Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Fall 2008)

Recitation 22
December 2, 2008

1. Sarah is working in a photon lab aboard a ship. She measures the number of photons emitted by a source every one second interval. The number of photons emitted at each interval i is random and denoted by the random variable X_i. These observations are assumed to be i.i.d. with a finite mean, λ, and a finite variance v. She calculates the sample mean,

$$M_n = \frac{1}{n}\sum_{i}^{n} X_i,$$

and the variance estimator

$$\hat{S}_n^2 = c\sum_{i}^{n}(X_i - M_n)^2.$$

   (a) What value of c will make the variance estimator, $\hat{S}_n^2$, unbiased?

   (b) She finds the sample mean $m_n = 7.1$. Using the estimator in part (a) she finds $\hat{s}_n^2 = 7.3$. A huge wave envelopes the ship sweeping away all the data. Luckily, Sarah had scribbled the values of the sample mean and variance in her notebook. Her boss wants her to model the distribution of X_i as a Poisson distribution and asks her to find the ML (Maximum Likelihood) estimate of λ. Sarah panics since she lost her data in the cold waters of Alaska. She knows from the 6.041 class she took that the variance of the Poisson and its mean are the same, but the two numbers she has are different. Is the data she has in her note book enough to calculate the the ML estimate, $\hat{\Lambda}_n$ of the parameter λ for the Poisson model? If so what is it?

2. 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 $\hat{\Theta}_n$.

3. (See Problem 9. Pg. 509) We are given i.i.d observations $X_1, \cdots, X_n$ that are uniformly distributed over the interval $[\theta, \theta + 1]$.
   (a) Find a ML (Maximum Likelihood) estimate for θ? If there's more than one, find all of them.
   (b) Consider the largest and smallest of the ML estimators. To what values do these two converge?
   (c) Is the largest one consistent? Is the smallest one consistent? What about the others ?