18.417 Introduction to Computational Molecular Biology	
Problem Set 6	Issued: November 4, 2004
Lecturer: Ross Lippert	Due: November 30, 2004

1. 11.2 of JP.
2. 11.4 of JP
3. 11.6 of JP
4. (*) In this problem, I would like you to prove a miniature version of the fundamental theorem of expectation maximization. It should require very little probability knowledge, but will require some clever inequalities, particularly Jensen's inequality.

   Let $\bar{\theta} = (\theta_1, \ldots, \theta_n)$ be n unknowns satisfying $\sum_i \theta_i = 1$. Let $L(\bar{\theta}) = \sum_{\bar{m}} c_{\bar{m}} \bar{\theta}^{\bar{m}}$ be a polynomial in the θ_i where the sum ranges over all multi-indices $\bar{m} = (m_1, \ldots, m_n)$ such that $\bar{\theta}^{\bar{m}} = \prod \theta_i^{m_i}$, and $c_{\bar{m}} \geq 0$.

   Prove: $L(\bar{\theta}') \geq L(\bar{\theta})$ where

   $$\theta_i' = \frac{\sum_{\bar{m}} m_i c_{\bar{m}} \bar{\theta}^{\bar{m}}}{\sum_i \sum_{\bar{m}} m_i c_{\bar{m}} \bar{\theta}^{\bar{m}}}$$

   Note: I have a suspicion that this proof also might require, $\sum_i m_i = M$ for all $\bar{m}$ where $c_{\bar{m}} \neq 0$, but if that is the case, I expect you to show a counter-example which necessitates this extra condition. Of course, if the proof still doesn't work, counter-examples to the weaker version are also welcome.

5. 12.3 from JP. The description of this problem and the next problem are a bit vague. It is clear that one could create very simple randomized methods do these problems, given enough time. For example, one could randomly generate a new guess from scratch and accept it if it improved the objective function (though hardly practical). Alternatively, one could be to pick random perturbation and accept only those ones which improve the objective function (but you need to show that

set of perturbations is rich enough that the optimum can always be reached by a non-decreasing sequence of steps).

I'd like to see if you can find methods which have provably interesting distributions on their answers. Yes, this is vague.

6. 12.5 from JP. See above.