Hw 2

1. X is multinomially distributed if its outcome is one of K disjoint states with probabilities $p_k,\ k = 1, \dots, K;\ \sum_k^K p_k = 1$. Let $\mathbf{x} = [x_1, \dots, x_k]$ be the indicator vector of the state of an outcome, with $x_k = \begin{cases} 1, \text{if the coutcome is state } k \\ 0, \text{otherwise} \end{cases}$. Derive the ML estimate of the multinomial sample $\hat{p}_k, k = 1, \dots, K$.

2. Given the data $X = \{x_1, \dots, x_N\}$ sampled from a lognormal distribution $p(x|\theta) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left(-\frac{(\ln x - \theta)^2}{2\sigma^2}\right), x > 0$. Show that the ML estimate is given by $\hat{\theta}_{ML} = \frac{1}{N}\sum_{i=1}^N \ln x_i$.

3. In the context of regression, we have the expected value of the squared error over X, :

   $E_X\left(\left(E(y|x) - g(x)\right)^2 \middle| x\right) = (E(y|x) - E_X(g(x)))^2 + E_X[(g(x) - E_X(g(x))^2]$.

   (a) The above equation expresses the mean-squared error as a sum of a $bias^2$ and *variance*. Expand the left-hand side of the equation to get the right-hand side.

   (b) Can *bias* ever be negative? Can *variance* ever be negative? Justify your answers.

4. If we have a two-class one-dimensional problem where both classes are normally distributed with the same mean $(\mu_1 = \mu_2 = \mu)$ but different variances $(\sigma_1^2 > \sigma_2^2)$. What will be the discriminant function look like in this case?