1

a) Der Maximum Likelihood Term wird logarithmiert, da viele Wahrscheinlichkeiten Potenzen enthalten, die durch das logarithmieren in eine andere Form (Multiplikation statt Potanz) gebracht werden, und man damit deutlich besser rechnen kann.
Die Position der Maxima ändert sich dadurch allerdings nicht, sonst würde das logarithmieren das Ergebnis verfälschen.
$p(D|\vartheta) = \prod_{k=1}^{n} p(x_k|\vartheta)$
$l(D|\vartheta) = log(\prod_{k=1}^{n} p(x_k|\vartheta))$
$= \sum_{k=1}^{n} log(p(x_k|\vartheta))$

b) $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$
$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$

c) 1. Es wird angenommen, das die gesuchte klassenabhängige Dichte nur von den Daten der jeweiligen Klassen abhängt.
2. Es wird angenommen, dass die gesuchte Verteilung parametrisch geschrieben werden kann, die Form von $p(x|\Theta)$ also bekannt ist.

d) 1. Der gewählte Datensatz ist zu klein und damit nicht repräsentativ für die tatsächlichen Parameter
2. Die getroffene Annahme über die Verteilung der Zufallsvariable ist falsch.
3. Die Stichprobe ist untypisch und es wird daraus ein stark von der Wirklichkeit abweichender Schätzer bestimmt.

2

Zunächst: Schätzer für μ bei unbekanntem σ:
$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(x-\mu)^2}{2\sigma^2})$
$l(\mu, \sigma) = -\frac{n}{2} log(2\pi\sigma^2) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}$
$\frac{\delta l(\mu,\sigma)}{\delta\mu} = \frac{\sum_{i=1}^{n} x_i - n\mu}{\sigma^2} \overset{!}{=} 0$
$\Rightarrow 0 = \sum_{i=1}^{n} x_i - n\mu \Rightarrow \sum_{i=1}^{n} x_i = n\mu$
$\hat{\mu} = \overline{x}$
Da der Erwartungswert nicht von der Varianz abhängt ist dies auch gleich dem Wert bei bekanntem σ.

3

a) i) $\hat{\mu_n}$ ist der (A-Posteriori) Schätzer. $\hat{\mu_0} = \mu_0$ (siehe unten).
Je größer n ist, desdo genauer nähert sich der Schätzer an den tatsächlichen Wert von μ an.

ii) μ_0 ist A-Priori Vermutung für den Erwartungswert. Er ist nicht abhängig von der Größe von n.

iii) σ_0^2 ist A-Priori Vermutung für die Varianz. Sie ist nicht abhängig von der Größe von n.

iv) σ^2 ist die bekannte Varianz. Sie verändert sich nicht, wenn mehr oder weniger Werte betrachtet werden, da sie eine Konstante ist.

b) σ_n^2 ist ein Wert, der mithilfe der A-Priori Vermutung für σ und dem tatsächlichen Wert berechnet wird und somit bei einer rekursiven Schätzung ein Maß für die Genauigkeit der Schätzung liefert.

e)

f)

g)