Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2011)

Problem Set 11
Never Due
Covered on Final Exam

1. **Problem 7, page 509 in textbook**
Derive the ML estimator of the parameter of a Poisson random variable based of i.i.d. observations $X_1, \ldots, X_n$. Is the estimator unbiased and consistent?

2. Caleb builds a particle detector and uses it to measure radiation from far stars. On any given day, the number of particles Y that hit the detector is conditionally distributed according to a Poisson distribution conditioned on parameter x. The parameter x is unknown and is modeled as the value of a random variable X, exponentially distributed with parameter μ as follows.

$$f_X(x) = \begin{cases} \mu e^{-\mu x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Then, the conditional PDF of the number of particles hitting the detector is,

$$p_{Y|X}(y \mid x) = \begin{cases} \frac{e^{-x}x^y}{y!} & \text{y = 0, 1, 2, ...} \\ 0 & \text{otherwise} \end{cases}$$

   (a) Find the MAP estimate of X from the observed particle count y.

   (b) Our goal is to find the conditional expectation estimator for X from the observed particle count y.

      i. Show that the posterior probability distribution for X given Y is of the form

$$f_{X|Y}(x \mid y) = \frac{\lambda^{y+1}}{y!} x^y e^{-\lambda x}, \quad x > 0$$

      and find the parameter λ. You may find the following equality useful (it is obviously true if the equation above describes a true PDF):

$$\int_0^\infty a^{y+1} x^y e^{-ax}\, dx = y! \quad \text{for any } a > 0$$

      ii. Find the conditional expectation estimate of X from the observed particle count y. *Hint:* you might want to express $x f_{X|Y}(x \mid y)$ in terms of $f_{X|Y}(x \mid y+1)$.

   (c) Compare the two estimators you constructed in part (a) and part (b).

3. Consider a Bernoulli process X_1, X_2, X_3, $\ldots$ with unknown probability of success q. As usual, define the kth inter-arrival time T_k as

$$T_1 = Y_1, \qquad T_k = Y_k - Y_{k-1}, \quad k = 2, 3, \ldots$$

where Y_k is the time of the kth success. This problem explores estimation of q from observed inter-arrival times $\{t_1, t_2, t_3, \ldots\}$.

You may find the following integral useful: For any non-negative integers k and m,

$$\int_0^1 q^k(1-q)^m dq = \frac{k!\, m!}{(k+m+1)!}$$

Assume q is sampled from the random variable Q which is uniformly distributed over $[0, 1]$.

(a) Compute the PMF of T_1, $p_{T_1}(t_1)$

(b) Compute the least squares estimate (LSE) of Q from the first recording $T_1 = t_1$.

(c) Compute the maximum a posteriori (MAP) estimate of Q given the k recordings, $T_1 = t_1, \ldots, T_k = t_k$.

For this part only assume q is sampled from the random variable Q which is now uniformly distributed over $[0.5, 1]$

(d) Find the linear least squares estimate (LLSE) of the second inter-arrival time (T_2), from the observed first arrival time ($T_1 = t_1$).

4. A blackbody at temperature θ radiates photons of all wavelengths, described by its characteristic spectrum. This problem will have you estimate θ, which is fixed but unknown. The PMF for the number of photons K in a given wavelength range and a fixed very short time interval is given by,

$$p_K(k;\theta) = \frac{1}{Z(\theta)} e^{-k/\theta}, k = 0, 1, 2, ...$$

$Z(\theta)$ is a normalization factor for the probability distribution (the physicists call it the partition function). You are given the task of determining the temperature of the body to two significant digits by photon counting in non-overlapping time intervals of duration one second. The photon emissions in non-overlapping time intervals are statistically independent from each other.

(a) Determine the normalization factor $Z(\theta)$.

(b) Compute the expected value of the photon number measured in any 1 second time interval, $\mu_K = \mathbf{E}_\theta[K]$, and its variance, $\text{var}_\theta(K) = \sigma_K^2$.

(c) You count the number k_i of photons detected in n non-overlapping 1 second time intervals. Find the maximum likelihood estimator, $\hat{\theta}_n$, for temperature θ. Note, it might be useful to introduce the average photon number $s_n = \frac{1}{n}\sum_{i=1}^n k_i$. In order to keep the analysis simple we assume that the body is hot, i.e. $\theta \gg 1$.
You may use the approximation: $\frac{1}{e^{1/\theta}-1} \approx \theta$ for $\theta \gg 1$.

In the following questions we wish to estimate the mean of the photon count in a one second time interval using the estimator $\hat{K}$, which is given by,

$$\hat{K} = \frac{1}{n}\sum_{i=1}^n K_i.$$

(d) Find the number of samples n for which the noise to signal ratio for $\hat{K}$, (i.e.,$\frac{\sigma_{\hat{K}}}{\mu_{\hat{K}}}$), is 0.01.

(e) Find a 95% confidence interval for the mean photon count estimate for the situation in part (d). (You may use the central limit theorem.)