Q1.
Proof:
Let $\varepsilon = y^* - Ay$
We choose A to make $cov(\varepsilon, y) = 0$ -> $cov(y^* - Ay, y) = K(X^*, X) - AK(X, X)$
Thus A = $K(X^*, X)K(X, X)^{-1}$
Therefore $y^* = \varepsilon + Ay$

$$E(y^*|X, y, X^*) = E(\varepsilon + Ay|X, y, X^*) = K(X^*, X)K(X, X)^{-1}y$$

$$\begin{aligned} Var(\varepsilon + Ay|X, y, X^*) &= Var(\varepsilon|X, y, X^*) + Var(Ay|X, y, X^*) + Cov(\varepsilon, Ay) + Cov(Ay, \varepsilon) \\ &= Var(\varepsilon|X, y, X^*) \\ &= K(X^*, X^*) - K(X^*, X)K(X, X)^{-1}K(X, X^*) \end{aligned}$$

The rest terms are all zeros

Q2.

a) $\begin{bmatrix} y1 \\ y2 \\ y3 \\ y4 \\ y5 \end{bmatrix} \sim N(\begin{bmatrix} 0 \\ ... \\ 0 \end{bmatrix}, K)$ K is calculated using different kernels

Then use Multivariate normal to calculate probability, and take log to get logP(D|M)

b) $P(M_i|D) = \frac{1/2 \cdot P(D|M_i)}{\sum_{j=1}^{2} 1/2 \cdot P(D|M_j)}$ (Plug in numbers)

c) Apply the equation from Q1, treat y as $\begin{bmatrix} y1 \\ y2 \\ y3 \\ y4 \\ y5 \end{bmatrix}$, K from different kernels

Q3.

(a) Let Y_i be the number of deaths on the i^{th} day, $i = 1, 2, \ldots, 1096$, where 1096 is the sum of n_k's.
Assume $Y_i \overset{\text{iid}}{\sim} Poisson(\lambda)$, i.e.

$$\mathbb{P}(Y_i = k) = \frac{\lambda^k}{k!} \exp(-\lambda).$$

Then the MLE for λ is

$$\hat{\lambda}_{\text{ML}} = \frac{\sum_{i=1}^{1096} Y_i}{1096} = \frac{\sum_{k=0}^{9} kn_k}{1096} = 2.157.$$

(b) $$P(y_i \mid \lambda, P) = P_1 \frac{\lambda_1^{y_i}}{y_i!}\exp(-\lambda_1) + P_2 \frac{\lambda_2^{y_i}}{y_i!}\exp(-\lambda_2)$$

where $\theta = (P_1, P_2, \lambda_1, \lambda_2)$ $\quad P_1 + P_2 = 1$

(c). Introducing latent variable $Z_i = 0, 1$

Complete-data Likelihood:

$$L(Y, Z \mid \theta) = \prod_{i=1} P_1^{Z_i}(1-P_1)^{1-Z_i}\left(\frac{\lambda_1^{y_i}}{y_i!}\right)\exp(-\lambda_1)\cdot\left(\frac{\lambda_2^{y_i}}{y_i!}\right)\exp(-\lambda_2)$$

$$\log L = \sum_{i=1}^{n} Z_i \log P_1 + (1-Z_i)\log(1-P_1) + y_i \log(\lambda_1) - \lambda_1 - \log(y_i!) + y_i \log(\lambda_2) - \log(y_i!) - \lambda_2$$

↑ convert the sum for each sample to the sum of K deaths $k=0,\dots$

$$= \sum_{K=0}^{\infty} m_K \log P_1 + (n_K - m_K)\log(1-P_1) + m_K \cdot K \log(\lambda_1) - m_K\lambda_1 + (n_K - m_K)(K\log\lambda_2 - \lambda_2) + \text{const}.$$

where: $n_K = \#\{i : Y_i = K\}$: count of deaths $= K$.

$m_K = \sum_{Y_i = K} Z_i$: count of deaths $= K$ & with parameter λ_1

We construct Q function by taking expectation w.r.t. $m_K \mid \theta^{(t)}$

E-Step: $$E(m_K \mid Y, \theta^{(t)}) = \frac{n_K \cdot P_1^{(t)} \frac{[\lambda_1^{(t)}]^K}{K!}\exp(-\lambda_1^{(t)})}{P_1^{(t)}\frac{(\lambda_1^{(t)})^K}{K!}\exp(-\lambda_1^{(t)}) + (1-P_1^{(t)})\frac{(\lambda_2^{(t)})^K}{K!}\exp(-\lambda_2^{(t)})}$$

$$\therefore Q = \sum_K n_K \log(1-P_1) + \log\frac{P_1}{1-P_1} E(m_K \mid Y, \theta^{(t)}) + (n_K - E(m_K \mid Y, \theta^{(t)}))(K\log\lambda_2 - \lambda_2) + E(m_K \mid Y, \theta^{(t)}) \cdot (K\log\lambda_1 - \lambda_1)$$

M-Step: $$\frac{\partial Q}{\partial P_1} = 0, \quad \frac{\partial Q}{\partial \lambda_1} = 0, \quad \frac{\partial Q}{\partial \lambda_2} = 0 \Rightarrow$$

$$P_1^{(t+1)} = \frac{\sum_K m_K^{(t)}}{n}$$

$$\lambda_1^{(t+1)} = \frac{\sum_K m_K^{(t)} \cdot K}{\sum m_K^{(t)}}$$

$$\lambda_2^{(t+1)} = \frac{\sum_K (n_K - m_K^{(t)}) \cdot K}{\sum_K (n_K - m_K^{(t)})}$$

Q4.

$$f(Y_i|\theta_i) = \frac{\exp(-\theta_i t_i)(\theta_i t_i)^{Y_i}}{Y_i!}$$

$$f(\theta_i|\beta) = \beta \cdot \exp(-\theta_i \beta)$$

$$f(\beta) = \frac{1}{\Gamma(1/20)} \beta^{1/20-1} \exp(-\beta)$$

(a)

$$f(\theta_i|Y_i,\beta) \propto f(Y_i|\theta_i) f(\theta_i|\beta)$$

$$\propto \exp[-\theta_i(t_i+\beta)] \cdot \theta_i^{Y_i} \cdot t_i^{Y_i} \cdot \beta / Y_i!$$

not relavent to θ_i

$\therefore \theta_i | Y_i, \beta \sim \text{Gamma}(Y_i+1, \beta+t_i)$

$$f(\beta|\vec{Y},\vec{\theta}) \propto f(\beta) f(\theta|\beta) f(Y|\theta,\beta) \propto \beta^{\frac{1}{20}-1} \exp(-\beta) \prod_{i=1}^{20} \beta \exp(-\theta_i \beta)$$

has nothing to do with β

$$\propto \beta^{20+\frac{1}{20}-1} \cdot \exp\left(-\beta\left(\sum_{i=1}^{20}\theta_i + 1\right)\right)$$

$\therefore \beta | Y, \theta \sim \text{Gamma}\left(20+\frac{1}{20}, \sum_{i=1}^{20}\theta_i + 1\right)$

(b). Since we have conditional pdf:

$\theta_i | \cdots$

$\beta | \cdots$ **we can design a gibbs Sampler**

Step 1: $\theta_i^{t+1} \sim \text{Gamma}(1+Y_i, \beta^{(t)}+t_i)$ $i = 1, 2, \cdots 20$

Step 2: $\beta^{t+1} \sim \text{Gamma}\left(20+\frac{1}{20}, \sum_{i=1}^{20}\theta_i^{t+1}+1\right)$

repeat Step 1 & Step 2.