LFD1 Problem Set for Week 4
Solutions

Tutor: Michael Schouten

Question 1

To evaluate $E(X) = \int\limits_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$
make the substitution $z = \frac{x-\mu}{\sigma} \Rightarrow dz = \frac{1}{\sigma} dx$ and rewrite the integral as:

$$\int\limits_{-\infty}^{\infty} (z\sigma + \mu) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

Denote N(z) as $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ and recognize that this is the p.d.f for a Gaussian with $\mu = 0$ and $\sigma^2 = 1 \Rightarrow \int\limits_{-\infty}^{\infty} N(z) dz = 1$
Using this notation and rearranging terms yields:

$$\mu \int\limits_{-\infty}^{\infty} N(z) dz + C_1 \int\limits_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz$$

where C_1 is a constant that does not depend on z. As noted above, the integral on the left evaluates to 1. The integral on the right can be expressed as:

$$\lim_{a \to -\infty} \int\limits_{a}^{0} z e^{-\frac{1}{2}z^2} dz + \lim_{b \to \infty} \int\limits_{0}^{b} z e^{-\frac{1}{2}z^2} dz$$

where both terms in the expression evaluate to 0. Hence:

$$E(X) = \mu \cdot 1 + C_1 \cdot 0 = \mu$$

Question 2

To evaluate Var(X)=$E(X-\mu)^2 = \int\limits_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$
begin by using the following relation:

$$Var(X) = E(X^2) - (E(X))^2$$

From question 1, $(E(X))^2 = \mu^2$. To evaluate $E(X^2) = \int\limits_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ make the same substitution as in question 1, i.e. $z = \frac{x-\mu}{\sigma}$, and rewrite the integral as:

$$\int\limits_{-\infty}^{\infty} (z\sigma + \mu)^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$= \mu^2 \int\limits_{-\infty}^{\infty} N(z) dz + C \int\limits_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz + \sigma^2 \int\limits_{-\infty}^{\infty} z^2 e^{-\frac{1}{2}z^2} dz$$

In question 1 it was shown that the first two integrals evaluate to 1 and 0 respectively. It can readily be shown that the third integral evaluates to 1 using integration by parts[1]. Hence:

$$Var(X) = \left[E(X^2)\right] - (E(X))^2 = \left[\mu^2 \cdot 1 - C \cdot 0 + \sigma^2 \cdot 1\right] - \mu^2 = \sigma^2$$

Question 3

The likelihood of the data $L(\mathbf{x}|\mu, \sigma^2)$ is specified as:

$$\begin{aligned} L(\mathbf{x}|\mu, \sigma^2) &= \prod_{i=1}^{P} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(x_i - \mu)^2\right] \\ &= \frac{1}{(2\pi\sigma^2)^{P/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{P}(x_i - \mu)^2\right] \end{aligned}$$

and the log likelihood,ℓ, is therefore

$$\begin{aligned} \ell(\mathbf{x}|\mu, \sigma^2) &= ln\left(\frac{1}{(2\pi\sigma^2)^{P/2}}\right) + ln\left(\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{P}(x_i - \mu)^2\right]\right) \\ &= (-P/2) ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{P}(x_i - \mu)^2 \end{aligned}$$

Since μ is not known, replace it with an estimate, $\hat{\mu}$. To find the value of $\hat{\mu}$ that maximizes ℓ (i.e. the maximum likelihood estimate of μ), solve the appropriate first order condition:

$$\frac{\partial \ell}{\partial \hat{\mu}} = \frac{1}{\sigma^2}\sum_{i=1}^{P}(x_i - \hat{\mu}) = 0$$

[1] Recall that the formula for this method is $\int udv = uv - \int vdu$. Let $u = z$ and $dv = ze^{-1/2z^2} \Rightarrow v = -N(z)$ The integral can then be written as $-ze^{-1/2z^2}\Big|_{-\infty}^{\infty} + \int N(z)dz$ The integral on the right evaluates to 1. You can use L'Hopital's Rule to show that the left term is 0.

$$\Longrightarrow \sum_{i=1}^{P} x_i - P\hat{\mu} = 0$$

$$\Longrightarrow \hat{\mu} = \frac{1}{P}\sum_{i=1}^{P} x_i$$

Similarly, to get the maximum likelihood estimate for σ^2:

$$\frac{\partial \ell}{\partial \hat{\sigma^2}} = \frac{P}{2\hat{\sigma^2}} - \frac{1}{\left[\hat{\sigma^2}\right]^2}\sum_{i=1}^{P}(x_i - \mu)^2 = 0$$

$$\Longrightarrow P - \frac{1}{\hat{\sigma^2}}\sum_{i=1}^{P}(x_i - \mu)^2 = 0$$

$$\Longrightarrow \hat{\sigma^2} = \frac{1}{P}\sum_{i=1}^{P}(x_i - \mu)^2$$

Question 4

Assume k Classes where observed data from each class is distributed from a Gaussian with unknown mean and variance. Maximum likelihood estimates for μ, σ^2 and p for Class i:

$$\hat{\mu_i} = \bar{x_i}$$

$$\hat{\sigma_i^2} = \frac{1}{N_i}\sum_{j=1}^{N_i}(x_j - \bar{x_i})$$

$$\hat{p_i} = \frac{N_i}{\sum_{j=1}^{k} N_j}$$

For two classes, the probability that a particular data point, **x**, belongs to class 1 is:

$$\frac{p(x|C_1)p_1}{p(x|C_1)p_1 + p(x|C_2)p_2}$$

where

$$p(x|C_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left[-\frac{1}{2}\left(\frac{x-\mu_i}{\sigma_i}\right)^2\right]$$

Applying the formulae above to the data given in the question,

$$\hat{\mu_1} = .26, \hat{\mu_1} = .8625, \hat{\sigma_1^2} = .0149, \hat{\sigma_2^2} = .0092, \hat{p_1} = .7143, \hat{p_2} = .2857$$

and the probability the .6 belongs to C_1 is therefore .6305.

Question 5

The decision boundary is given by the value(s) of x that satisfy:

$$\frac{1}{\sqrt{2\pi\sigma_1^2}} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] p_1 = \frac{1}{\sqrt{2\pi\sigma_2^2}} exp\left[-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right] p_2$$

Taking logs of both sides and rearranging terms yields:

$$\left[-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right] - \left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] + ln\left(\frac{p_2\sigma_1}{p_1\sigma_2}\right) = 0$$

This will simply to the following quadratic equation:

$$ax^2 + bx + c = 0$$

where

$$a = \sigma_1^2 - \sigma_2^2$$

$$b = 2(\sigma_2^2\mu_1 - \sigma_1^2\mu_2)$$

$$c = \sigma_2^2\mu_1^2 + \sigma_1^2\mu_2^2 - 2\sigma_1^2\sigma_2^2 ln\left(\frac{p_2\sigma_1}{p_1\sigma_2}\right)$$

The solutions are given by the quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Hence, if $b^2 - 4ac > 0$ there will be two decision boundaries (consider the case where μ_1 is close to μ_2 and σ_1^2 is large relative to σ_2^2).