MATH 534 Homework 3

Due: Thursday, 9/7/2017

September 13, 2017

1. a. Use qqplots to find a plausible distribution for the the following 40 data.

   b. Can you do some test to check your conclusion ?

   c. Based on your study on the possible distribution, estimate the parameters in the distribution.

2. Skip

3. In Example 2 of §5.1, The MLE of $F(t)$ under the Weibull assumption is $\hat{F}(t) = 1 - \exp(-t/\hat{\tau})^{\hat{\gamma}})$, $t > 0$, where $(\hat{\gamma}, \hat{\tau})$ is asymptotically normal with mean y$e and Cov=y$vcov. Using the delta method to estimate the standard deviation of $\hat{F}(2) - \hat{F}(1)$. Check whethter $\hat{P}(X \in (1, 2])$ falls in the CI of $P(X \in 1, 2])$ under the EDF or under the Weibull assumption ?

   Answer: So the standard error will be about 0.0028, for edf, the 95% CI is (0.20,0.37), and for weibull distribution, the 95% CI is (0.17,0.19). As $P(X \in (1, 2]) \approx 0.175$, falls in the 95% CI of $P(X \in (1, 2])$ under weibull.

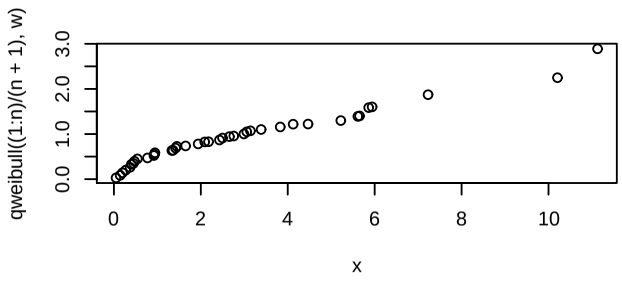
   Comments: Note that the CI due to Weibull distribution is $\hat{\theta} \pm 1.96 \times \hat{\sigma}_{\hat{\theta}}$, where $\theta = P(X \in (1, 2])$ here. Thus the last statement is nothing to do with the question.

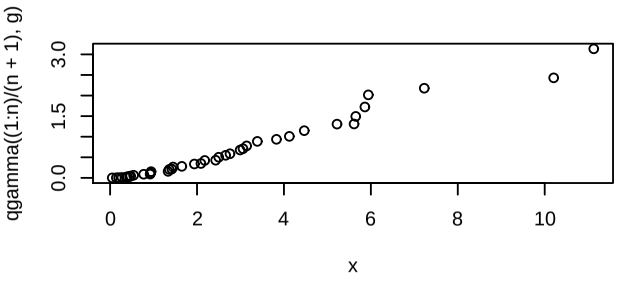
   By data analysis, it seems that $X \sim U(1, 5)$, thus $\hat{P}(X \in (1, 2]) = 0.25$.

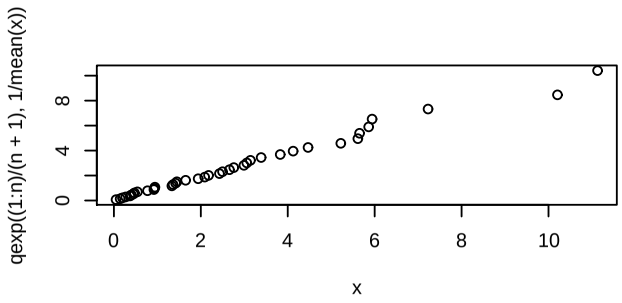
   CI of θ is $[0.20, 0.37]$ and length 0.17 for edf,

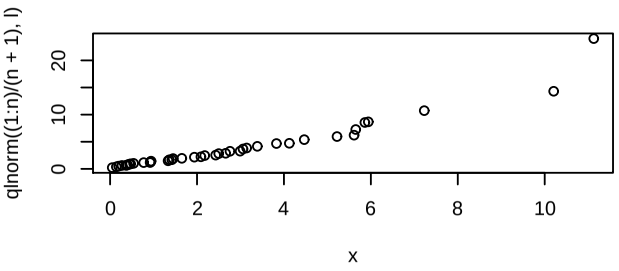
   CI of θ is $[0.17, 0.19]$ and length 0.02 for the Weibull distribution.

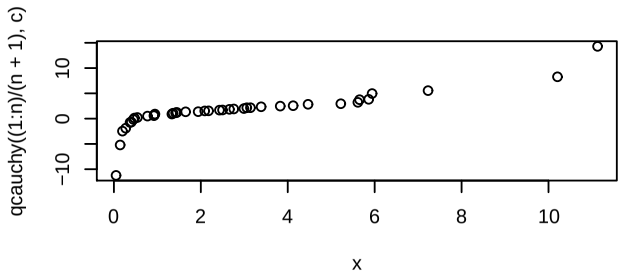
   CI of θ is $[0.216, 0.252]$ and length is 0.036 due to U(a,b)

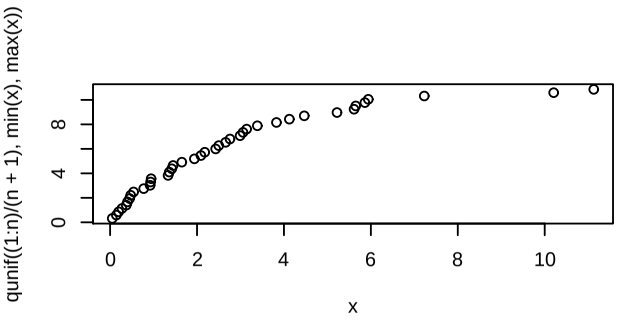

qweibull((1:n)/(n + 1), w)
0.0
1.0
2.0
3.0
0
2
4
6
8
10
x

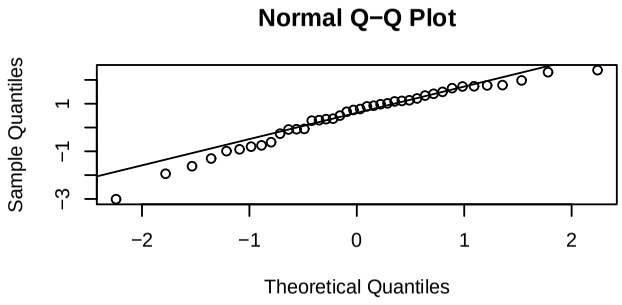

qgamma((1:n)/(n + 1), g)
0.0
1.5
3.0
0
2
4
6
8
10
x


qexp((1:n)/(n + 1), 1/mean(x))
0
4
8
0
2
4
6
8
10
x


qlnorm((1:n)/(n + 1), l)
0
10
20
0
2
4
6
8
10
x


qcauchy((1:n)/(n + 1), c)
−10
0
10
0
2
4
6
8
10
x


qunif((1:n)/(n + 1), min(x), max(x))
0
4
8
0
2
4
6
8
10
x


Normal Q−Q Plot
Sample Quantiles
−3
−1
1
−2
−1
0
1
2
Theoretical Quantiles

$\tilde{\theta} = (2 - X_{(1)})/(X_{(n)} - X_{(1)})$.

Bootstraping:

1. Generate 1000 times of n observations from U(1,5) or $U(min(x), max(x))$, ...

2. Derive 1000 $\hat{\theta}$,

3. Find the 2.5% and 97.5% quantiles.

Summary:

1. The parametric approach results in shorter CI, but it often leads to wrong estimate if the assumption is wrong.

2. The non-parametric approach results in wider CI, but it leads to a consistent estimate though the SD is bigger.