Exercise 1

(a)

$p(x|\theta) = \theta^2 \cdot x \cdot e^{-\theta \cdot x} \cdot u(x)$ όπου $u(x) = \begin{cases} 1 & \text{για } x > 0 \\ 0 & x < 0 \end{cases}$

$L(\theta|x) = f(x_1, x_2, \dots, x_n | \theta)$

$= \prod_{i=1}^{N} f(x_i|\theta) = \theta^2 \cdot x_1 \cdot e^{-\theta x_1} \cdot u(x_1) \times \theta^2 \cdot x_2 \cdot e^{-\theta x_2} \cdot u(x_2)$

$\times \dots \theta^2 x_N \cdot e^{-\theta x_N} \cdot u(x_N)$

$= \theta^{2N} \cdot e^{-\theta \cdot \sum_{i=1}^{N} x_i} \cdot \prod_{i=1}^{N} u(x_i)$

$\ln\left(L(\theta|x)\right) = \ln\left(\theta^{2N} \cdot e^{-\theta \cdot \sum_{i=1}^{N} x_i} \cdot \prod_{i=1}^{N} u(x_i)\right)$

$= \ln \theta^{2N} + \ln e^{-\theta \cdot \sum_{i=1}^{N} x_i} + \ln \prod_{i=1}^{N} u(x_i)$

$= 2N \cdot \ln\theta - \theta \cdot \sum_{i=1}^{N} x_i \cdot \cancel{\ln e}^{\,1} + \ln \prod_{i=1}^{N} u(x_i)$

Επομένως:

$\frac{\partial \ln(L(\theta|x))}{\partial \theta} = \frac{1}{\theta} \cdot 2n - \sum_{i=1}^{N} x_i$

Και άρα: $\tilde{\theta}_{ML} = \arg\max_\theta p(X|\theta)$

$\hat{\theta}_{ML}: \frac{1}{\theta} \cdot 2n - \sum_{i=1}^{N} x_i = 0 \Leftrightarrow \boxed{\hat{\theta} = \frac{2N}{\sum_{i=1}^{N} x_i}}$

(b) $E[\hat{\theta}_{ML}] = E\left[\frac{2N}{\sum x_i}\right] = 2N \cdot \frac{1}{\sum_{i=1}^{N} E[x_i]} = 2\cancel{N} \cdot \frac{1}{\cancel{N} E[x_i]}$

$= \frac{2}{\frac{5 \cdot 2}{11,9}} = \frac{11,9 \cdot 2}{10} = \underline{2,38}$

Exercise 2

(α)

$$\hat{\theta}_{map} = \arg\max_\theta p(\theta|X) = \arg\max_\theta \frac{p(\theta)\cdot P(X|\theta)}{p(X)}$$

$$= \arg\max_\theta \left(p(\theta)\cdot P(X|\theta)\right) = \arg\max_\theta \ln\left(p(\theta)\cdot P(X|\theta)\right)$$

Συνεπώς:

$$L\left(p(\theta)\cdot P(X|\theta)\right) = p(\theta)\cdot\prod_{i=1}^{N} p(x_i|\theta)$$

$$= \frac{1}{2\pi\sigma_0^2}\cdot e^{-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}}\cdot \theta^2\cdot x_1\cdot e^{-\theta x_1}\cdot u(x_1)\times\ldots\times\theta^2\cdot x_n\cdot e^{-\theta x_n}\cdot u(x_n)$$

$$= \frac{1}{2\pi\sigma_0^2}\cdot e^{-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}}\cdot\prod_{i=1}^{N}\theta^2\cdot x_i\cdot e^{-\theta x_i}\cdot u(x_i)$$

$$= \frac{1}{2\pi\sigma_0^2}\cdot e^{-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}}\cdot\theta^{2N}\cdot\prod_{i=1}^{N}x_i\cdot e^{-\theta\sum_{i=1}^{N}x_i}\cdot\prod_{i=1}^{N}u(x_i)$$

Άρα:

$$\ln(L) = \ln\left[\frac{1}{2\pi\sigma_0^2}\cdot e^{-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}}\cdot\theta^{2N}\cdot\prod_{i=1}^{N}x_i\cdot e^{-\theta\sum_{i=1}^{N}x_i}\cdot\prod_{i=1}^{N}u(x_i)\right]$$

$$= \ln\frac{1}{2\pi\sigma_0^2} + \frac{-(\theta-\theta_0)^2}{2\sigma_0^2}\cdot\cancelto{1}{\ln e} + 2N\cdot\ln\theta + \ln\prod_{i=1}^{N}x_i - \theta\sum_{i=1}^{N}x_i$$

$$+ \ln\prod_{i=1}^{N}u(x_i)$$

Οπότε:

$$\frac{\partial \ln(L)}{\partial \theta} = \frac{-(2\theta - 2\theta_0)}{2\sigma_0^2} + 2N\frac{1}{\theta} - \sum_{n=1}^{N} x_i$$

Θέλουμε $\frac{\partial \ln(L)}{\partial \theta} = 0 \iff$

$$\frac{2\theta_0 - 2\theta}{2\sigma_0^2} + 2N \cdot \frac{1}{\theta} - \sum_{n=1}^{N} x_i \overset{(\times 2\sigma_0^2 \cdot \theta)}{\iff}$$

$$\theta(2\theta_0 - 2\theta) + 2N\sigma_0^2 - 4\sigma_0^2 \theta \cdot \sum_{i=1}^{N} x_i = 0 \iff$$

$$-\cancel{2}\theta^2 + \cancel{2}\theta_0\theta + \cancel{2}N\sigma_0^2 - \cancel{4}^{2}\sigma_0^2 \sum_{i=1}^{N} x_i \cdot \theta = 0 \iff$$

$$-\theta^2 + (\theta_0 - \sigma_0^2 \cdot \sum_{i=1}^{N} x_i) \cdot \theta + 2N\sigma_0^2 = 0 \iff$$

$$\theta^2 - (\theta_0 - \sigma_0^2 \cdot \sum_{i=1}^{N} x_i) \cdot \theta - 2N\sigma_0^2 = 0 \quad (1)$$

$$\Delta = (\theta_0 - \sigma_0^2 \cdot \sum_{i=1}^{N} x_i)^2 + 8N\sigma_0^2 > 0$$

$$\theta_{1,2} = \frac{(\theta_0 - \sigma_0^2 \sum_{i=1}^{N} x_i) \pm \sqrt{(\theta_0 - \sigma_0^2 \sum x_i)^2 + 8N\sigma_0^2}}{2} \quad (*)$$

$$\theta_1 = \hat{\theta} = \frac{(\theta_0 - \sigma_0^2 \sum x_i) + \sqrt{(\theta_0 - \sigma_0^2 \sum x_i)^2 + 8N\sigma_0^2}}{2}$$

(*) !! Έστω $c = (\theta_0 - \sigma_0^2 \sum x_i)$ και $a = 8N\sigma_0^2 > 0$
Συνεπώς $c < \sqrt{c^2 + a}$ όπου a ένας θετικός αριθμός.
Άρα $\hat{\theta}_2 < 0$, απορ. γιατί θέλουμε $\hat{\theta} > 0$

b) $\frac{(1)}{\sigma_0^2} \Leftrightarrow \frac{\theta^2}{\sigma_0^2} - \theta\left(\frac{\theta_0}{\sigma_0^2} - \frac{\cancel{\sigma_0^2}\sum x_i}{\cancel{\sigma_0^2}}\right) - 2N\cancel{\sigma_0^2}\cdot\frac{1}{\cancel{\sigma_0^2}} = 0 \Leftrightarrow$

$$\frac{\theta^2}{\sigma_0^2} - \theta\left(\frac{\theta_0}{\sigma_0^2} - \sum_{i=1}^{N} x_i\right) - 2N = 0 \quad (2)$$

i) Όταν το $N \to \infty$ το το $\hat{\theta}$ συγκλίνει στο πραγματικό θ γιατί είναι "σαν" να έχουμε όλα τα σημεία του πληθυσμού!

ii) Όταν το $\sigma_0^2 >>$ τότε η (2) γίνεται:

$$\cancel{\frac{\theta^2}{\sigma_0^2}}^{\to 0} - \theta\left(\cancel{\frac{\theta_0}{\sigma_0^2}}^{\to 0} - \sum_{i=1}^{N} x_i\right) - 2N = 0 \Leftrightarrow$$

$$\theta_{MAP} = \frac{2N}{\sum_{i=1}^{N} x_i} = \hat{\theta}_{ML} \quad \text{(Exercise 1)}$$

Το παραπάνω αποτέλεσμα μπορούμε να το αντιληφθούμε και με απλή λογική αφού όταν το σ_0^2 από τη "prior" γνώση είναι πολύ μεγάλο τότε και η "αβεβαιότητα" για το $\hat{\theta}_0$ είναι τόσο μεγάλη που δεν επηρεάζει τη $\hat{\theta}_{MAP}$ εκτιμήτρια.

iii) Αντίθετα, όταν $\sigma_0^2 <<$ τότε $\hat{\theta}_{MAP} = \theta_0$ γιατί είμαστε πολύ "σίγουροι" για τη γνώση που έχουμε από την prior πληροφορία.