MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

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. 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 $\{t_1, t_2, t_3, \ldots\}$. In problem set 10, we solved the problem using Bayesian inference. Our focus here will be on classical estimation.

We assume that q is an unknown parameter in the interval $(0, 1]$. Denote the true parameter by q^*. Denote by $\widehat{Q}_k$ the maximum likelihood estimate (MLE) of q given k recordings, $T_1 = t_1, \ldots, T_k = t_k$.

(a) Compute $\widehat{Q}_k$. Is this different from the MAP estimate you found in problem set 10?

(b) Show that for all $\epsilon > 0$

$$\lim_{k\to\infty} \mathbf{P}\left(\left|\frac{1}{\widehat{Q}_k} - \frac{1}{q^*}\right| > \epsilon\right) = 0$$

(c) Assume $q^* \geq 0.5$. Give a lower bound on k such that

$$\mathbf{P}\left(\left|\frac{1}{\widehat{Q}_k} - \frac{1}{q^*}\right| \leq 0.1\right) \geq 0.95$$

4. A body at temperature θ radiates photons at a given wavelength. 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 time interval of one second is given by,

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

$Z(\theta)$ 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, $\mathrm{var}_\theta(K) = \sigma_K^2$.

(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: $\frac{1}{e^{\frac{1}{\theta}} - 1} \approx \theta$ 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 the 95% confidence interval for the mean photon count estimate for the situation in part (d). (You may use the central limit theorem.)

5. The RandomView window factory produces window panes. After manufacturing, 1000 panes were loaded onto a truck. The weight W_i of the i-th pane (in pounds) on the truck is modeled as a random variable, with the assumption that the W_i's are independent and identically distributed.

(a) Assume that the measured weight of the load on the truck was 2340 pounds, and that $\mathrm{var}(W_i) \leq 4$. Find an approximate 95 percent confidence interval for $\mu = \mathbf{E}[W_i]$, using the Central Limit Theorem.

(b) Now assume instead that the random variables W_i are i.i.d., with an exponential distribution with parameter $\theta > 0$, i.e., a distribution with PDF

$$f_W(w; \theta) = \theta e^{-\theta w}.$$

What is the maximum likelihood estimate of θ, given that the truckload has weight 2340 pounds?

6. Given the five data pairs (x_i, y_i) in the table below,

x	0.8	2.5	5	7.3	9.1
y	-2.3	20.9	103.5	215.8	334

we want to construct a model relating x and y. We consider a linear model

$$Y_i = \theta_0 + \theta_1 x_i + W_i, \qquad i = 1, \cdots, 5,$$

and a quadratic model

$$Y_i = \beta_0 + \beta_1 x_i^2 + V_i, \qquad i = 1, \cdots, 5.$$

where W_i and V_i represent additive noise terms, modeled by independent normal random variables with mean zero and variance σ_1^2 and σ_2^2, respectively.

(a) Find the ML estimates of the linear model parameters.

(b) Find the ML estimates of the quadratic model parameters.

Note: You may use the regression formulas and the connection with ML described in pages 478-479 of the text. However, the regression material is outside the scope of the final.