STA-3001/8001 : COMPUTER-INTENSIVE STATISTICS
ASSIGNMENT 3: APPROXIMATION AND OPTIMIZATION

1: EM, MCEM, DA

Let us assume we have data fram the failure times of $n = 20$ of street lamps. Each street lamp contains two light bulbs, where one is stand by and is automatically lit when the first one fails. If we assume that the failure times of individual light bulbs are independent and exponentially distributed with mean $\frac{1}{\lambda}$ then the failure times of the street lamps are gamma distributed with $\alpha = 2$:

$$f(x|\lambda) = \lambda^2 \, x \, e^{-\lambda \, x}, \qquad x > 0, \quad \lambda > 0$$

This is a censored situation. Some of the failure times are recorded, but for the other we only know that they had not failed yet at a certain time $c > 0$. Assume we have a random sample of n observations of failure times for these lamps, where $X_1, \ldots, X_m$ are observed, and $X^*_{i+1}, \ldots X^*_n$ are censored.

a) Explain why the completed likelihood is

$$L(\lambda|\boldsymbol{y}) \;\; = \;\; \left\{ \prod_{i=1}^{m} \lambda^2 \, x_i \, e^{-x_i \lambda} \right\} \left\{ \prod_{i=m+1}^{n} \lambda^2 \, z_i \, e^{-z_i \lambda} \right\}$$

where $Z_{i+1}; \ldots, Z_n$ are the (unobserved) censored failure time.
Also find the completed log-likelihood.

b) Compute

$$E(Z_i \,|\, \boldsymbol{x}, \lambda)$$

c) Derive expressions for the E-step and the M-step of the EM algorithm.

Below are the recorded failure/censoring times of the the $n = 20$ lamps (not the individual bulbs). If a time is censored (not observed) the censoring time is denoted by *.

162	145	118	200*	115	76	52	12	187	200*
200*	106	61	200*	63	163	134	200*	133	63

d) Find the MLE of λ by the EM algorithm. Illustrate the convergence by a plot.

e) Find an etimate of the standard deviation (error) of the MLE of λ.

f) Explain how you can sample from the distribution of $Z_i|\boldsymbol{x}, \lambda$.
Use this to find the MLE of λ by the Monte Carlo EM algorithm.

g) Use data augmentation to give an estimate of the distribution of $\lambda|\mathbf{x}$.
Use this to find estimates of expectation and standard deviation.

2: EM

In this problem we will use a mixed model We assume that an observation, X, can come from one of two distributions with probablity p and $1-p$, respectively. Let $f_1(\cdot)$ and $f_2(\cdot)$ be the density functions of the two distributions, then the density function of X is

$$f(x|p) = p\, f_1(x) + (1-p)\, f_2(x)$$

We assume $f_1(\cdot)$ and $f_2(\cdot)$ are known, and we want to get an estimate of p from a random sample $X_1, \ldots, X_n$.

The likelihood in this case, which we could try to optimize numerically, is:

$$L(p|\mathbf{x}) = \prod_{i=1}^{n} [p\, f_1(x_i) + (1-p)\, f_2(x_i)]$$

But instead we will try to introduce additional (latent)data, and utilize the EM-algorithm.

Let $Z_1, \ldots, Z_n$ be Bernoulli stochastic variables such that:

$$f(x|z=1) = f_1(x) \qquad \text{and} \qquad f(x|z=0) = f_2(x)$$

and the probability mass function of Z is then $g(z|p) = p^z(1-p)^{1-z},\ z = 0, 1$.

a) Show the likelihood of the completed data is

$$L(p|\mathbf{y}) = L(p|\mathbf{x}, \mathbf{z}) = \left\{ \prod_{i=1}^{n} [z_i\, f_1(x_i) + (1-z_i)\, f_2(x_i)] \right\} p^{\sum_{i=1}^{n} z_i} (1-p)^{n-\sum_{i=1}^{n} z_i}$$

Hint: $f(x, z|p) = f(x|z, p)\, g(z|p)$.

b) Show that the conditional expectation of Z given $X = x$ (and p) is

$$E(Z|x) = \frac{f_1(x)\, p}{p\, f_1(x) + (1-p)\, f_2(x)}$$

Hint: $E(Z|x) = \sum_{i=0}^{1} z\, g(z|x)$ and $g(z|x) = \frac{f(x|z)\, g(z)}{f(x)}$.

c) Use these to find an expression for $Q(p \mid p^{(t)})$. Find an iteration EM sequence for p that converges to a maximum likelihood estimator.

d) Simulate 20 independent data from $\mathrm{N}(10, 2)$ and 10 from $\mathrm{N}(14, 2)$. Use the method above to estimate p (which is here $2/3$).

e) What would the MAP estimator of p be with a beta$(3, 3)$ prior distribution and the same data you used in the previous point?

3: MCI, Riemann, Laplace

Assume X is gamma distributed with parameters $\alpha = 3.7$, $\lambda = 1$:

$$f(x|\alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \qquad x > 0, \quad \lambda > 0,\ \alpha > 0$$

Use Monte Carlo integration, Riemann-sums and Laplace approximations to find the expectation of $E[w(X)]$, where

$$w(x) = x \log(x)$$

Compare the results of MCI and Riemann with regard to the speed of convergence to the correct number.
Note: For the Laplace approximation you may assume that $\Gamma(z)$ is known. Alternatively you may use Laplace once more to prove that $\Gamma(z+1) = \int_0^\infty x^z e^{-x} dx$ may be approximated by $\sqrt{2\pi} z^{z+\frac{1}{2}} e^{-z}$.

4: Simulated Annealing

The table below shows the distance between some cities. Suppose a salesman has to visit all these cities once and return to the city he started in. Moreover, he wants to choose his route (path) so that the total distance is as short as possible. Our aim is to solve this problem by the use of Simulated Annealing. We have the following distances:

	London	Mexico City	New York	Paris	Peking	Tokyo
London	0	5558	3469	214	5074	5959
Mexico City		0	2090	5725	7753	7035
New York			0	3636	6844	6757
Paris				0	5120	6053
Peking					0	1307
Tokyo						0

Implement and solve this problem by using Simulated Annealing. Here the state is the path travelled, and the function is the total distance of the path.