Problem 3.

a. Bernoulli distribution

b/. $p(y|x) \propto p(x|y) \cdot p(y)$

$\hat{y} = \arg\max_k p(y=k|x)$

$p(y=1|x) \propto p(x|y=1) \cdot p(y=1)$

$= \lambda_1 e^{-\lambda_1 x} \cdot \frac{1}{2}$

$= \frac{1}{2} \lambda_1 e^{-\lambda_1 x}$

$p(y=0|x) \propto p(x|y=0) \cdot p(y=0)$

$= \lambda_0 e^{-\lambda_0 x} \cdot \frac{1}{2}$

$= \frac{1}{2} \lambda_0 e^{-\lambda_0 x}$

if $\hat{y}=1 \Rightarrow p(y=1|x) > p(y=0|x)$

$\frac{1}{2} \lambda_1 e^{-\lambda_1 x} > \frac{1}{2} \lambda_0 e^{-\lambda_0 x}$ $\quad e^{-\lambda_1 x}, e^{-\lambda_0 x} > 0$

$\log \frac{1}{2} \lambda_1 e^{-\lambda_1 x} - \log \frac{1}{2} \lambda_0 e^{-\lambda_0 x} > 0$

$\log \frac{\lambda_1}{\lambda_0} + \log e^{(\lambda_0 - \lambda_1)x} > 0$

$\log \frac{\lambda_1}{\lambda_0} + (\lambda_0 - \lambda_1)x > 0$

$(\lambda_0 - \lambda_1)x > -\log \frac{\lambda_1}{\lambda_0}$

$x > \frac{-\log \frac{\lambda_1}{\lambda_0}}{\lambda_0 - \lambda_1}$ if $\lambda_0 > \lambda_1$

$0 \leq x < \frac{-\log \frac{\lambda_1}{\lambda_0}}{\lambda_0 - \lambda_1}$ else

Problem 4:

$$P(y|w,X) = \prod_{i=1}^{N} \sigma(w^T x_i)^{y_i} (1-\sigma(w^T x_i))^{1-y_i}$$

$$E(w) = -\log P(y|w,X) = -\sum_{i=1}^{N} [y_i \log\sigma(w^T x_i) + (1-y_i)\log(1-\sigma(w^T x_i))]$$

$$\frac{dE(w)}{dw} = -\sum_{i=1}^{N} [y_i x_i \cdot \frac{1}{\sigma(w^T x_i)} \sigma(w^T x_i)(1-\sigma(w^T x_i)) + (1-y_i) \cdot \frac{1}{(1-\sigma(w^T x_i))} \cdot (-\sigma(w^T x_i)) (1-\sigma(w^T x_i))]$$

$$= -\sum_{i=1}^{N} [y_i x_i \cdot (1-\sigma(w^T x_i)) - (1-y_i) \cdot \sigma(w^T x_i)]$$

$$= -\sum_{i=1}^{N} [y_i x_i - y_i x_i \sigma(w^T x_i) - \sigma(w^T x_i) + y_i \sigma(w^T x_i)]$$

$$y_i = \sigma(w^T x_i) \quad \text{if } ||w|| \to \infty$$

$$-\sum_{i=1}^{N} [y_i x_i - y_i^2 x_i - y_i + y_i^2]$$

$$= -\sum_{i=1}^{N} [y_i(x_i - 1) + y_i^2(1 - x_i)] \quad y_i^2 = y_i \text{ for } y_i = 1 \text{ or } y_i = 0$$

$$= -\sum_{i=1}^{N} [y_i(x_i - 1) - y_i(x_i - 1)] = 0$$

so for maximum likelihood parameter w, $||w|| \to \infty$.

we can add regularization term $\frac{\lambda}{2}||w||^2$,

$$\hat{E}(w) = -\sum_{i=1}^{N} [y_i \log\sigma(w^T x_i) + (1-y_i)\log(1-\sigma(w^T x_i))] + \frac{\lambda}{2}||w||^2$$

add $||w||^2$ can penalize the weights, w can't be ∞, otherwise $\hat{E}(w)$ will be infinitive

Problem 5:

$$\frac{e^{(w_1^T x)}}{e^{(w_1^T x)} + e^{(w_2^T x)}} = \frac{1}{1+e^{(w_2^T - w_1^T)x}}$$

$$= \sigma\left((w_2^T - w_1^T)x\right)$$

Problem 6:

$$\frac{d\sigma(a)}{da} = \frac{e^{-a}}{(1+e^{-a})^2} = \frac{1}{1+e^{-a}} \cdot \frac{1+e^{-a}-1}{1+e^{-a}}$$

$$= \sigma(a) \cdot (1-\sigma(a))$$

Problem 7:

use $\Phi = x_1 x_2$ if $y =$ cross, $\Phi < 0$, $w^T x < 0$ for $w > 0$

if $y =$ circle, $\Phi > 0$, $w^T x > 0$ for $w > 0$

Problem 8:

$$p(y=c|x) = \frac{p(x|y=c)p(y=c)}{p(x)}$$

$$p(y=1|x) = p(y=0|x)$$

$$p(x|y=1)\pi_1 = p(x|y=0)\pi_0$$

$$\log p(x|y=1) + \log \pi_1 = \log p(x|y=0) + \log \pi_0$$

$$\log N(x|u_1, \Sigma_1) + \log \pi_1 = \log N(x|u_0, \Sigma_0) + \log \pi_0$$

$$-\frac{1}{2}\log|\Sigma_1| - \frac{1}{2}(x-u_1)^T \Sigma_1^{-1}(x-u_1) + \log \pi_1 = -\frac{1}{2}\log|\Sigma_0| - \frac{1}{2}(x-u_0)^T \Sigma_0^{-1}(x-u_0) + \log \pi_0$$

$$(x-u_1)^T \Sigma_1^{-1}(x-u_1) - (x-u_0)^T \Sigma_0^{-1}(x-u_0) + \log\frac{|\Sigma_1|}{|\Sigma_0|} + 2\log\frac{\pi_0}{\pi_1} = 0$$

$$x^T(\Sigma_1^{-1} - \Sigma_0^{-1})x + (2u_0\Sigma_0^{-1} - 2u_1\Sigma_1^{-1})x + u_1^2\Sigma_1^{-1} - u_0^2\Sigma_0^{-1} + \log\frac{|\Sigma_1|}{|\Sigma_0|} + 2\log\frac{\pi_0}{\pi_1} = 0$$

$$T = \{x \mid x^T A x + b^T x + c = 0\}$$

$$A = \Sigma_1^{-1} - \Sigma_0^{-1}$$

$$b^T = (2\mu_0 \Sigma_0^{-1} - 2\mu_1 \Sigma_1^{-1})$$

$$c = \mu_1^2 \Sigma_1^{-1} - \mu_0 \Sigma_0^{-1} + \log \left| \frac{\Sigma_1}{\Sigma_0} \right| + 2 \log \frac{\pi_0}{\pi_1}$$