HW2.

problem 1 (a)

$$\hat{\pi}, \hat{\lambda}_{0,1:D}, \hat{\lambda}_{1,1:D} = \arg\max \sum_{i=1}^{n} \ln p(y_i|\pi) + \sum_{d=1}^{D} \Big(\ln p(\lambda_{0,d}) + \ln p(\lambda_{1,d})$$
$$+ \sum_{i=1}^{n} \ln p(x_{i,d}|\lambda_{y_i d}) \Big)$$

From above $L = \sum_{i=1}^{n} \ln p(y_i|\pi) + \sum_{d=1}^{D} \big(\ln p(\lambda_{0,d}) + \ln p(\lambda_{1,d}) + \sim \big)$.

$$= \sum_{i=1}^{n} \ln \pi^{y_i}(1-\pi)^{1-y_i} + \sum_{d=1}^{D} \Big(\ln \frac{\lambda_{0d} e^{-\lambda_{0d}}}{\Gamma(2)} + \ln \frac{\lambda_{1d} e^{-\lambda_{1d}}}{\Gamma(2)} + \sum_{i=1}^{n} \ln \frac{\lambda_{y_i d}^{x_i} e^{-\lambda_{y_i d}}}{x_i!} \Big)$$

$$\Rightarrow (\textstyle\sum y_i) \ln \pi + (n - \sum y_i) \ln(1-\pi).$$

$$\frac{\partial f}{\partial \pi} = (\textstyle\sum y_i) \frac{1}{\pi} + (n - \sum y_i) \frac{-1}{1-\pi} = 0$$

$$= \frac{\sum y_i}{\pi} - \frac{n - \sum y_i}{1-\pi} = 0$$

$$\Rightarrow (\textstyle\sum y_i)(1-\pi) = \pi(n - \sum y_i)$$

$$\Rightarrow \hat{\pi} = \frac{\sum y_i}{n}$$

(b) for $\hat{\lambda}_{y,d} = \sum_{i=1}^{n} \ln \pi^{y_i}(1-\pi)^{1-y} + \sum_{d=1}^{P} \Big(\ln \frac{\lambda_{0d} e^{-\lambda_{0d}}}{\Gamma(2)} + \ln \frac{\lambda_{1d} e^{-\lambda_{1d}}}{\Gamma(2)} + \sum_{i=1}^{n} \ln \frac{\lambda_{y_i d}^{x_i} e^{-\lambda_{yd}}}{x_i} \Big)$

denote: $\lambda_{y_i d} = y_i \lambda_{1,d} + (1-y_i)\lambda_{0d}$.

$$\Big(\ln(\lambda_{0,d}) - \lambda_{0,d} + \ln(\lambda_{1,d}) - \lambda_{1d} + \sum_{i=1}^{n} \big(\overbrace{x_{id} \ln(y_i \lambda_{1,d} + (1-y_i)\lambda_{0d})}^{\lambda_{y_i d}^{x_i}}$$
$$- \underbrace{(y_i \lambda_{1d} + (1-y_i)\lambda_{0d})}_{\ln e^{-\lambda_{yd}}} - \ln(x_{id}!) \big) \Big)$$

$$\frac{\partial}{\partial \lambda_{0,d}} = \frac{1}{\lambda_{0,d}} - 1 + \sum_{i=1}^{n} \left(\frac{1-y_i}{y_i \lambda_{1,d} + (1-y_i)\lambda_{0,d}} \cdot x_{i,d} - (1-y_i) \right) = 0$$

$y_i \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$

$$\frac{1}{\lambda_{0}d} - 1 + \sum_{i=1}^{n} \left(\frac{x_{i}d}{\lambda_{0}d} \cdot - 1 \right) = 0 \qquad \frac{1}{\lambda_{1,d}} - 1 + \sum_{i \mid y=1} \left(\frac{x_{i}d}{\lambda_{1,d}} - 1 \right) = 0$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\hat{\lambda}_{0,d} = \frac{1 + \sum_{i \mid y_i = 0} x_{i,d}}{1 + \sum_{i \mid y_i = 0} 1} \qquad \hat{\lambda}_{1,d} = \frac{1 + \sum_{i \mid y_i = 1} x_{i}d}{1 + \sum_{i \mid y_i = 1} 1}$$

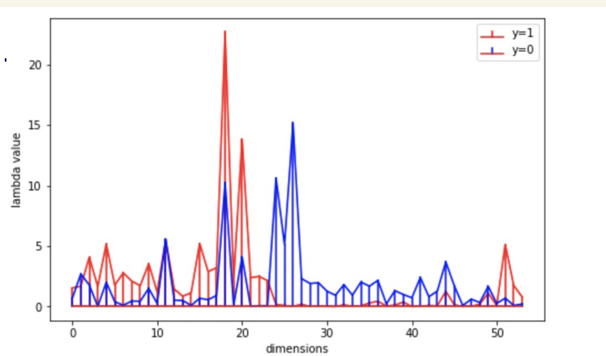
y

$$\hat{\lambda}_{y,d} = \frac{1 + \sum_{i \mid y_i = y} x_i d}{1 + \sum_{i \mid y_i = y} 1}$$

problem 2. (a)

True positive	False Positive
1702	560
False negative	True negative
611	2227

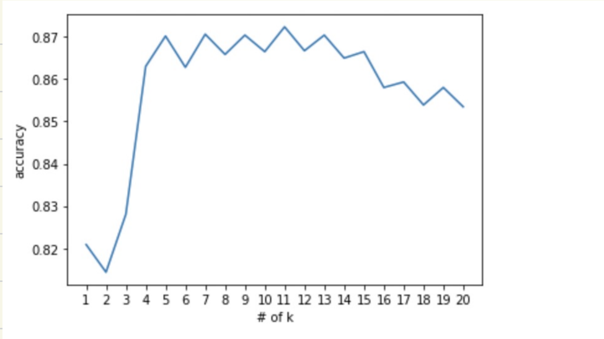
(b)



λ is the mean of poission

bayes classifier based on 16 and 52 dimesion it tend to decide the value belongs to y_1, the probability is higher, (for y_1 compare ~~to~~ y_0

(c)



Problem 3. (a).

(b)

(c). The drawback is Gaussian doesn't have a penalization term as RR in first homework thus it may overfit

(d)