Homework Assignment #1

Due date – Oct. 18, 2011 (Tue), 5:45PM in class.

Problem 1. Thumbtack Experiment. (20 points)
Consider a thumbtack tossing experiment, where the probability of obtaining a head (H) is \( \theta \) and that of obtaining a tail (T) is \( 1 - \theta \). Suppose that in the observation data \( D = \{x[1], x[2], \ldots, x[M]\} \), there were exactly \( M[1] \) heads and \( M[0] \) tails. Find the likelihood function \( L(\theta : D) \) and prove that the maximum likelihood estimate \( \hat{\theta} \) can be written as:

\[
\hat{\theta} = \arg \max_{\theta} L(\theta : D) = \frac{M[1]}{M[1] + M[0]}.
\]

Problem 2. Thumbtack Experiment (20 points)
Now, suppose that we want to use the Bayesian estimation technique, using a uniform prior, i.e., \( P(\theta) = 1 \) for all \( \theta \in [0, 1] \). Suppose that we observed \( M[1] \) heads and \( M[0] \) tails in the past \( M = M[1] + M[0] \) experiments. Prove that the probability of observing a head in the \((M + 1)\)th experiment can be written as:

\[
\]

Problem 3. MLE for Multinomial Distribution (20 points)
Consider a multinomial model, where the random variable \( X \) can take \( K \) distinct values \( \{x_1, x_2, \ldots, x_K\} \) with probability \( P(X = x_k) = \theta_k \). Suppose that we made \( M = \sum_k M[k] \) observations of \( X \), where \( M[k] \) is the number of times \( X[m] = x_k \) for \( m = 1, \ldots, M \). Find the likelihood function \( L(\theta : D) \), where \( \theta = (\theta_1, \ldots, \theta_K) \). Prove that the maximum likelihood estimate \( \hat{\theta} \) can be written as:

\[
\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_K),
\]

where \( \hat{\theta}_k = M[k]/M \).

(Hint: maximize the log-likelihood function using a Lagrange coefficient \( \lambda \) to enforce \( \sum_k \theta_k = 1 \).)
**Problem 4. Parameter Estimation in Bayesian Network**  (40 points)

Consider the Bayesian network shown in Figure 1. All variables take binary values, with the following probabilities:

\[
\begin{align*}
P(X_1 = 1) &= \theta_1 \\
P(X_2 = 1) &= \theta_2 \\
P(X_3 = 1 | X_1 = x, X_2 = y) &= \theta_{3x,y} \\
P(X_4 = 1) &= \theta_4 \\
P(X_5 = 1 | X_3 = x) &= \theta_{5x} \\
P(X_6 = 1 | X_3 = x, X_4 = y) &= \theta_{6x,y}.
\end{align*}
\]

Let \( \theta \) be the parameter vector that contains all the above parameters. Based on the given BN, perform the following experiments (with \( M = 200 \)):

1. Randomly generate each parameter in \( \theta \) from a uniform distribution over \([0,1]\).

2. Use the above \( \theta \) to generate a dataset \( D = \{x[1], \ldots, x[M]\} \) with \( M \) observations \( x[m] = (x_1[m], \ldots, x_6[m]) \), for \( m = 1, \ldots, M \).

3. Use the first \( L \) observations \( (L = 1, \ldots, M) \) to estimate the parameter vector \( \theta \), which we denote as \( \hat{\theta}[L] \). Plot all parameters in \( \hat{\theta}[L] \) as a function of \( L \) (along with the true parameter values).

(a) Use MLE to estimate the parameters of the BN. Plot the estimated parameters.

(b) Use Bayesian approach to estimate the parameters of the BN. Plot the estimated parameters.

(c) Let us change Step-1 in the experimental procedure as follows: “Randomly generate each parameter in \( \theta \) from a uniform distribution over \([0.5, 1.0]\).” Repeat parts (a) and (b) according to the new procedure.

(d) Discuss your results.

---

Figure 1: Bayesian network for problem 4.