Exercise 5.

$$\log p_k(x) = \log\left(\frac{\pi_k \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{1}{2\sigma^2}(x-\mu_k)^2)}{\sum_{l=1}^{K} \pi_l \frac{1}{\sqrt{2\pi}\sigma} \exp(\frac{1}{2\sigma^2}(x-\mu_l)^2)}\right)$$

$$= \log(\pi_k \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{1}{2\sigma^2}(x-\mu_k)^2)) - \log(\sum_{l=1}^{K} \pi_l \frac{1}{\sqrt{2\pi}\sigma} \exp(\frac{1}{2\sigma^2}(x-\mu_l)^2))$$

$$= -\frac{1}{2\sigma^2}(x-\mu_k)^2 + \log(\pi_k) + \log\frac{1}{\sqrt{2\pi}\sigma}) - \log(\frac{1}{\sqrt{2\pi}\sigma} \sum_{l=1}^{K} \pi_l \exp(\frac{1}{2\sigma^2}(x-\mu_l)^2))$$

$$= -\frac{1}{2\sigma^2}(x^2 - 2x\mu_k + \mu_k^2) + \log(\pi_k) + \log(\frac{1}{\sqrt{2\pi}\sigma}) - \log(\frac{1}{\sqrt{2\pi}\sigma}) + \log(\sum_{l=1}^{K} \pi_l \exp(\frac{1}{2\sigma^2}(x-\mu_l)^2))$$

$$= -\frac{x^2}{2\sigma^2} + \frac{x\mu_k}{\sigma^2} - \frac{\mu_k^2}{2\sigma^2} + \log(\pi_k) + \log(\sum_{l=1}^{K} \pi_l \exp(\frac{1}{2\sigma^2}(x-\mu_l)^2))$$

Assumption: $N(\mu_k, \sigma^2)$ distribution.

After taking the log of $p_k(x)$, we get the above results. Since the two parts $-\frac{x^2}{2\sigma^2}$ and $\log(\sum_{l=1}^{K} \pi_l \exp(\frac{1}{2\sigma^2}(x-\mu_l)^2))$ keep constant for the kth class, only the middle part $x\frac{\mu_k}{\sigma^2} - \frac{\mu_k^2}{2\sigma^2} + \log(\pi_k)$ left.

So when (4.12) $p_k(x)$ is largest, (4.13) $x\frac{\mu_k}{\sigma^2} - \frac{\mu_k^2}{2\sigma^2} + \log(\pi_k)$ is largest.

Exercise 6. a). π: probability that a woman develops an infection after abdominal delivery. $\Rightarrow \pi = P(\text{Infection} = \text{Yes} \mid \text{antibiotics})$

Likelihood: $P(\text{infection} = \text{Yes} \mid \text{antibiotics} = \text{Yes}) = \frac{11+1}{87+11+2+17+1} = 0.102$

$P(\text{infection} = \text{Yes} \mid \text{antibiotics} = \text{No}) = \frac{23+8+28}{9+3+23+32+30+8+28} = 0.444$

b).

	coefficient	odds Ratio = exp(coefficient)	p-value
Intercept	-0.8207	0.4401	0.0971
antibiotics	-3.2544	0.0386	1.37 e-11
plan	-1.0720	0.3423	0.0117
risk	2.0299	7.6133	8.25 e-06

Give the 95% CI for the odds ratio, antibiotics, plan and risk with p-value < 0.05 are significant associated with infection

- Antibiotics: Mothers dose antibiotics have 0.0386 times the odds of infection than those do not does antibiotics, controlling other variables in the model
- Plan: Mothers plan abdominal delivery have 0.3423 times the odds of infection than those do not plan abdominal delivery, controlling other variables.
- Risk: Mothers have a risk factor have 7.6133 times the odds of infection than those do not have a risk factor, controlling other variables in the model.

$$\text{log-odds} = -0.8207 - 3.2544 \cdot \text{antibiotics} + 2.0299 \cdot \text{risk} - 1.0720 \cdot \text{plan}$$

c). risk = 1, plan = 1, anitibiotics = 1

$\text{log-odds} = -0.8207 - 3.2544 + 2.0299 - 1.0720 = -3.1172$

$\Pr(\text{infection} = 1) = \frac{e^{-3.1172}}{1 + e^{-3.1172}} = 0.0424$ ← probability of an infection d). From the coefficients of LDA, "antibiotics" has the strongest associated weight with LD1. $Y=0$: no infection, $Y=1$: infection

$$P(Y=1 \mid X) = \frac{\pi_1 f_1(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)} \quad \text{——— ①}$$

Given the LDA output, $\pi_0 = 0.7171$, $\pi_1 = 0.2829$.

With risk=1, plan=1, antibiotics=1,
$$\begin{cases} f_0(x) = P(X=x \mid Y=0) = \dfrac{17}{9+3+87+32+30+2+17} = 0.0944 \\ f_1(x) = P(X=x \mid Y=1) = \dfrac{1}{23+4+8+28+1} = 0.0141 \end{cases}$$

Plug the above numbers into ①, and we get the probability of a woman describe in 1(c) is 0.0556