Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

Tutorial 11 Solutions

1. (a) The LMS estimator is

$$g(x) = \mathbf{E}[Y|X] = \begin{cases} \frac{1}{2}X & 0 \le X < 1 \\ X - \frac{1}{2} & 1 \le X \le 2 \\ \text{Undefined} & \text{Otherwise} \end{cases}$$

(b) If $x \in [0,1]$, the conditional PDF of Y is uniform over the interval $[0,x]$, and

$$\mathbf{E}\left[(Y - g(X))^2 \mid X = x\right] = \frac{x^2}{12}.$$

Similarly, if $x \in [1,2]$, the conditional PDF of Y is uniform over $[1-x, x]$, and

$$\mathbf{E}\left[(Y - g(X))^2 \mid X = x\right] = 1/12.$$

(c) The expectations $\mathbf{E}\left[(Y - g(X))^2\right]$ and $\mathbf{E}\left[\text{var}(Y|X)\right]$ are equal because by the law of iterated expectations,

$$\mathbf{E}\left[\left(Y - g(X)\right)^2\right] = \mathbf{E}\left[\mathbf{E}\left[(Y - g(X))^2 \mid X\right]\right] = \mathbf{E}[\text{var}(Y \mid X)].$$

Recall from part (b) that

$$\text{var}(Y|X = x) = \begin{cases} \frac{x^2}{12} & 0 \le x < 1, \\ \frac{1}{12} & 1 \le x \le 2. \end{cases}$$

It follows that

$$\mathbf{E}[\text{var}(Y \mid X)] = \int_x \text{var}(Y \mid X = x) f_X(x) dx = \int_0^1 \frac{x^2}{12} . \frac{2}{3} x dx + \int_1^2 \frac{1}{12} . \frac{2}{3} dx = \frac{5}{72}.$$

Note that

$$f_X(x) = \begin{cases} 2x/3 & 0 \le x < 1, \\ 2/3 & 1 \le x \le 2. \end{cases}$$

(d) The linear LMS estimator is

$$L(X) = \mathbf{E}[Y] + \frac{\text{cov}(X,Y)}{\text{var}(X)}[X - \mathbf{E}[X]].$$

In order to calculate $\text{var}(X)$ we first calculate $\mathbf{E}[X^2]$ and $\mathbf{E}[X]^2$.

$$\begin{aligned} \mathbf{E}[X^2] &= \int_0^2 x^3 \frac{2}{3} dx + \int_1^2 x^2 \frac{2}{3} dx, \\ &= \frac{31}{18}, \\ \mathbf{E}[X] &= \int_0^2 x^2 \frac{2}{3} dx + \int_1^2 x \frac{2}{3} dx, \\ &= \frac{11}{9} \end{aligned}$$

$\text{var}(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{37}{162}$.

$$\mathbf{E}[Y] = \int_0^1 \int_0^x \frac{2}{3} y \; dydx + \int_1^2 \int_{x-1}^x \frac{2}{3} y \; dydx = \frac{1}{9} + \frac{2}{3} = \frac{7}{9}.$$

To determine $\text{cov}(X, Y)$ we need to evaluate $\mathbf{E}[XY]$.

$$\begin{aligned} \mathbf{E}[YX] &= \int_x \int_y xy f_{X,Y}(x, y) dydx \\ &= \int_0^1 \int_0^x yx \frac{2}{3} dydx + \int_1^2 \int_{x-1}^x yx \frac{2}{3} dydx \\ &= \frac{41}{36} \end{aligned}$$

Therefore $\text{cov}(X, Y) = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = \frac{61}{324}$. Therefore,

$$L(X) = \frac{7}{9} + \frac{61}{74}[X - \frac{11}{9}].$$

(e) The LMS estimator is the one that minimizes mean squared error (among all estimators of Y based on X). The linear LMS estimator, therefore, cannot perform better than the LMS estimator, i.e., we expect $\mathbf{E}[(Y - L(X))^2] \geq \mathbf{E}[(Y - g(X))^2]$. In fact,

$$\begin{aligned} \mathbf{E}[(Y - L(X))^2] &= \sigma_Y^2 (1 - \rho^2), \\ &= \sigma_Y^2 (1 - \frac{\text{cov}(X, Y)^2}{\sigma_X^2 \sigma_Y^2}), \\ &= \frac{37}{162} \left(1 - (\frac{61}{74})^2\right), \\ &= 0.073 \geq \frac{5}{72} \end{aligned}$$

(f) For a single observation x of X, the MAP estimate is not unique since all possible values of Y for this x are equally likely. Therefore, the MAP estimator does not give meaningful results.

2. (a) X is a binomial random variable with parameters $n = 3$ and given the probability p that a single bit is flipped in a transmission over the noisy channel:

$$p_X(k; p) = \begin{cases} \binom{3}{k} p^k (1-p)^{3-k}, & k = 0, 1, 2, 3 \\ 0 & \text{o.w.} \end{cases}$$

(b) To derive the ML estimator for p based on $X_1, \ldots, X_n$, the numbers of bits flipped in the first n three-bit messages, we need to find the value of p that maximizes the likelihood function:

$$\hat{p}_n = \arg\max_p \; p_{X_1, \ldots, X_n}(k_1, k_2, \ldots, k_n; p)$$

Since the X_i's are independent, the likelihood function simplifies to:

$$p_{X_1,\ldots,X_n}(k_1,k_2,\ldots,k_n;p) = \Pi_{i=1}^n p_{X_i}(k_i;p) = \Pi_{i=1}^n \binom{3}{k_i} p^{k_i}(1-p)^{3-k_i}$$

The log-likelihood function is given by

$$\log(p_{X_1,\ldots,X_n}(k_1,k_2,\ldots,k_n;p)) = \sum_{i=1}^n \left(k_i \log(p) + (3-k_i)\log(1-p) + \log\binom{3}{k_i}\right)$$

We then maximize the log-likelihood function with respect to p:

$$\begin{aligned}
\frac{1}{p}\left(\sum_{i=1}^n k_i\right) - \frac{1}{1-p}\left(\sum_{i=1}^n (3-k_i)\right) &= 0 \\
\left(3n - \sum_{i=1}^n k_i\right) p &= \left(\sum_{i=1}^n k_i\right)(1-p) \\
\hat{p}_n = \frac{1}{3n}\sum_{i=1}^n k_i &
\end{aligned}$$

This yields the ML estimator:

$$\hat{P}_n = \frac{1}{3n}\sum_{i=1}^n X_i$$

(c) The estimator is unbiased since:

$$\begin{aligned}
\mathbf{E}_p[\hat{P}_n] &= \frac{1}{3n}\sum_{i=1}^n \mathbf{E}_p[X_i] \\
&= \frac{1}{3n}\sum_{i=1}^n 3p \\
&= p
\end{aligned}$$

(d) By the weak law of large numbers, $\frac{1}{n}\sum_{i=1}^n X_i$ converges in probability to $\mathbf{E}_p[X_i] = 3p$, and therefore $\hat{P}_n = \frac{1}{3n}\sum_{i=1}^n X_i$ converges in probability to p. Thus $\hat{P}_n$ is consistent.

(e) Sending 3 bit messages instead of 1 bit messages does not affect the ML estimate of p. To see this, let Y_i be a Bernoulli RV which takes the value 1 if the ith bit is flipped (with probability p), and let $m = 3n$ be the total number of bits sent over the channel. The ML estimate of p is then

$$\hat{P}_n = \frac{1}{3n}\sum_{i=1}^n X_i = \frac{1}{m}\sum_{i=1}^m Y_i.$$

Using the central limit theorem, $\hat{P}_n$ is approximately a normal RV for large n. An approximate 95% confidence interval for p is then,

$$\left[\hat{P}_n - 1.96\sqrt{\frac{v}{m}}, \hat{P}_n + 1.96\sqrt{\frac{v}{m}}\right]$$

where v is the variance of Y_i.

As suggested by the question, we estimate the unknown variance v by the convervative upper bound of 1/4. We are also give that $n = 100$ and the number of bits flipped is 20, yielding $\hat{P}_n = \frac{2}{30}$. Thus, an approximate 95% confidence interval is $[0.01, 0.123]$.

(f) Other estimates for the variance are the sample variance and the estimate $\hat{P}_n(1-\hat{P}_n)$. They potentially result in narrower confidence intervals than the conservative variance estimate used in part (e).