Automatic Speech Recognition
Question Set 3

1. We have shown that the maximum-likelihood (ML) estimator for the mean and variance of a multi-variate Gaussian vector is given by the sample mean and variance. Here we look at the case where the random variable X is discrete. In this case, the model parameters are the probabilities in the probability mass function. Suppose there are n elements in the support set of X. Let p_i be the probability of the i-th element, $p_i = Pr(X = x_i)$.

   (a) Given N samples for X, show that the ML estimator for p_i is $\frac{n_i}{N}$, where n_i is the number of samples equal to x_i.

   (b) What is the maximum likelihood?

2. Suppose we have N *labelled* data samples, $(x_j, l_j), j = 1, \ldots, N$, where $x \in R$ and $l \in \{1, \ldots, L\}$. Assume that X_j is Gaussian given l_j.

   (a) Describe how one can obtain the ML parameters for the class probabilities p_l, the class means μ_l, and the class variances Σ_l, for $l = 1, \ldots, L$.

   (b) If the data is *not* labelled, while still assuming that there are L classes and that the class-conditional distribution is Gaussian, state how to use the EM algorithm to obtain ML estimates.