Department of EECS - University of California at Berkeley

EECS 126 - Probability and Random Processes - Fall 2008

Final: 12/20/2008

SOLUTIONS

---

1. LLSE (5%)

Let X, Y be i.i.d. and uniformly distributed in $[-1, 1]$. Find $L[X|(X+Y)^2]$.

Answer. *Let* $Z = (X+Y)^2$. *We know that*

$$L[X|Z] = E(X) + \frac{cov(X,Z)}{var(Z)}(Z - E(X)).$$

Now,

$$cov(X,Z) = E(XZ) - E(X)E(Z) = E(XZ) = E(X(X^2 + 2XY + Y^2)) = 0.$$

Hence,

$$L[X|(X+Y)^2] = 0.$$

2. LLSE (10%)

Let X, Y be i.i.d. $N(0,1)$. Find $L[X|(X+Y)^3]$.
(*Hint:* Let V be a $N(0,1)$ random variable. Then $E(V^4)=3, E(V^6)=15, E(V^8)=105$.)

Answer. *Let* $Z=(X+Y)^3$*. Then*

$$L[X|Z]=E(X)+\frac{cov(X,Z)}{var(Z)}(Z-E(Z)).$$

Now,

$$\begin{aligned}
&E(X)=E(Z)=0\\
&cov(X,Z)=E(XZ)-E(X)E(Z)=E(XZ)=E(X(X^3+3X^2Y+3XY^2+Y^3))=3+0+3+0=6\\
&var(Z)=E(Z^2)=E((\sqrt{2}V)^6)=(\sqrt{2})^6E(V^6)=8\times 15=120.
\end{aligned}$$

Hence,

$$L[X|Z]=\frac{6}{120}Z=\frac{1}{20}Z.$$

3. MMSE (5%) Let X, Y be i.i.d. $N(0,1)$. Find $E[X|(X+Y)^3]$.

Answer. *Let* $Z = (X+Y)^3$. *Given* Z, *one finds* $X + Y = Z^{1/3}$. *By symmetry,* $E[X|X+Y] = (X+Y)/2$. *Hence,*

$$E[X|Z] = \frac{1}{2} Z^{1/3}.$$

4. MMSE (10%)

Let X, Y, Z be i.i.d. $N(0,1)$. Calculate $E[X|X+3Y, X+5Z]$.

Answer. *Let* $V_1 = X+3Y, V_2 = X+5Z$, *and* $\mathbf{V} = [V_1, V_2]^T$. *Then, since* X, V_1, V_2 *are zero-mean,*

$$E[X|\mathbf{V}] = cov(X, \mathbf{V})\, cov(\mathbf{V})^{-1}\mathbf{V}.$$

Now,

$$cov(X, \mathbf{V}) = E(X\mathbf{V}^T) = E(X[V_1, V_2]) = E([X^2 + 3XY, X^2 + 5XZ]) = [1, 1].$$

Also,

$$\begin{aligned} var(V_1) &= var(X) + var(3Y) = 1 + 9 = 10 \\ var(V_2) &= var(X) + var(5Z) = 1 + 25 = 26 \\ cov(V_1, V_2) &= E((X+3Y)(X+5Z)) = E(X^2 + 3XY + 5XZ + 15YZ) = 1, \end{aligned}$$

so that

$$\Sigma_{\mathbf{V}} := cov(\mathbf{V}) = \begin{bmatrix} 10 & 1 \\ 1 & 26 \end{bmatrix}.$$

Hence,

$$E[X|\mathbf{V}] = [1,1] \begin{bmatrix} 10 & 1 \\ 1 & 26 \end{bmatrix}^{-1} \mathbf{V} = [1,1]\frac{1}{259} \begin{bmatrix} 26 & -1 \\ -1 & 10 \end{bmatrix} \mathbf{V} = \frac{1}{259}[25, 9]\mathbf{V}.$$

5. MLE (10%)

For $i \in \{1, 2, 3, 4\}$, if $X = i$, then $\mathbf{Y} = N(\mu_i, \sigma^2 I)$ where $\mu_1 = (-1, -1), \mu_2 = (1, -1), \mu_3 = (1, 1), \mu_4 = (-1, 1)$. Find the Maximum Likelihood Estimate of X given $\mathbf{Y} = \mathbf{y} \in \Re^2$.

Answer. *We saw in class that* $MLE[X|\mathbf{Y} = \mathbf{y}] = argmin_i ||\mathbf{y} - \mu_i||$ *where the norm is the standard Euclidean norm. Accordingly,*

$$MLE[X|\mathbf{Y} = \mathbf{y}] = \begin{cases} 1, & \text{if } y_1 < 0 \text{ and } y_2 < 0 \\ 2, & \text{if } y_1 \geq 0 \text{ and } y_2 < 0 \\ 3, & \text{if } y_1 \geq 0 \text{ and } y_2 \geq 0 \\ 4, & \text{if } y_1 < 0 \text{ and } y_2 \geq 0. \end{cases}$$

6. HT (10%) If $X = 1$, then Y is $N(1, 1)$. If $X = 0$, then Y is exponentially distributed with mean 1. Find the function $\hat{X}$ of Y that maximizes $P[\hat{X} = 1|X = 1]$ subject to $P[\hat{X} = 1|X = 0] \leq 0.1$.
(*Hint:* $\cosh(x) = (e^x + e^{-x})/2, \sinh(x) = (e^x - e^{-x})/2$. You may give the answer in terms of a basic function without evaluating the numerical value.)

Answer. *The likelihood ratio is*

$$L(y) = \frac{\phi_1(y)}{\phi_0(y)} = e^y \frac{1}{\sqrt{2\pi}} e^{-(y-1)^2/2}.$$

Thus,

$$L(y) = Ae^{-(y-2)^2/2}$$

with $A = e^{-3/2}/\sqrt{2\pi}$.

Accordingly,

$$\hat{X} = 1\{L(Y) > \lambda\} = 1\{|y - 2| < \alpha\}$$

where α *is such that*

$$P[|Y - 2| < \alpha|X = 0] = 0.1.$$

Now,

$$\begin{aligned} P[|Y - 2| < \alpha|X = 0] &= P[Y > 2 - \alpha|X = 0] - P[Y > 2 + \alpha|X = 0] \\ &= e^{-2-\alpha} - e^{-2-\alpha} = e^{-2}[e^{\alpha} - e^{-\alpha}] = 2e^{-2}\sinh(\alpha). \end{aligned}$$

Hence,

$$\alpha = \sinh^{-1}(0.05e^2) \approx 0.36.$$

(The last value is found using a calculator.)

Finally,

$$\hat{X} = 1\{1.64 < Y < 2.36\}.$$

7. CLT (10%) We want to estimate the probability p that a coin flip yields 'Tail.' We flip the coin $n \gg 1$ times and we let X_n denote the fraction of 'Tails.' Assume that we know that p is not far from 0.5 but we want a more precise estimate. Using the Central Limit Theorem, find a value of α so that

$$P(X_n - \alpha < p < X_n + \alpha) \approx 0.95.$$

(*Hint:* If $V = N(0,1)$, then $P(|V| < 1.96) \approx 0.95$).)

Answer. *We know, from the CLT, that*

$$(X_n - p)\sqrt{n} \approx N(0, p(1-p)) \approx N(0, \frac{1}{4}) = \frac{1}{2}V.$$

Thus,

$$X_n - p \approx \frac{1}{2\sqrt{n}}V.$$

Consequently,

$$P(|X_n - p| < \alpha) \approx P(\frac{1}{2\sqrt{n}}|V| < \alpha) = P(|V| < 2\sqrt{n}\alpha).$$

Accordingly, we need

$$2\sqrt{n}\alpha \approx 1.96, \ \textit{i.e.}, \alpha \approx \frac{1}{\sqrt{n}}.$$

Hence, for $n \gg 1$,

$$P(X_n - \frac{1}{\sqrt{n}} < p < X_n + \frac{1}{\sqrt{n}}) \approx 0.95.$$

8. Markov Chain (10%)

Consider a Markov X_n on $\{1, 2, 3, 4\}$ such that

$$P(1,2) = P(2,3) = P(3,4) = P(4,1) = p \text{ and } P(2,1) = P(3,2) = P(4,3) = P(1,4) = 1-p$$

where $p \in (0,1)$ is given.

a) Is the Markov chain irreducible? Aperiodic?

b) Find the invariant distribution π.

c) Calculate the average time it takes the Markov chain to go from 1 to 3.

d) What is the probability that the Markov chain goes from 1 to 3 without first visiting 4?

Answer. *a) The Markov chain is irreducible and periodic with period 2.*

b) By symmetry, it must be that $\pi = [1/4, 1/4, 1/4.1/4]$.

c) Let $\beta(i)$ *be the mean time from* $X_0 = i$ *to* 3. *We find*

$$\begin{aligned}\beta(1) &= 1 + p\beta(2) + (1-p)\beta(4)\\ \beta(2) &= 1 + p \times 0 + (1-p)\beta(1)\\ \beta(4) &= 1 + p\beta(1) + (1-p) \times 0.\end{aligned}$$

Substituting the values for $\beta(2)$ *and* $\beta(4)$ *found the the last two equations into the first one, we find*

$$\beta(1) = 1 + p[1 + (1-p)\beta(1)] + (1-p)[1 + p\beta(1)].$$

Regrouping terms, we find

$$\beta(1) = 2[1 - 2p(1-p)]^{-1},$$

which is the desired answer.

d) Let $\alpha(i)$ *denote the probability that the Markov chain started in* $X_0 = i$ *visits state 3 before visiting state 4. We find*

$$\begin{aligned}\alpha(1) &= p\alpha(2) + (1-p) \times 0\\ \alpha(2) &= p + (1-p)\alpha(1)\end{aligned}$$

Substituting the expression for $\alpha(2)$ *given by the last equation into the first one, we find*

$$\alpha(1) = p[p + (1-p)\alpha(1)] = p^2 + p(1-p)\alpha(1).$$

That is,

$$\alpha(1) = p^2[1 - p(1-p)]^{-1}.$$

9. Conditional Expectation (10%) Random variables X and Y are jointly distributed such that X is uniformly distributed on $[0, 2]$ and, conditioned on X, Y is then uniform on $[0, X]$. What is $E[X|Y]$?

(*Note:* Be sure you solve for $E[X|Y]$, not $E[Y|X]$.)

Answer. *Start from the definition of conditional expectation:*

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx$$
$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

We will solve for marginal distribution $f_Y(y)$ by way of the joint distribution $f_{XY}(x,y)$:

$$\begin{aligned} f_{XY}(x,y) =& f_{Y|X}(y|x) f_X(x) \\ =& \frac{1}{x}\frac{1}{2}, \quad 0 \le y \le x, \quad 0 \le x \le 2 \end{aligned}$$

Note that this triangle can also be described by $0 \le y \le 2, y \le x \le 2$.

$$\begin{aligned} f_Y(y) =& \int_{-\infty}^{\infty} f_{XY}(x,y)dx \\ =& \int_y^2 \frac{1}{2x}dx \\ =& \frac{1}{2}\ln x\Big|_{x=y}^{2} \\ =& \frac{\ln 2 - \ln y}{2}, \quad 0 \le y \le 2 \\ f_{X|Y}(x|y) =& \frac{1}{x}\frac{1}{\ln 2 - \ln y}, \quad 0 \le y \le 2, \quad y \le x \le 2 \\ E[X|Y=y] =& \int_y^2 x \frac{1}{x}\frac{1}{\ln 2 - \ln y}dx \\ =& \frac{2-y}{\ln 2 - \ln y} \end{aligned}$$

10. MAP (10%) You are told that random variables X and Y are distributed as follows:

$$f_X(x) = \begin{cases} ax^2, & 0 \le x \le 3 \\ 0, & \text{otherwise} \end{cases}$$
$$f_{Y|X}(y|x) = xe^{-xy}, \quad 0 \le x \le 3, \quad 0 < y$$

where a is some constant. Find a and the MAP estimate for X based on observation Y.

Answer. *Since $f_X(x)$ is a probability distribution, its integral should equal 1.*

$$\begin{aligned} \int_{-\infty}^{\infty} f_X(x)dx &= \int_0^3 ax^2 dx \\ &= a\frac{x^3}{3}\Big|_{x=0}^{3} \\ &= 9a = 1 \qquad \Rightarrow a = \frac{1}{9} \end{aligned}$$

The MAP estimate is:

$$\begin{aligned} \hat{x} &= \operatorname*{argmax}_x f_{X|Y}(x|y) \\ &= \operatorname*{argmax}_x \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)} \\ &= \operatorname*{argmax}_x f_{Y|X}(y|x) f_X(x) \\ &= \operatorname*{argmax}_{0 \le x \le 3} xe^{-xy}\frac{x^2}{9} \\ &= \operatorname*{argmax}_{0 \le x \le 3} e^{-xy}\frac{x^3}{9} \end{aligned}$$

We maximize this by differentiating with respect to x and setting equal to 0:

$$\begin{aligned} \frac{\partial}{\partial x} e^{-xy}\frac{x^3}{9} &= -ye^{-xy}\frac{x^3}{9} + e^{-xy}\frac{x^2}{3} \\ &= (-xy+3)e^{-xy}\frac{x^2}{9} = 0 \end{aligned}$$

For $y > 0$, this has solutions $x = \frac{3}{y}, +\infty$, and 0. However, only for the case $x = \frac{3}{y}$ is $f_{X|Y}(x|y)$ positive (the other two are 0). Therefore, $\hat{x} = \frac{3}{y}$ if $0 \le \frac{3}{y} \le 3 \rightarrow y \ge 1$. For $0 < y < 1$, $\frac{3}{y} > 3$ so we know $f_{X|Y}(x|y)$ is increasing over $0 \le x \le 3$, and we choose $\hat{x} = 3$. Therefore the MAP estimator is:

$$\hat{x} = \begin{cases} 3, & 0 < y < 1 \\ \frac{3}{y}, & 1 \le y \end{cases}$$

11. (10%). Note: This problem is quite difficult and its solution is not very short.

When $X = 1$, a machine produces electric light bulbs with i.i.d. exponential lifetimes with a mean less than one year. When $X = 0$, their mean lifetime is equal to one year. You measure the lifetimes $Y_1, \ldots, Y_n$ of $n \gg 1$ light bulbs. You decide $\hat{X} \in \{0, 1\}$ based on the observed lifetimes.

1) Construct $\hat{X}$ to maximize $P[\hat{X} = 1|X = 1]$ subject to $P[\hat{X} = 1|X = 0] \leq 5\%$.

2) Assume that when $X = 1$ the mean life time is 0.9 year. How large should n be so that $P[\hat{X} = 1|X = 1] \geq 95\%$?

(*Hints:* Use the CLT. Tables show that $P(N(0,1) > 1.64) = 5\%, P(N(0,1) > 1.96) = 2.5\%, P(N(0,1) > 2.33) = 1\%, P(N(0,1) > 2.53) = 0.5\%$. Also, the variance of an exponentially distributed random variable with rate μ is $1/\mu^2$.)

Answer. *1) For any given mean lifetime less than one year when $X = 1$, the likelihood ratio is decreasing in $Y_1 + \cdots + Y_n$. Consequently,*

$$\hat{X} = 1\{\frac{Y_1 + \cdots + Y_n}{n} \leq A\}$$

where A is such that

$$P[\frac{Y_1 + \cdots + Y_n}{n} \leq A|X = 0] = 5\%.$$

Now,

$$P[\frac{Y_1 + \cdots + Y_n}{n} \leq A|X = 0] = P(\frac{Z_1 + \cdots + Z_n}{n} \leq A)$$

where $Z_1, Z_2, \ldots$ are i.i.d. exponentially distributed with mean 1. Also,

$$P(\frac{Z_1 + \cdots + Z_n}{n} \leq A) = P([\frac{Z_1 + \cdots + Z_n}{n} - 1]\sqrt{n} \leq (A - 1)\sqrt{n}).$$

But, by the CLT,

$$[\frac{Z_1 + \cdots + Z_n}{n} - 1]\sqrt{n} \approx N(0, var(Z_1)) = N(0, 1).$$

Hence,

$$P(\frac{Z_1 + \cdots + Z_n}{n} \leq A) \approx P(N(0,1) \leq (A - 1)\sqrt{n}).$$

For this probability to equal to 5%, we need

$$(A - 1)\sqrt{n} = -1.64, \text{ or } A = 1 - \frac{1.64}{\sqrt{n}}.$$

Hence,

$$\hat{X} = 1\{\frac{Y_1 + \cdots + Y_n}{n} \leq 1 - \frac{1.64}{\sqrt{n}}\}.$$

2) We find that

$$P[\hat{X} = 1|X = 1] = P[\frac{Y_1 + \cdots + Y_n}{n} \leq A|X = 1] = P(\frac{V_1 + \cdots + V_n}{n} \leq A)$$

where the V_m are i.i.d. exponentially distributed with mean 0.9 year. Now,

$$P(\frac{V_1 + \cdots + V_n}{n} \le A) = P([\frac{V_1 + \cdots + V_n}{n} - 0.9]\sqrt{n} \le (A - 0.9)\sqrt{n}).$$

But, by the CLT,

$$[\frac{V_1 + \cdots + V_n}{n} - 0.9]\sqrt{n} \approx N(0, var(V_1)) = N(0, (0.9)^2).$$

Hence,

$$P[\hat{X} = 1 | X = 1] \approx P(N(0, (0.9)^2) \le (A - 0.9)\sqrt{n}) = P(N(0,1) \le \frac{(A - 0.9)\sqrt{n}}{0.9}).$$

Thus, for this probability to be 95%, we need

$$\frac{(A - 0.9)\sqrt{n}}{0.9} = 1.64,$$

so that

$$A = \frac{1.64 \times 0.9}{\sqrt{n}} + 0.9.$$

Combining with the expression for A derived in part 1), we find

$$1 - \frac{1.64}{\sqrt{n}} = \frac{1.64 \times 0.9}{\sqrt{n}} + 0.9.$$

Solving for n we find

$$n = [\frac{1.64 \times 1.9}{0.1}]^2 \approx 971.$$