MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 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 $X_1, \dots, X_n$. 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, $\dots$ with unknown probability of success q. Define the kth inter-arrival time T_k as

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

   where Y_k is the time of the kth success. This problem explores estimation of q from observed inter-arrival times $\{t_1, t_2, t_3, \dots\}$. 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. Let $X_1, \ldots, X_n$ be i.i.d. samples of a Gaussian random variable with an unknown common mean θ, and an unknown variance σ^2. Suppose we have sample values $X_1 = x_1, \ldots, X_n = x_n$. The mean estimator is

$$\hat{\Theta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

(a) Find the mean and variance of $\hat{\Theta}_n$. Is $\hat{\Theta}_n$ Gaussian?

(b) A common approximation (which is not exactly correct, but is close for large values of n) is that the unbiased variance estimator $\hat{S}_n^2$ is exactly equal to σ^2. Using this approximation, find the probability distribution for the random variable

$$T_n = \frac{\sqrt{n}(\hat{\Theta}_n - \theta)}{\hat{S}_n},$$

where $\hat{S}_n = +\sqrt{\hat{S}_n^2}$.
Write the event that θ lies in the confidence interval

$$[\hat{\Theta}_n - z\frac{\hat{S}_n}{\sqrt{n}}, \hat{\Theta}_n + z\frac{\hat{S}_n}{\sqrt{n}}]$$

in terms of a range of possible values for T_n. Using the approximation above, find the 95 % confidence interval for Θ , i.e., find the value of z for which

$$\mathbf{P}_\theta\left(\hat{\Theta}_n - z\frac{\hat{S}_n}{\sqrt{n}} < \theta < \hat{\Theta}_n + z\frac{\hat{S}_n}{\sqrt{n}}\right) = 0.95.$$

Find the confidence interval for $n = 4$ and $n = 16$ in terms of $\hat{S}_n$ and $\hat{\Theta}_n$.

(c) When the X_i's are iid normal, the random variable T_n is called the "t-distribution with n-1 degrees of freedom," and it has a known probability distribution. The distribution is symmetric about the origin and broadly resembles the standard normal density, $N(\mu = 0, \sigma = 1)$, but with "fatter tails." Find values of z that give a more accurate 95 % confidence interval for θ for $n = 4$ and $n = 16$. Give the confidence intervals for both values of n in terms of $\hat{S}_n$ and $\hat{\Theta}_n$.

(d) Compare your answers to parts (b) and (c). Which method gives a wider confidence interval? How does this behavior depend on n?

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 $\text{var}(W_i) \leq 4$. 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

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

   and a quadratic model

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

   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.