Estimating COVID-19 Cases in Shenzhen

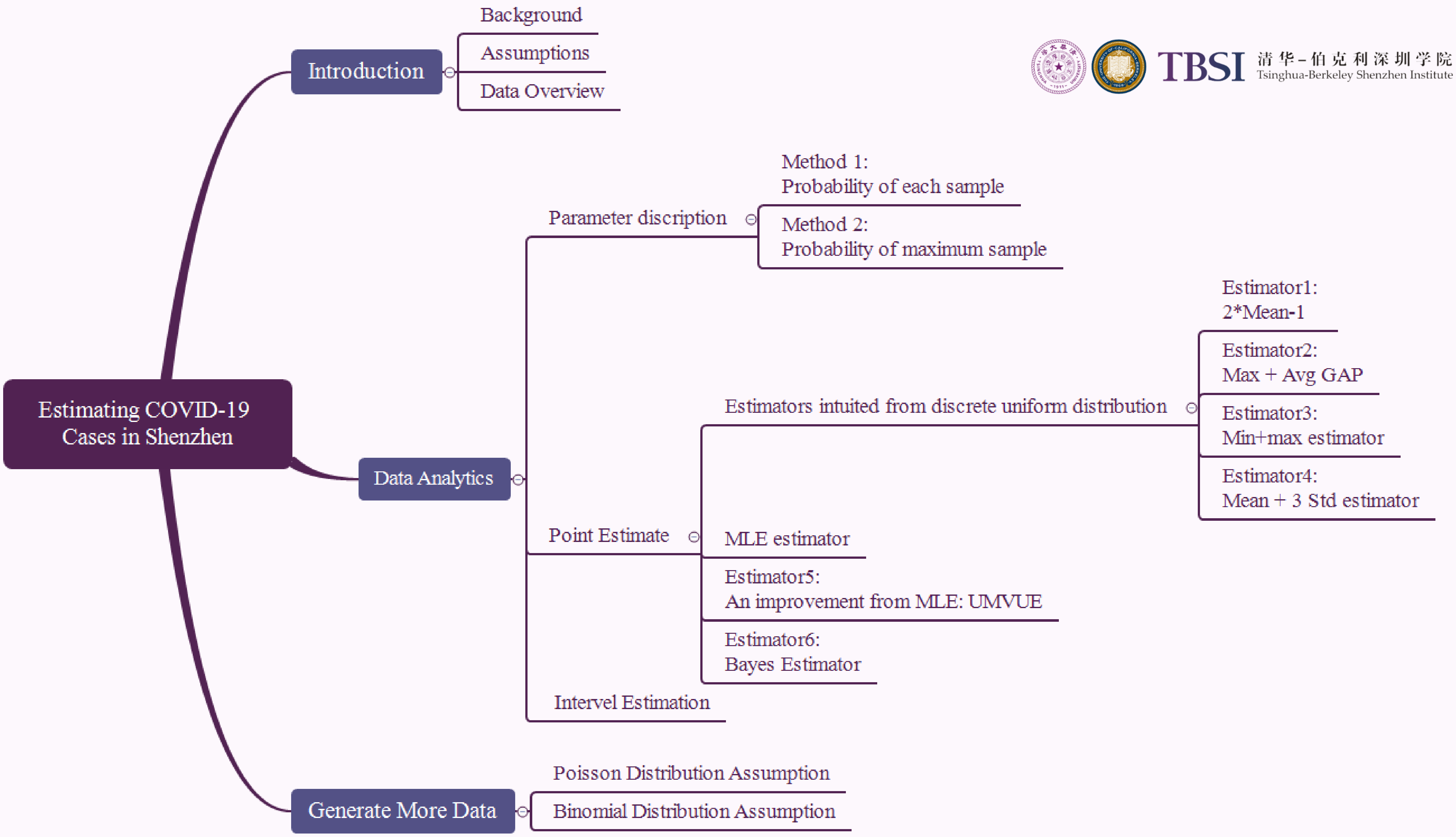
Mathematical Statistics & Application in R Take-Home Quiz-1

Members：郭俊麟 (GUO JUNLIN) / 陈梦玄 (CHEN MENGXUAN)

Student ID：2019270146 / 2019214596

Date：2020 / 03 / 16

TBSI 清华-伯克利深圳学院 Tsinghua-Berkeley Shenzhen Institute



Parameter Description

Method 1: Probability of each sample

$$P(x) = \frac{1}{\theta}$$

Estimating the upper limit of discrete uniform distribution

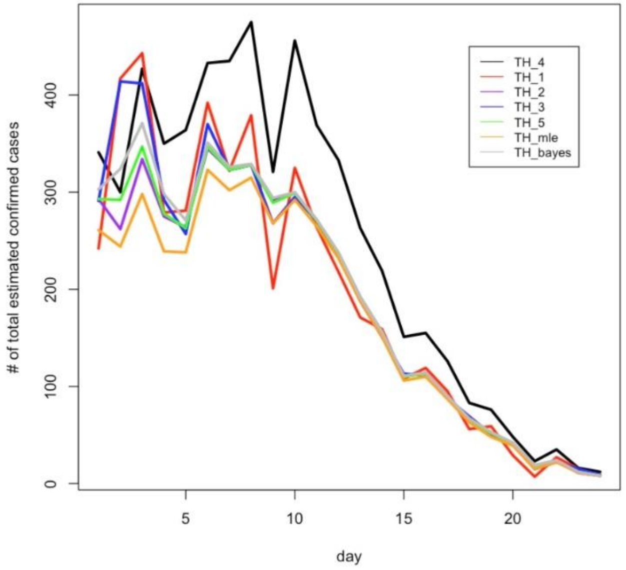
Method 2: Probability of maximum sample

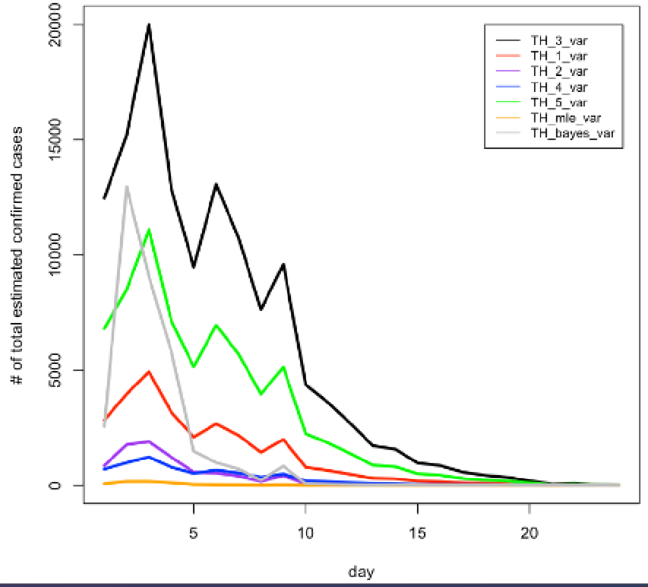
$$P(M = m) = \frac{C_{k-1}^{m-1}}{C_k^N}$$

The probability mass function (PMF) of getting the maximum ID

Point Estimation Conclusion

No.	Function	$E(\hat\theta)$	$Var(\hat\theta)$	TN
$\hat\theta_1$	$\frac{2}{n}\sum_{i=1}^{n} X_i - 1$	θ	$\frac{\theta^2}{3n}$	4616
$\hat\theta_2$	$X_{n:n} + \frac{1}{n-1}\sum_{i>j}(X_i - X_j - 1)$	$\frac{n\theta}{n+1}$	$\frac{n\theta^2}{(n+1)(n-1)(n+2)}$	4404
$\hat\theta_3$	$x_{1:n} + x_{n:n}$	θ	$\frac{2\theta^2}{n(n+2)} + \frac{2n^2\theta^2}{(n+1)^2(n+2)}$	4627
$\hat\theta_4$	$\frac{\sum_{i=1}^{n} X_i}{n} + 3\sqrt{\frac{\sum E(X_i-\bar{X})^2}{n-1}}$	$\frac{1}{2}\theta + \frac{2\sqrt{3}}{3}\theta$	$Var(\hat\theta_4) > \frac{\theta^2}{12n}$	5811
$\hat\theta_5$	$\frac{n+1}{n}x_{n:n} - 1$	θ	$\frac{\theta^2}{n(n+2)}$	4447
$\hat\theta_{MLE}$	$X_{n:n}$	$\frac{n}{n+1}\theta$	$\frac{n}{(n+1)^2(n+2)}\theta^2$	4125
$\hat\theta_{Bayes}$	$\frac{(X_{n:n}-1)(n-1)}{n-2}$	$\frac{n}{n+2}\theta$	$\frac{(x_{n:n}-1)(n-1)(x_{n:n}+1-n)}{(n-3)(n-2)^2}$	4565

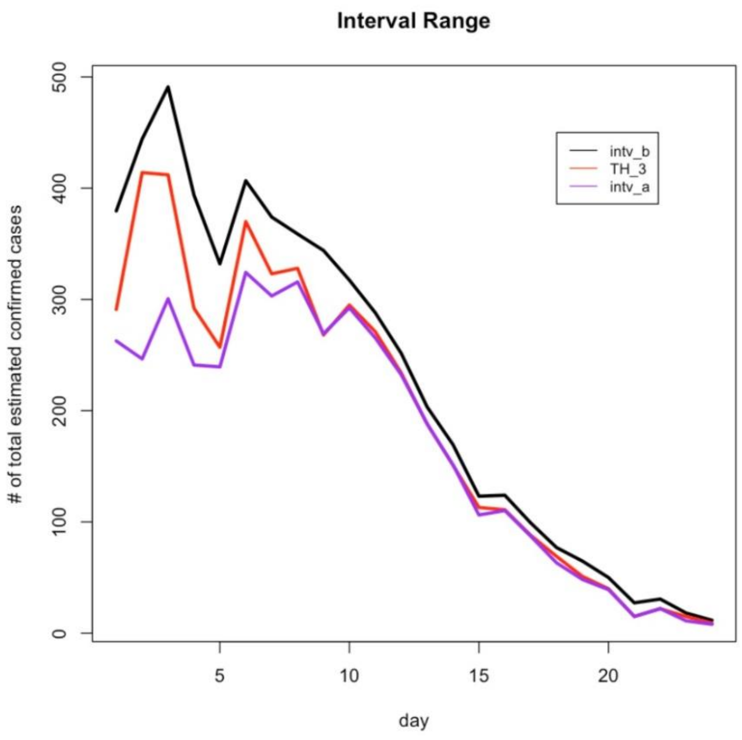




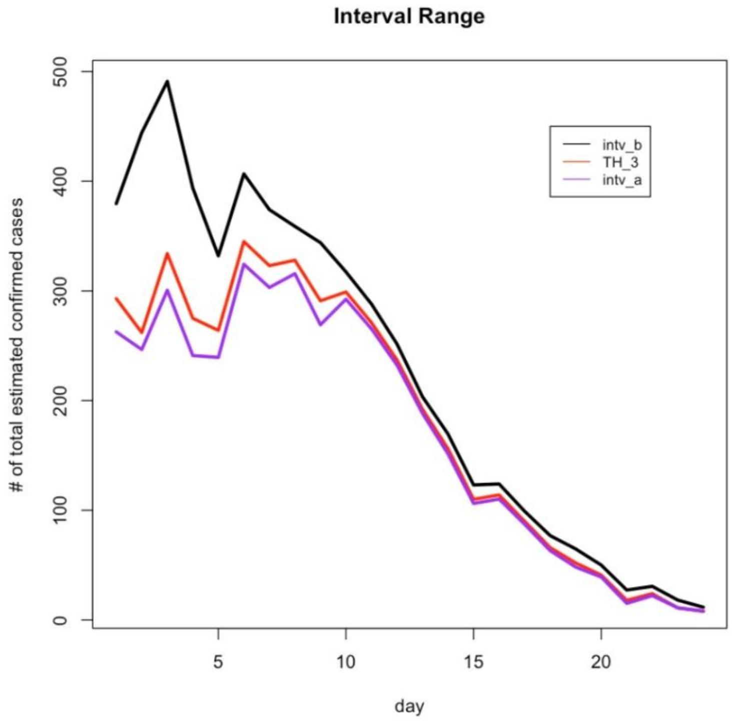
Interval Estimation Conclusion

清华-伯克利深圳学院
Tsinghua-Berkeley Shenzhen Institute

The sampling distribution of the quantile of the sample maximum is the graph $x^{1/k}$ from 0 to 1: the *p*-th to *q*-th quantile of the sample maximum *m* are the interval $[p^{1/k}N, q^{1/k}N]$. Inverting this yields the corresponding confidence interval for the population maximum of $[m/q^{1/k}, m/p^{1/k}]$.



P=5% Q=95% for theta_3

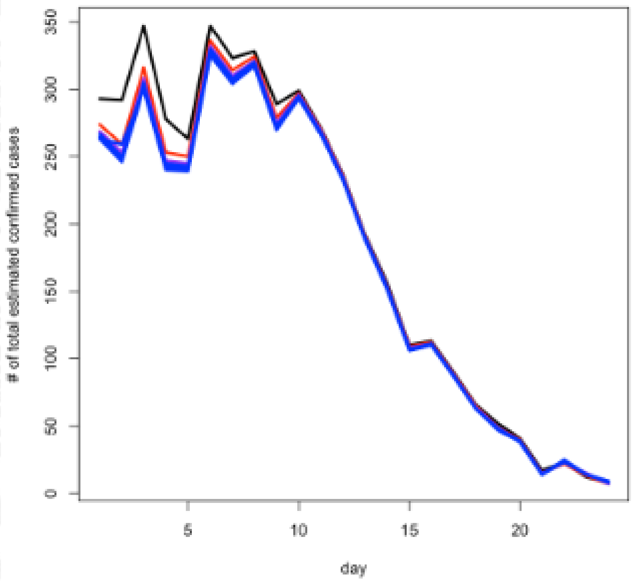


P=5% Q=95% for theta_5

Generate more Data

Here are the steps to iterate through data:The random ID would be generated under certain distribution.The new sample ID would update the sample mean $\bar{x}$ .The new sample ID might also update the maximum ID.the new $\bar{x}$ and maximum ID can further update the parameter of a assumed distribution.The random ID would be generated under an adjusted distribution. (Go back to the $2nd$ step.

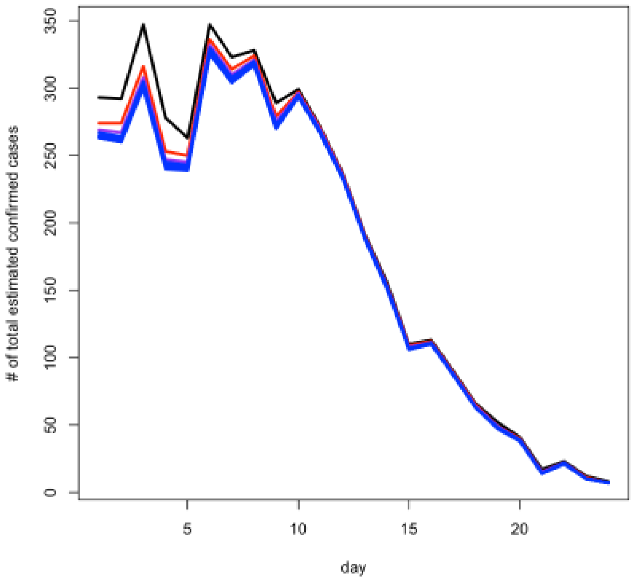
Poisson Distribution Assumption



$$P(X = x_{id}) = \frac{e^{-\lambda}\lambda^{x_{id}}}{x_{id}!}$$

$$\hat{\lambda}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_{id_i}$$

Binomial Distribution Assumption



$$P(X = x_{id}) = C_{x_{id}}^{N} p^{x_{id}}(1-p)^{N-x_{id}}$$

$$\hat{p} = \frac{\sum_{i=1}^{n} x_{id_i}}{nN} = \frac{\bar{x}_{id}}{N}$$