EEL 709 MINOR I (T16)

1. (a)

$$\mathcal{L} = \theta^{100}(1-\theta)^{120}\, q_b^{50}(1-q_b)^{70}\, q_g^{90}(1-q_g)^{10}\, p_{bw}^{8}(1-p_{bw})^{42}\, p_{gw}^{15}(1-p_{gw})^{75}\, p_{bh}^{32}(1-p_{bh})^{38}\, p_{gh}^{5}(1-p_{gh})^{5}$$

(b)

$$\frac{d\mathcal{L}}{dp_{gw}} = \frac{15\mathcal{L}}{p_{gw}} - \frac{75\mathcal{L}}{1-p_{gw}}$$

$$= \frac{15\mathcal{L} - 15\mathcal{L}p_{gw} - 75\mathcal{L}p_{gw}}{p_{gw}(1-p_{gw})} \overset{\text{set}}{=} 0$$

$[>0 \text{ (for ML)}]$

$$\Rightarrow 15\mathcal{L}(1-p_{gw}-5p_{gw}) = 0$$

$$\Rightarrow \hat{p}_{gw} = \frac{1}{6} \quad \text{(for max. likelihood)}$$

(c) ML estimates:

$$\hat{\theta} = \frac{100}{220} = \frac{5}{11}$$

$$\hat{q}_g = \frac{90}{100} = \frac{9}{10}$$

$$\hat{q}_b = \frac{50}{120} = \frac{5}{12}$$

$$\hat{p}_{bw} = \frac{8}{50} = \frac{4}{25}$$

$$\hat{p}_{gh} = \frac{5}{10} = \frac{1}{2}$$

$$\hat{p}_{bh} = \frac{32}{70} = \frac{16}{35}$$

2.(a) a must be ≤ 82, otherwise the likelihood becomes 0 (the reading 82 has 0 probability). Similarly, b must be ≥ 89.
Given these constraints, minimising $(b-a)$ will maximise the likelihood. $\Rightarrow$ Maximise a, minimise b.
$\therefore$ The ML estimates are simply $a=82$, $b=89$.

(b) Not good, as no margin of uncertainty for unseen data. In other words, these estimates overfit the data.

(c) Prior on μ:

$$p(\mu) = \frac{1}{\sqrt{2\pi}\,\beta} e^{-\frac{(\mu-\alpha)^2}{2\beta^2}}$$

(Given: $\alpha = 75$, $\beta = 5$)

Likelihood: Let $X = \{x_1, \ldots x_5\}$

$$p(X \mid \mu, \sigma) = \prod_{i=1}^{5} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

Posterior:

$$p(\mu \mid X) = \frac{p(X\mid\mu)\cdot p(\mu)}{p(X)}$$

$p(X)$ → not a fn. of μ

$$\propto \prod_i \left\{ \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \right\} \frac{1}{\sqrt{2\pi}\beta} e^{-\frac{(\mu-\alpha)^2}{2\beta^2}}$$

Log posterior:

$$\log p(\mu|X) \underset{(=)}{\propto} \sum_i \left\{ -\frac{1}{2}\log 2\pi - \log\sigma - \frac{(x_i-\mu)^2}{2\sigma^2} \right\}$$

$$-\frac{1}{2}\log 2\pi - \log\beta - \frac{(\mu-\alpha)^2}{2\beta^2} - \log p(X)$$

For MAP estimate of μ:

$$\frac{\partial \log p(\mu|X)}{\partial\mu} = 0$$

$$\Rightarrow \sum_i \left(\frac{x_i-\mu}{\sigma^2}\right) - \frac{\mu-\alpha}{\beta^2} = 0$$

$$\frac{\sum_i x_i - 5\mu}{\sigma^2} - \frac{\mu-\alpha}{\beta^2} = 0$$

$$\frac{\sum_i x_i}{\sigma^2} + \frac{\alpha}{\beta^2} = \mu\left(\frac{5}{\sigma^2} + \frac{1}{\beta^2}\right)$$

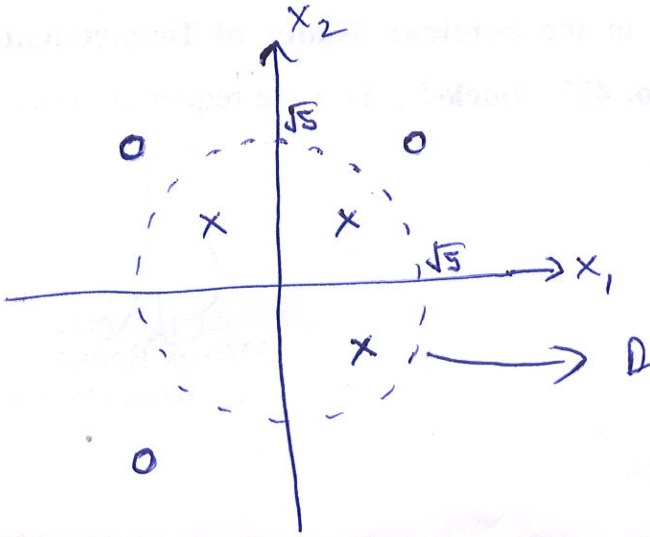
$$\Rightarrow \boxed{\hat{\mu}_{MAP} = \frac{\frac{\sum_i x_i}{\sigma^2} + \frac{\alpha}{\beta^2}}{\frac{5}{\sigma^2} + \frac{1}{\beta^2}}}$$

Plug in $\alpha = 75$, $\beta = 5$, $\sum_i x_i = 84+82+87+89+85 = 427$

$$\hat{\mu}_{MAP} = \frac{\frac{427}{\sigma^2} + 3}{\frac{5}{\sigma^2} + \frac{1}{25}} \Rightarrow \boxed{\begin{array}{l} \sigma = 1: \ \hat{\mu}_{MAP} = \underline{85.32} \\ \sigma = 10: \ \hat{\mu}_{MAP} = \underline{80.78} \end{array}}$$

2.(d) Clearly 85.32 seems a better estimate; the other estimate lies outside the range of the data entirely! It is being driven too strongly by the prior on μ, and hence is <u>underfitting</u> the data.

3.



O : $t = 1$

X : $t = -1$

Decision boundary:

$$\underline{x_1^2 + x_2^2 = 5}$$

(a) Not linearly separable, hence hard-margin linear SVM will fail, as there is no solution that will satisfy all its constraints.

(b) $\Phi(\underline{x}) = (x_1^2 + x_2^2)$ Just one dimension needed in new space

(c) For class 1, $x_1^2 + x_2^2 = 8$ $\forall$ points

For class −1, $x_1^2 + x_2^2 = 2$ "

$\Rightarrow$ Max. margin SVM will have the decision boundary $\underline{x_1^2 + x_2^2 = 5}$

(d) $K(\underline{x}, \underline{x}') = \Phi(\underline{x}) \cdot \Phi(\underline{x}')$

$$= (x_1^2 + x_2^2) \cdot (x_1'^2 + x_2'^2) = x_1^2 x_1'^2 + x_1^2 x_2'^2 + x_2^2 x_1'^2 + x_2^2 x_2'^2$$