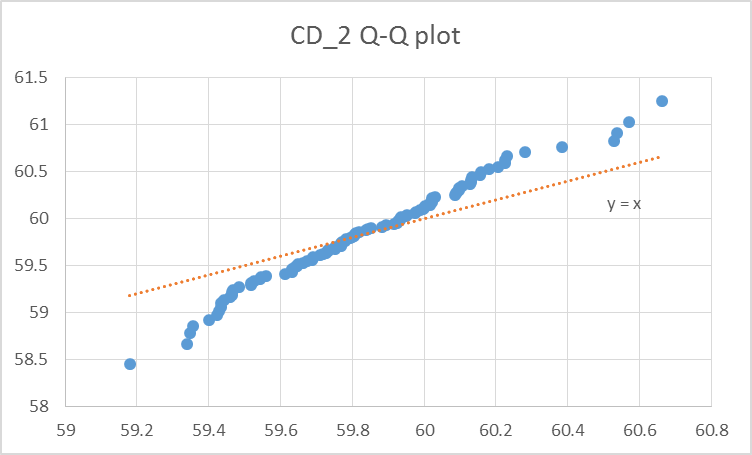


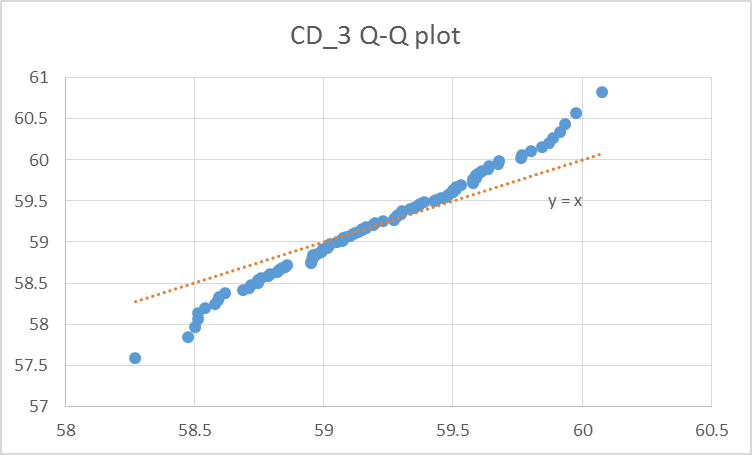
Q-Q plot by using gamma distribution, first is to find and by moment of estimates. The parameters for each sample are shown as below:

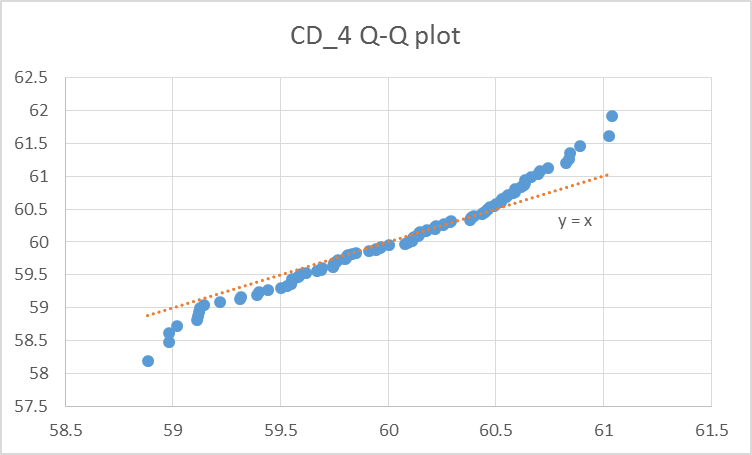
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| alpha | 9930.627 | 12105.1 | 8829.739 | 6846.763 | 11894.88 |
| beta | 0.006069 | 0.004943 | 0.006704 | 0.008768 | 0.005005 |

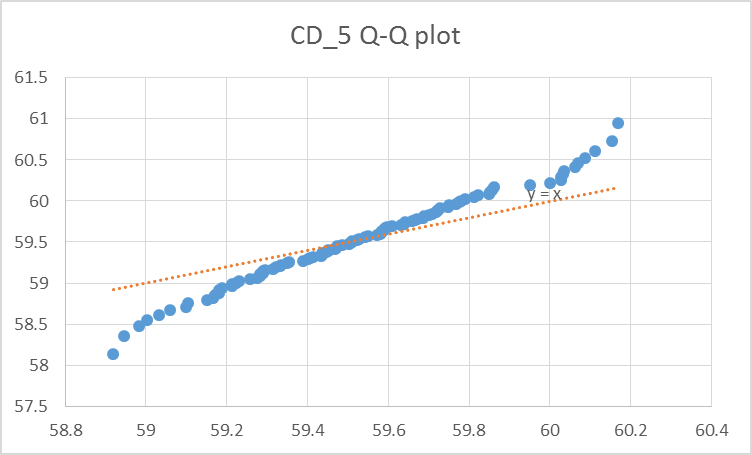
The Q-Q plot accordingly are displayed:











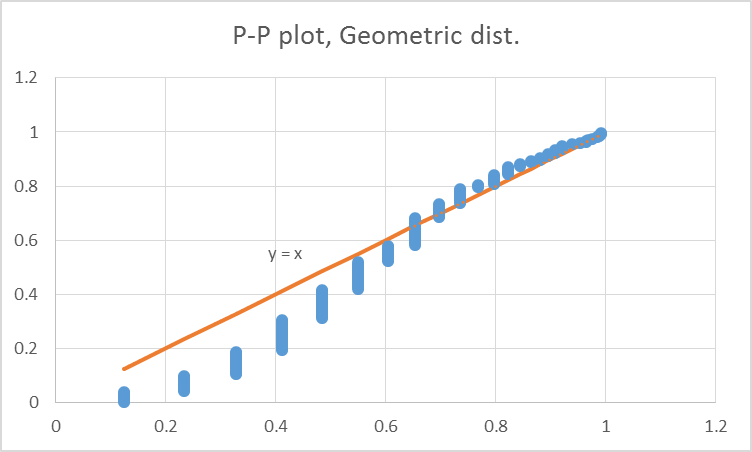
Compare the above, we can see that estimation under normal distribution, the errors to the line are smaller. By this observation, we can say that the sample is more fit by using normal distribution.

2. To do P-P plot under Geometric distribution, we first need to decide the parameter p in this distribution. By using MLE, we can see that Geometric distribution, , the likelihood function is:

For this, we can have .

So, , the cumulative distribution function for geometric distribution are , where k = 1, 2, 3, ….. ,n.

The P-P plot under geometric distribution is shown in below:



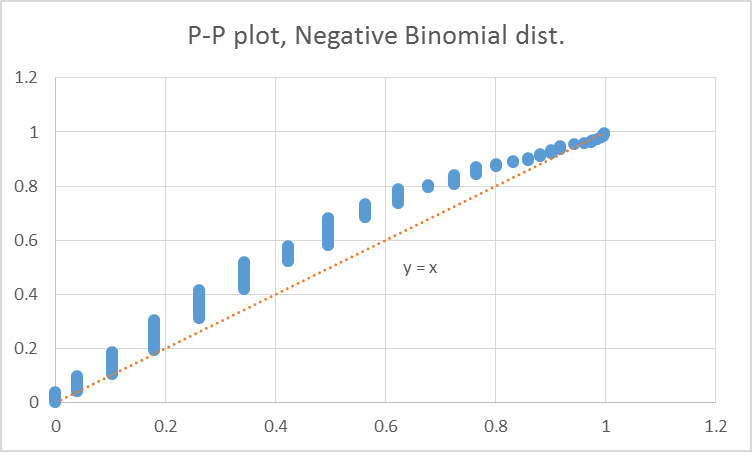
For negative binomial distribution,

so we have likelihood function

to satisfy

We have

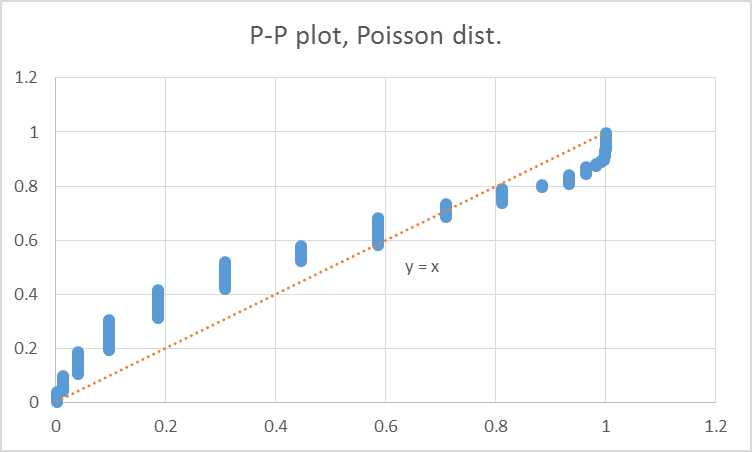
For this, we have and . From moment of estimate, we have , and . With these parameters, we can have P-P plot as shown in below. Note that p=0 when k<r.



For Poisson distribution, taking MLE for Poisson distribution and we have

in this problem, we have , and CDF =

According to the CDF, we can have the P-P plot in below:



1. From the comparison between three chart, we can say that Geometric distribution fits best because the error between sample and the 45 degree line are the smallest.
2. Z value of 75% CI = P(-1.15035<Z<1.150349). For CD\_1, the CI = ;  
   For CD\_3, the CI =
3. For t-statistic, internal for 75% CI = ;  
   For CD\_3, the CI =
4. The two means are not equal, because the CI will not overlap at all even when both the CI of mean have 99% confidence.
5. For CD\_5, with and with from t-distribution we can have critical value , in other word, the tolerance for to pass is when the observed estimate . from this data, the sample mean = 59.5364, which is larger than the critical value. So we can say that CD\_5 passed the hypothesis testing, and can say that is true with 75% confidence under single tail test.
6. For CD\_1, with and with from t-distribution we can have critical value , in other word, the tolerance for to pass is when the observed estimate . from this data, the sample mean = 60.27011, which is larger than the critical value. So we can say that CD\_1 passed the hypothesis testing, and can say that is true with 75% confidence under single tail test.
7. Yes, we can say that under the estimate that , both CD\_1 and CD\_5 passed the hypothesis test, so the hypothesis is true with a CI of 75%.