CS 194-10, Fall 2011
Assignment 0

This assignment is to be done individually; its purpose is to refresh those parts of your brain that have gone to sleep over the summer and remind you of the mathematical topics we need for the course that you might have forgotten. It also gives you a chance to practice with some of the Python packages that are useful for machine learning. The idea is that everyone should get full (well, nearly full) credit; so if you are having any problems, seek help from the GSIs and/or the newsgroup. When you are ready, `submit a0` as described **here**.

1. (25 pts) In this question you will write a simple program in python that produces samples from various distributions. A variety of methods for sampling from various distributions are available in the module **scipy.stats**. In this question you will write your own implementations for some of them, using *only* samples from the uniform distribution over the unit interval (that is, the only "source of randomness" you may use is calls to **numpy.random.uniform()**). The functions you will implement will enable sampling from four distributions: **categorical** (see also the **multinomial** for $n=1$), **univariate Gaussian**, **multivariate Gaussian** and **general mixture** distributions. You may find the linked Wikipedia pages helpful in designing your sampling algorithms.

   Write your implementations in the file sampler.py. You may add arguments to the function signatures as long as your implementation supports the interface as it is specified in the documentation of the functions. Save and submit the fully implemented sampler.py.

   Test your code as follows: Consider an equal-weighted mixture of four Gaussians in 2 dimensions, centered at $(\pm 1, \pm 1)$ and having covariance I. Estimate the probability that a sample from this distribution lies within the unit circle centered at $(0.1, 0.2)$.

   Note: Every function *can* be implemented in less than 10 lines of readable code.

2. (10 pts) Prove that a sum of two independent Poisson random variables is also a Poisson random variable.

3. (10 pts) Let X_0 and X_1 be continuous random variables. Show that if

$$p(X_0 = x_0) = \alpha_0 e^{-\frac{(x_0-\mu_0)^2}{2\sigma_0^2}}$$
$$p(X_1 = x_1 | X_0 = x_0) = \alpha_1 e^{-\frac{(x_1-x_0)^2}{2\sigma^2}}$$

   then there exists $\alpha,\ \mu_1$ and σ_1 such that

$$p(X_1 = x_1) = \alpha e^{\frac{(x_1-\mu_1)^2}{2\sigma_1^2}} .$$

   Write explicitly the values of α, μ_1 and σ_1 that satisfy the above relations.

4. (10 pts) Find the eigenvalues and eigenvectors of the following matrix:

$$A = \begin{pmatrix} 13 & 5 \\ 2 & 4 \end{pmatrix} .$$

5. (10 pts) Provide one example for each of the following cases, where A and B are 2×2 matrices:
   (a) $(A+B)^2 \neq A^2 + 2AB + B^2$,
   (b) $AB = 0,\ A \neq 0,\ B \neq 0$

6. (10 pts) Let u denote a real vector normalized to unit length, i.e., $u^T u = 1$. Show that

$$A = I - 2uu^T$$

is orthogonal, i.e., that $A^T A = I$.

7. (15 pts) A function f is convex on a given set iff for $\lambda \in [0, 1]$ and for all x, y in the set, the following holds,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)\,.$$

Moreover, a univariate function $f(x)$ is convex on a set iff its second derivative $f''(x)$ is non-negative everywhere in the set. Prove the following assertions:

(a) $f(x) = x^3$ is convex for $x \geq 0$.

(b) $f(x_1, x_2) = \max(x_1, x_2)$ is convex on $\mathbb{R}^2$.

(c) If univariate f and g are convex on a set, $f + g$ is convex on the same set.

(d) If univariate f and g are convex and nonnegative on a set, and have their minimum within the set at the same point, then fg is convex on the same set.

8. (10 pts) The *entropy* of a categorical distribution on K values is defined as

$$H(p) = -\sum_{i=1}^{K} p_i \log(p_i)\,.$$

Using the method of Lagrange multipliers, find the categorical distribution that has the highest entropy.