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

Recitation 25 Solutions
December 8, 2011

1. (a) The PDF of X_i is

$$f_{X_i}(x_i) = \begin{cases} 1, & \text{if } \theta \le x_i \le \theta + 1, \\ 0, & \text{otherwise.} \end{cases}$$

The likelihood function is

$$\begin{aligned} f_X(x_1, \ldots, x_n; \theta) &= f_{X_1}(x_1; \theta) \cdots f_{X_n}(x_n; \theta) \\ &= \begin{cases} 1 & \text{if } \theta \le \min_{i=1,\ldots,n} x_i \le \max_{i=1,\ldots,n} x_i \le \theta + 1 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

Any value in the feasible interval

$$[\max_{i=1,\ldots,n} X_i - 1, \min_{i=1,\ldots,n} X_i]$$

maximizes the likelihood function and is therefore a ML estimator.

(b) $\min_{i=1,\ldots,n} X_i$ converges in probability to θ, while $\max_{i=1,\ldots,n} X_i - 1$ converges in probability to θ too. We will show the first result. Let $L_n = \min_{i=1,\ldots,n} X_i$.

We intuitively expect that the series converges to θ. Note that L_n cannot increase as n increases, therefore the series is monotonic. Moreover, it is lower bounded by θ, therefore a limit exists. We therefore make a guess that it converges to θ.

Indeed for $\epsilon > 0$, we have using the independence of the X_i's,

$$\begin{aligned} \mathbf{P}(|L_n - \theta| \ge \epsilon) &= \mathbf{P}(X_1 \ge \theta + \epsilon, \cdots, X_n \ge \theta + \epsilon) \\ &= \mathbf{P}(X_1 \ge \theta + \epsilon) \cdots \mathbf{P}(X_n \ge \theta + \epsilon) \\ &= (1 - \epsilon)^n. \end{aligned}$$

Since

$$\lim_{n \to \infty} \mathbf{P}(|L_n - \theta| \ge \epsilon) = \lim_{n \to \infty} (1 - \epsilon)^n = 0$$

for every $\epsilon > 0$ it follows that L_n converges to θ in probability. A similar argument can be applied to show $\max_{i=1,\ldots,n} X_i$ converges in probability to $\theta + 1$.

(c) Any choice of estimator within the above interval is consistent. The reason is that $\min_{i=1,\ldots,n} X_i$ converges in probability to θ, while $\max_{i=1,\ldots,n} X_i$ converges in probability to $\theta + 1$. Thus, both endpoints of the above interval converge to θ.

2. The parameter vector here is $\theta = (\mu, v)$. The corresponding likelihood function is

$$f_X(x; \mu, v) = \prod_{i=1}^{n} f_{X_i}(x_i; \mu, v) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi v}} e^{-(x_i - \mu)^2/2v}.$$

After some calculation it can be written as

$$f_X(x; \mu, v) = \frac{1}{(2\pi v)^{n/2}} \cdot \exp\left\{-\frac{ns_n^2}{2v}\right\} \cdot \exp\left\{-\frac{n(m_n - \mu)^2}{2v}\right\},$$

where m_n is the realized value of the random variable

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

and s_n^2 is the realized value of the random variable

$$\bar{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_n)^2.$$

The log-likelihood function is

$$\log f_X(x;\mu,v) = -\frac{n}{2}\cdot\log\,(2\pi) - \frac{n}{2}\cdot\log\,v - \frac{ns_n^2}{2v} - \frac{n(m_n-\mu)^2}{2v}.$$

Setting to zero the derivatives of this function with respect to μ and v, we obtain the estimate and estimator, respectively,

$$\hat{\theta}_n = (m_n, s_n^2), \qquad \hat{\Theta}_n = (M_n, \bar{S}_n^2).$$

Note that M_n is the sample mean, while $\bar{S}_n^2$ may be viewed as a "sample variance." As will be shown shortly, $\mathbf{E}_\theta[\bar{S}_n^2]$ converges to v as n increases, so that $\bar{S}_n^2$ is asymptotically unbiased. Using also the weak law of large numbers, it can be shown that M_n and $\bar{S}_n^2$ are consistent estimators of μ and v, respectively.

3. (a) If we use the variance estimate

$$\begin{aligned}\hat{S}_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\Theta}_n)^2\\ &= \frac{1}{1199}\left(684\cdot\left(1-\frac{684}{1200}\right)^2 + (1200-684)\cdot\left(0-\frac{684}{1200}\right)^2\right)\\ &\approx 0.245,\end{aligned}$$

and treat $\hat{\Theta}_n$ as a normal random variable with mean θ and variance 0.245, we obtain the 95% confidence interval

$$\begin{aligned}\left[\hat{\Theta}_n - 1.96\frac{\hat{S}_n}{\sqrt{n}},\ \hat{\Theta}_n + 1.96\frac{\hat{S}_n}{\sqrt{n}}\right] &= \left[0.57 - \frac{1.96\cdot\sqrt{0.245}}{\sqrt{1200}},\ 0.57 + \frac{1.96\cdot\sqrt{0.245}}{\sqrt{1200}}\right]\\ &= [0.542,\ 0.598].\end{aligned}$$

(b) The variance estimate

$$\hat{\Theta}_n(1-\hat{\Theta}_n) = \frac{684}{1200}\cdot\left(1-\frac{684}{1200}\right) = 0.245$$

is the same as the previous one (up to three decimal place accuracy), and the resulting 95% confidence interval

$$\left[\hat{\Theta}_n - 1.96\frac{\sqrt{\hat{\Theta}_n(1-\hat{\Theta}_n)}}{\sqrt{n}},\ \hat{\Theta}_n + 1.96\frac{\sqrt{\hat{\Theta}_n(1-\hat{\Theta}_n)}}{\sqrt{n}}\right]$$

is again [0.542, 0.598].

(c) The conservative upper bound of 1/4 for the variance results in the confidence interval

$$\begin{aligned}\left[\hat{\Theta}_n - 1.96\frac{1/2}{\sqrt{n}},\ \hat{\Theta}_n + 1.96\frac{1/2}{\sqrt{n}}\right] &= \left[0.57 - \frac{1.96\cdot(1/2)}{\sqrt{1200}},\ 0.57 + \frac{1.96\cdot(1/2)}{\sqrt{1200}}\right]\\ &= [0.542,\ 0.599],\end{aligned}$$

which is only slightly wider, but practically the same as before.