Part B

1-Sample K-S Test:

First we need to estimate parameters of Poisson, Geometric and Binomial distribution, separately. The following are the formulations:

Assume random variable X ~ Pois(λ), E[X] = λ and Var(X) = λ. We have $\lambda_{MME} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Assume random variable X ~ Geo(p), E[X] = 1/p and $Var(X) = \frac{1-p}{p^2}$. We have $p_{MME} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Assume random variable X ~ Bino(n, p), E[X] = np and Var(X) = npq where q = 1-p. We have

MME for binomial distribution:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

Since $E[X] = np$ and $Var(X) = np(1-p)$, we have.

$$np = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

$$np(1-p) = E[X^2] - (E[X])^2 = S^2$$

Therefore $1-p = \frac{S^2}{\bar{x}}$

$$\hat{p} = 1 - \frac{S^2}{\bar{x}} = \frac{\bar{x} - S^2}{\bar{x}}$$

$$\hat{n} = \frac{\bar{x}}{\hat{p}} = \frac{\bar{x}^2}{\bar{x} - S^2}$$

In our dataset, there are negative values for the number of new case and death. According to the definition of three distributions, we remove negative values before using MME to estimate parameters. For geometric distribution, we also remove zero values. The results of our experiments are as follows:

Please note that the estimated parameters of binomial distribution are negative. We just set the d as 1.

```
K-S Test for new_case :
try Poisson :
d =  0.96 reject H0 (The distribution of new_case in AL is the same as the distribution of new_case in AZ)
try Geometric :
d =  0.857241666501106 reject H0 (The distribution of new_case in AL is the same as the distribution of new_case in AZ)
try Binomial :
d =  1.0 reject H0 (The distribution of new_case in AL is the same as the distribution of new_case in AZ)


K-S Test for pnew_case :
try Poisson :
d =  0.33262588485348554 reject H0 (The distribution of pnew_case in AL is the same as the distribution of pnew_case in AZ)
try Geometric :
d =  0.3853414198776546 reject H0 (The distribution of pnew_case in AL is the same as the distribution of pnew_case in AZ)
try Binomial :
d =  1.0 reject H0 (The distribution of pnew_case in AL is the same as the distribution of pnew_case in AZ)


K-S Test for new_death :
try Poisson :
d =  0.4374328757733563 reject H0 (The distribution of new_death in AL is the same as the distribution of new_death in AZ)
try Geometric :
d =  0.34299463555022003 reject H0 (The distribution of new_death in AL is the same as the distribution of new_death in AZ)
try Binomial :
d =  1.0 reject H0 (The distribution of new_death in AL is the same as the distribution of new_death in AZ)


K-S Test for pnew_death :
try Poisson :
d =  0.6991087468651928 reject H0 (The distribution of pnew_death in AL is the same as the distribution of pnew_death in AZ)
try Geometric :
d =  0.5602295684113866 reject H0 (The distribution of pnew_death in AL is the same as the distribution of pnew_death in AZ)
try Binomial :
d =  1.0 reject H0 (The distribution of pnew_death in AL is the same as the distribution of pnew_death in AZ)
```

2-Samples K-S Test:

In our dataset, there are negative values for the number of new case and death. According to the practical meaning of new case and death, we remove negative values before using K-S test. The results are as follows:

```
K-S Test for new_case :
d =  0.9344715447154478 reject H0 (The distribution of new_case in AL is the same as the distribution of new_case in AZ)


K-S Test for pnew_case :
d =  0.2579185520361992 reject H0 (The distribution of pnew_case in AL is the same as the distribution of pnew_case in AZ)


K-S Test for new_death :
d =  0.32342657342657344 reject H0 (The distribution of new_death in AL is the same as the distribution of new_death in AZ)


K-S Test for pnew_death :
d =  0.5978260869565217 reject H0 (The distribution of pnew_death in AL is the same as the distribution of pnew_death in AZ)
```

Permutation Test:

In our dataset, there are negative values for the number of new case and death. According to the practical meaning of new case and death, we remove negative values before using K-S test. The results are as follows:

```
Permutation Test for new_case :
d =  0.0 reject H0 (The distribution of new_case in AL is the same as the distribution of new_case in AZ)


Permutation Test for pnew_case :
d =  0.693 accept H0 (The distribution of pnew_case in AL is the same as the distribution of pnew_case in AZ)


Permutation Test for new_death :
d =  0.007 reject H0 (The distribution of new_death in AL is the same as the distribution of new_death in AZ)


Permutation Test for pnew_death :
d =  0.0 reject H0 (The distribution of pnew_death in AL is the same as the distribution of pnew_death in AZ)
```

Part C

Before coding, we need to get the formulations of posterior distribution and MAP. The formulations are as follows:

Posterior distribution of Poisson distribution:

The pmf of Poisson distribution is $f(k)=Pr(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$ $(k\geq 0)$

Assume $X\sim Exp(\frac{1}{\beta})$ then $E[X]=\beta$ and $Var(x)=\beta^2$. The pdf is $f(x)=\frac{1}{\beta}e^{-\frac{1}{\beta}x}$ $(x>0)$

$$f(\lambda|D) \propto f(D|\lambda)\, f(\lambda)$$

$$=\left(\prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}\right)\cdot \frac{1}{\beta}e^{-\frac{1}{\beta}\lambda} = \frac{\lambda^{\sum x_i}\, e^{-(n+\frac{1}{\beta})\lambda}}{\prod_{i=1}^{n} x_i!}\cdot\frac{1}{\beta}$$

Since $\int f(\lambda|D)\, d\lambda = 1$, we have

$$C\int_0^\infty \frac{1}{\beta(\prod_{i=1}^{n} x_i!)}\lambda^{\sum x_i} e^{-(n+\frac{1}{\beta})\lambda}\, d\lambda = 1$$

$$C=\frac{\beta(\prod_{i=1}^{n} x_i!)}{\int_0^\infty \lambda^{\sum x_i} e^{-(n+\frac{1}{\beta})\lambda}\, d\lambda}$$

We know that $\int_0^\infty \lambda^a e^{b\lambda}\, d\lambda = (-b)^{-a-1}\,\Gamma(a+1) = (-b)^{-a-1}\, a!$

So $$C = \frac{\beta(\prod_{i=1}^{n} x_i!)}{\int_0^\infty \lambda^{\sum x_i} e^{-(n+\frac{1}{\beta})\lambda}\, d\lambda} = \frac{\beta(\prod_{i=1}^{n} x_i!)}{(n+\frac{1}{\beta})^{-\sum x_i - 1}\,(\sum x_i)!}$$

Therefore $$f(\lambda|D) = C\cdot \frac{\lambda^{\sum x_i} e^{-(n+\frac{1}{\beta})\lambda}}{\prod_{i=1}^{n} x_i!}\,\frac{1}{\beta} = \frac{(n+\beta)^{\sum x_i+1}\,\lambda^{\sum x_i}\, e^{-(n+\frac{1}{\beta})\lambda}}{\Gamma(\sum x_i+1)}$$

Therefore. we know that $f(\lambda|D)$ is a pdf of Gamma distribution.

where $\alpha = \sum x_i + 1$ and $\beta_G = n + \frac{1}{\beta}$ (β_G represents the β in Gamma distribution)

This is the ~~result~~ posterior distribution after 1-st iteration.

Similarly we have the posterior distribution after k-th iteration:

$$f(\lambda|D_1,\cdots D_k) = \frac{(\sum_{j=1}^{k} n_j + \beta)^{(\sum_{j=1}^{k}\sum_{i=1}^{n_j} x_{ij}+1)}\,\lambda^{\sum_{j=1}^{k}\sum_{i=1}^{n_j} x_{ij}}\, e^{-(\sum_{j=1}^{k} n_j + \frac{1}{\beta})\lambda}}{\Gamma(\sum_{j=1}^{k}\sum_{i=1}^{n_j} x_{ij}+1)}$$

This is also a pdf of Gamma distribution where $\alpha = \sum_{j=1}^{k}\sum_{i=1}^{n_j} x_{ij} + 1$ and $\beta_G = \sum_{j=1}^{k} n_j + \frac{1}{\beta}$ (β_G is the β in Gamma distribution and $\frac{1}{\beta}$ can be obtained from MME of data in the first 4 weeks)

MAP: $\log(f(\lambda|D)) = \log C + (\sum x_i)\log\lambda - (n+\frac{1}{\beta})\lambda$

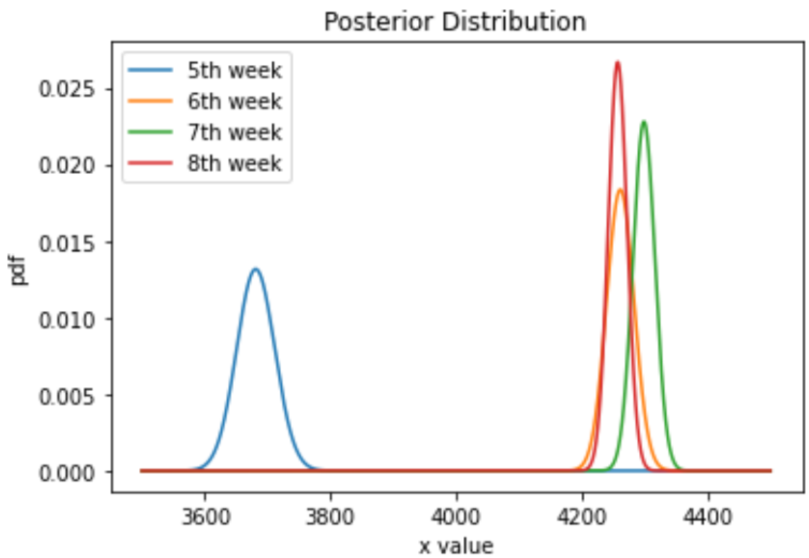
$$\log(f(\lambda|D))' = (\sum x_i)\frac{1}{\lambda} - (n+\frac{1}{\beta}) = 0$$

$$\lambda_{MAP} = \frac{\sum x_i}{n+\frac{1}{\beta}}$$

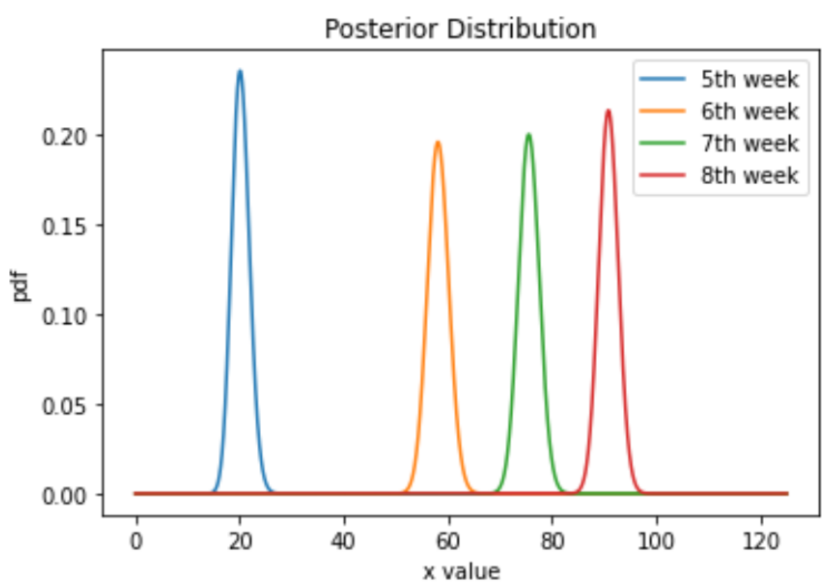
After K iterations, $$\lambda_{MAP} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} x_{ij}}{\sum_{j=1}^{k} n_j + \frac{1}{\beta}}$$

In our dataset, there are negative values for the number of new case and death. According to the practical meaning of new case and death, we remove negative values. The results are as follows:

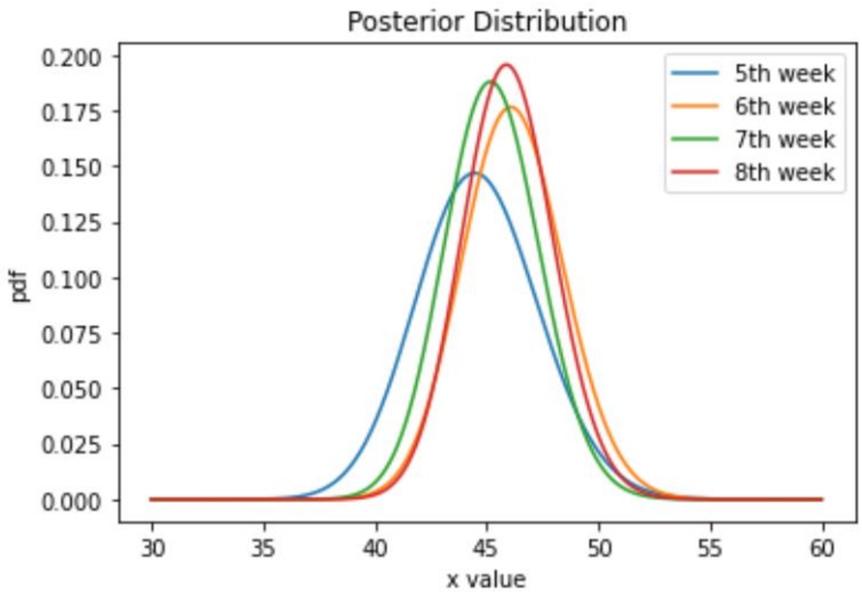
```
For new_case :
After inputting  5 th week, MAP of lambda is:  3681.108629698987
After inputting  6 th week, MAP of lambda is:  4261.131983302408
After inputting  7 th week, MAP of lambda is:  4298.797987952483
After inputting  8 th week, MAP of lambda is:  4257.009979089709
```




```
For pnew_case :
After inputting  5 th week, MAP of lambda is:  20.0931165757794
After inputting  6 th week, MAP of lambda is:  58.07097970536726
After inputting  7 th week, MAP of lambda is:  75.51008010934356
After inputting  8 th week, MAP of lambda is:  90.78564688904923
```



```
For new_death :
After inputting  5 th week, MAP of lambda is:  44.46989720998532
After inputting  6 th week, MAP of lambda is:  46.08627450980393
After inputting  7 th week, MAP of lambda is:  45.18005295675199
After inputting  8 th week, MAP of lambda is:  45.889245585874804
```




```
For pnew_death :
After inputting  5 th week, MAP of lambda is:  2.6335877862595423
After inputting  6 th week, MAP of lambda is:  2.650557620817844
After inputting  7 th week, MAP of lambda is:  3.1855955678670362
After inputting  8 th week, MAP of lambda is:  3.921455938697318
```