Department of EECS - University of California at Berkeley

EECS 126 - Probability and Random Processes - Spring 2008

Midterm 1: 4/10/2008

Name (Last, First):

SID:

---

1. (15%)

For each of the following questions, choose the correct answer:

1. A random variable is *defined* as
   - a function of a random variable;
   - a cumulative probability distribution function;
   - **a (deterministic) function of the outcome of a random experiment;**
   - a set of possible values with their probabilities.
2. Two random variables X and Y are independent if and only if
   - they are uncorrelated;
   - the variance of their sum is the sum of their variances;
   - $E[X|Y] = E(X)$;
   - $g(X)$ and $h(Y)$ are uncorrelated for all functions $g(.)$ and $h(.)$;
   - **none of the above.**
3. Two random variables X and Y are jointly Gaussian if and only if
   - each is Gaussian;
   - they are independent and Gaussian;
   - **they are linear combinations of independent $N(0,1)$ random variables;**
   - their sum is Gaussian;
   - their mean values and covariance matrix are given.
4. The Maximum A Posteriori estimate $\hat{X}$ of X given Y is such that
   - $P(\hat{X} = X)$ is maximized over all the random variables $\hat{X}$;
   - $\hat{X} = g(Y)$ where $g(y)$ is the value of x that maximizes $P[Y = y|X = x]$;
   - **$\hat{X} = g(Y)$ where $g(y)$ is the value of x that maximizes $P[X = x|Y = y]$;**
   - $X - \hat{X} \perp h(Y)$ for all function $h(\cdot)$.

5. Three events A, B, C are mutually independent if and only if

- $P[A|B \cap C] = P(A)$ and $P(B \cap C) = P(B)P(C)$;
- $P(A \cap B \cap C) = P(A)P(B)P(C)$;
- $P[A \cap B|C] = P(C)$ and $P[C|A \cap B] = P(A \cap B)$;
- **None of the above.**

---

2. (20%)

Assume that X is Geometric(p), so that $P(X = n) = (1-p)^{n-1}p$ for $n \geq 1$. (Here, $p \in (0,1)$ is given.) Assume also that $Y_1, Y_2, \ldots$ are independent and Poisson(λ), so that $P(Y_k = n) = \lambda^n e^{-\lambda}/n!$ for $n \geq 0$ where $\lambda > 0$ is given. Moreover, assume that $\{X, Y_1, Y_2, \ldots\}$ are mutually independent. Define $Z = Y_1 + Y_2 + \cdots + Y_X$.

a) Calculate $E(Z)$;

b) Calculate $\mathrm{var}(Z)$;

c) Calculate $E[X|Z]$.

For part (a), we use the fact that $E[E[Z|X]] = E[Z]$. So to begin, we need to calculate the conditional expectation of Z given X. So, suppose that $X = x$. Then we have

$$\begin{aligned} E[Z|X=x] &= E[Y_1 + Y_2 + \ldots + Y_X | X = x] \\ &= E\left[\sum_{i=1}^{x} Y_i\right] \\ &= \sum_{i=1}^{x} E[Y_i] \\ &= \sum_{i=1}^{x} \lambda \\ &= x\lambda \end{aligned}$$

where in the fourth line we used the fact that the expectation of a Poisson random variable is λ. So from this, it is clear that

$$E[Z|X] = X\lambda$$

Therefore, we get

$$\begin{aligned} E[Z] &= E[E[Z|X]] \\ &= E[X\lambda] \\ &= E[X]\lambda \\ &= \frac{\lambda}{p} \end{aligned}$$

where in the fourth line we used that the expectation of a geometric random variable is $\frac{1}{p}$.

For part (b), we will again use the fact that $E[E[Z^2|X]] = E[Z^2]$, and combine that with the fact that we already know $E[Z]$ to calculate $\mathrm{var}(Z)$. So, suppose that $X = x$. Then we have

$$
\begin{aligned}
E[Z^2|X=x] &= E[(Y_1+Y_2+\ldots+Y_X)^2|X=x] \\
&= E[(Y_1+Y_2+\ldots+Y_x)^2] \\
&= E\left[\sum_{i=1}^{x} Y_i^2 + \sum_{i<j} 2Y_iY_j\right] \\
&= \sum_{i=1}^{x} E[Y_i^2] + 2\sum_{i<j} E[Y_iY_j] \\
&= \sum_{i=1}^{x} E[Y_i^2] + 2\sum_{i<j} E[Y_i]E[Y_j] \\
&= xE[Y_i^2] + 2\cdot\frac{x(x-1)}{2}\cdot E[Y_i]E[Y_j] \\
&= xE[Y_i^2] + x(x-1)E[Y_i]^2 \\
&= x(E[Y_i^2]-E[Y_i]^2) + x^2E[Y_i]^2 \\
&= x\cdot\mathrm{var}(Y_i) + x^2E[Y_i]^2 \\
&= x\lambda + x^2\lambda^2
\end{aligned}
$$

where in the fifth line we used the fact that the Y_i's are independent, in the sixth line we used the fact that they are identically distributed, in the seventh line we used the fact that they are identically distributed, and in the last line we used the fact that the mean and variance of a Poisson random variable is λ. So from this, it is clear that

$$E[Z^2|X] = X\lambda + X^2\lambda^2$$

Therefore, we get

$$
\begin{aligned}
E[Z^2] &= E[E[Z^2|X]] \\
&= E[X\lambda + X^2\lambda^2] \\
&= E[X]\lambda + E[X^2]\lambda^2 \\
&= E[X]\lambda + (\mathrm{var}(X)+E[X]^2)\lambda^2 \\
&= \frac{\lambda}{p} + \left(\frac{1-p}{p^2}+\frac{1}{p^2}\right)\lambda^2 \\
&= \frac{\lambda p + (2-p)\lambda^2}{p^2}
\end{aligned}
$$

where in the fifth line we used the fact that the mean of a geometric random variable is $\frac{1}{p}$ and the variance of a geometric random variable is $\frac{1-p}{p^2}$. Therefore, we get

$$
\begin{aligned}
\mathrm{var}(Z) &= E[Z^2] - E[Z]^2 \\
&= \frac{\lambda p + (2-p)\lambda^2}{p^2} - \frac{\lambda^2}{p^2} \\
&= \frac{\lambda p + (1-p)\lambda^2}{p^2}
\end{aligned}
$$

For part (c), we note that given $X = n$ the random variable Z is Poisson($n\lambda$). Thus,

$$P[Z = z | X = n] = \frac{(n\lambda)^z}{z!} e^{-n\lambda}, z \geq 0.$$

Hence,

$$P[X = n | Z = z] = \frac{n^z e^{-n\lambda}(1-p)^n}{\sum_{m=0}^{\infty} m^z e^{-m\lambda}(1-p)^m},$$

so that

$$E[X | Z = z] = \frac{\sum_{n=0}^{\infty} n^{z+1} e^{-n\lambda}(1-p)^n}{\sum_{m=0}^{\infty} m^z e^{-m\lambda}(1-p)^m} = \frac{H(z+1,\alpha)}{H(z,\alpha)}$$

where

$$H(z,\alpha) = \sum_{m=0}^{\infty} m^z \alpha^m \text{ and } \alpha = e^{-\lambda}(1-p).$$

Observe that

$$\frac{\partial}{\partial \alpha} H(z,\alpha) = \sum_m m^{z+1} \alpha^{m-1} = \alpha^{-1} H(z+1,\alpha).$$

Hence,

$$H(z+1,\alpha) = \alpha \frac{\partial}{\partial \alpha} H(z,\alpha).$$

Since $H(0,\alpha) = (1-\alpha)^{-1}$, we find successively

$$H(1,\alpha) = \alpha(1-\alpha)^{-2}, H(2,\alpha) = 2\alpha(1-\alpha)^{-3}, ...$$

so that

$$H(z,\alpha) = z!\alpha^z(1-\alpha)^{-z-1}.$$

This allows to compute

$$E[X | Z = z] = \frac{(z+1)!\alpha^{z+1}(1-\alpha)^{-z-2}}{z!\alpha^z(1-\alpha)^{-z-1}} = \frac{(z+1)\alpha}{1-\alpha} \text{ with } \alpha = e^{-\lambda}(1-p).$$

3. (15%)

Let $(X, \mathbf{Y})^T$ be $N([1,2,3]^T, \Sigma)$ where $\Sigma = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

a) Find the matrix A such that $X - A\mathbf{Y}$ is independent of $\mathbf{Y}$;
Hint: $Cov(Y, X - AY)$ might be helpful.

b) Calculate $E[X|\mathbf{Y}]$;

c) Show that, given $\mathbf{Y}$, $X = N(g(\mathbf{Y}), \sigma^2)$ and determine $g(\mathbf{Y})$ and σ^2.

For part (a), we follow the hint about looking at the covariance of $\mathbf{Y}$ and $X - A\mathbf{Y}$. First, let's consider the quantity $X - A\mathbf{Y}$. Since we are adding X to $A\mathbf{Y}$, we know that $A\mathbf{Y}$ must be a scalar (or a 1×1 matrix), therefore A must be a 1×2 matrix, and

$$X - A\mathbf{Y} = X - \left[\begin{array}{cc} a_1 & a_2 \end{array}\right] \left[\begin{array}{c} Y_1 \\ Y_2 \end{array}\right] = X - a_1 Y_1 - a_2 Y_2$$

From this, we also know that

$$E[X - A\mathbf{Y}] = E[X] - a_1 E[Y_1] - a_2 E[Y_2] = 1 - 2a_1 - 3a_2$$

Also, we know that $X - A\mathbf{Y}$ is Gaussian, since it is the sum of jointly Gaussian random variables, and also that $X - A\mathbf{Y}$ and $\mathbf{Y}$ are jointly Gaussian. From this, if we show that $Cov(\mathbf{Y}, X - A\mathbf{Y}) = 0$, then we know that $X - A\mathbf{Y}$ and $\mathbf{Y}$ are independent. So we get

$$\begin{aligned}
Cov(\mathbf{Y}, X - A\mathbf{Y}) &= E[(\mathbf{Y} - E[\mathbf{Y}])(X - A\mathbf{Y} - E[X - A\mathbf{Y}])^T] \\
&= E\left[\left(\left[\begin{array}{c} Y_1 \\ Y_2 \end{array}\right] - \left[\begin{array}{c} 2 \\ 3 \end{array}\right]\right)\left(\left[\begin{array}{c} X - a_1Y_1 - a_2Y_2 \end{array}\right] - \left[\begin{array}{c} 1 - 2a_1 - 3a_2 \end{array}\right]\right)^T\right] \\
&= E\left[\left[\begin{array}{c} Y_1 - 2 \\ Y_2 - 3 \end{array}\right]\left[\begin{array}{c} (X-1) - a_1(Y_1 - 2) - a_2(Y_2 - 3) \end{array}\right]\right] \\
&= E\left[\left[\begin{array}{c} (Y_1-2)(X-1) - a_1(Y_1-2)(Y_1-2) - a_2(Y_1-2)(Y_2-3) \\ (Y_2-3)(X-1) - a_1(Y_2-3)(Y_1-2) - a_2(Y_2-3)(Y_2-3) \end{array}\right]\right] \\
&= \left[\begin{array}{c} Cov(Y_1, X) - a_1 Cov(Y_1, Y_1) - a_2 Cov(Y_1, Y_2) \\ Cov(Y_2, X) - a_1 Cov(Y_2, Y_1) - a_2 Cov(Y_2, Y_2) \end{array}\right] \\
&= \left[\begin{array}{c} 2 - 2a_1 - a_2 \\ 1 - a_1 - a_2 \end{array}\right]
\end{aligned}$$

So we weant

$$\begin{aligned}
2 - 2a_1 - a_2 &= 0 \\
1 - a_1 - a_2 &= 0
\end{aligned}$$

From this we get

$$A = \left[\begin{array}{cc} a_1 & a_2 \end{array}\right] = \left[\begin{array}{cc} 1 & 0 \end{array}\right]$$

For part (b), we want the best possible estimate of X given that we know $\mathbf{Y}^T = (Y_1, Y_2)$. Since X and $\mathbf{Y}$ are jointly Gaussian, then we know that this estimate is going to be a linear function of $\mathbf{Y}$, that is

$$E[X|\mathbf{Y}] = E[X] + \Sigma_{X,\mathbf{Y}}\Sigma_{\mathbf{Y}}^{-1}(\mathbf{Y} - E[\mathbf{Y}])$$

First, we know that

$$\Sigma_{\mathbf{Y}} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

and so we get

$$\Sigma_{\mathbf{Y}}^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

We also get

$$\begin{aligned} \Sigma_{X,\mathbf{Y}} &= \begin{bmatrix} Cov(X, Y_1) & Cov(X, Y_2) \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \end{bmatrix} \end{aligned}$$

Therefore, we get

$$\begin{aligned} E[X|\mathbf{Y}] &= E[X] + \Sigma_{X,\mathbf{Y}}\Sigma_{\mathbf{Y}}^{-1}(\mathbf{Y} - E[\mathbf{Y}]) \\ &= 1 + \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} Y_1 - 2 \\ Y_2 - 3 \end{bmatrix} \\ &= 1 + \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} Y_1 - 2 \\ Y_2 - 3 \end{bmatrix} \\ &= 1 + Y_1 - 2 \\ &= Y_1 - 1 \end{aligned}$$

For part (c), we already know what $g(\mathbf{Y})$ is; it is $E[X|\mathbf{Y}] = Y_1 - 1$. We just need to figure out what σ^2 is; we calculate

$$\sigma^2 = E[(X - E[X|\mathbf{Y}])^2|\mathbf{Y}]$$

But, we know that $X - E[X|\mathbf{Y}]$ is independent of bfY; therefore, we get

$$\begin{aligned} \sigma^2 &= E[(X - E[X|\mathbf{Y}])^2|\mathbf{Y}] \\ &= E[(X - E[X|\mathbf{Y}])^2] \\ &= E[(X - 1 + Y_1)^2] \\ &= E[(X - 1) + (Y_1 - 2) + 2)^2] \\ &= Var(X) + Var(Y_1) + 2Cov(X, Y_1) + 2E[X - 1] + 2E[X - 2] + 4 \\ &= 4 + 2 + 4 + 0 + 0 + 4 \\ &= 14 \end{aligned}$$

Therefore, $X = N(Y_1 - 1, 14)$.

4. (15%)

Let X and Y be independent random variables having joint pdf $f_{XY}(x,y)$. We would like to compute the pdf of $Z = XY$.

1. Let $W = X$. Show that the joint pdf of Z and W can be written as

$$f_{ZW}(z,w) = \left|\frac{1}{w}\right| f_{XY}(w, \frac{z}{w})$$

   Hint: Think of (Z,W) as a function of (X,Y).

2. Assume that X and Y are uniform $U(0,1)$. Use the previous question to compute the pdf $f_Z(z)$ of Z.

---

1. We define the function $h(\cdot)$ as

$$\begin{array}{rcl} h : \mathbf{R}^2 & \rightarrow & \mathbf{R}^2 \\ (z,w) & \mapsto & h(x,y) = (xy, x) \end{array}$$

   First notice that $h(\cdot)$ is a 1-to-1 mapping.
   Now we can use the formula for the density of a function of a random variable to compute $f_{ZW}(z,w)$.

$$f_{ZW}(z,w) = \frac{1}{|J(h^{-1}(z,w))|} f_{XY}(h^{-1}(z,w))$$

   The Jacobian $J(\cdot)$ is given by

$$J(x,y) = \left(\begin{array}{cc} y & x \\ 1 & 0 \end{array} \right) = -x$$

   By applying the formula, we get

$$f_{ZW}(z,w) = \frac{1}{|w|} f_{XY}(w, z/w)$$

2. X and Y being independent uniform random variables, their joint pdf is given by

$$f_{XY}(x,y) = \begin{cases} 1 & \text{if } 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   This gives that

$$f_{ZW}(z,w) = \begin{cases} 1 & \text{if } 0 < z \leq w \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   The marginal pdf of Z can be computed by integrating the joint pdf over all possible values of w.

$$f_Z(z) = \int_0^1 f_{ZW}(z,w)dw = \int_z^1 \frac{1}{w}dw = ln(w)|_z^1 = -ln(z) \quad 0 < z \leq 1$$

5. (10%)

Given that $X = 0$, the random variables $Y_1, \ldots, Y_n$ are i.i.d Poisson with parameter λ. If, in the other hand $X = 1$, $Y_1, \ldots, Y_n$ are i.i.d Poisson with rate $\lambda_1 > \lambda_0$.
You observe $Y_1, \ldots, Y_n$, and based on that observation, you would like to decide whether $X = 0$ or $X = 1$.
You can assume that $Pr(X = 0) = Pr(X = 1) = 0.5$.
What is your decision rule if you are interested in maximizing the probability $P(\hat{X} = X|Y)$?

Notice that we have several observations Y_j, $j = 1, \ldots, n$ based upon which we would like to detect X. Also, since we want to maximize the posterior probability $P(\hat{X} = X|Y)$, our detection rule corresponds to the maximum a posteriori (MAP) rule. However, the events $X = 0$ and $X = 1$ are equally likely, which simplifies the detection rule to the maximum likelihood (ML).
To compute the likelihood ratio, we need the conditional joint distributions of the observations $P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n|X = 1)$ and $P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n|X = 0)$. But, given X, the observations are independent. Hence $P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n|X = i) = P(Y_1 = y_1|X = i)P(Y_2 = y_2|X = i) \ldots P(Y_n = y_n|X = i)$. This gives

$$\begin{aligned}
\Lambda(y_1, \ldots, y_n) &= \frac{P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n|X = 1)}{P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n|X = 0)} \\
&= \frac{P(Y_1 = y_1|X = 1)P(Y_2 = y_2|X = 1) \ldots P(Y_n = y_n|X = 1)}{P(Y_1 = y_1|X = 0)P(Y_2 = y_2|X = 0) \ldots P(Y_n = y_n|X = 0)}
\end{aligned}$$

Also given $X = i$, the observations have the same distribution $P(Y_j = y_j|X = i) = e^{-\lambda_i} \frac{(\lambda_i)^{y_j}}{y_j!}$.
Putting all this together, we get

$$\Lambda(y_1, \ldots, y_n) = \frac{e^{-\lambda_1} \frac{(\lambda_1)^{y_1}}{y_1!} \ldots e^{-\lambda_1} \frac{(\lambda_1)^{y_n}}{y_n!}}{e^{-\lambda_0} \frac{(\lambda_0)^{y_1}}{y_1!} \ldots e^{-\lambda_0} \frac{(\lambda_0)^{y_n}}{y_n!}}$$

After some simplifications, we get

$$\Lambda(y_1, \ldots, y_n) = e^{-n(\lambda_1 - \lambda_0)} \left(\frac{\lambda_1}{\lambda_0}\right)^{\sum_{j=1}^{n} y_j}$$

$\Lambda(y_1, \ldots, y_n)$ is to be compared to 1, and $\hat{X} = 1$ is $\Lambda(y_1, \ldots, y_n) \geq 1$ ($\hat{X} = 0$ otherwise). Note that it is easier to work with the log-likelihood-ratio

$$l(y_1, \ldots, y_n) = ln(\Lambda(y_1, \ldots, y_n)) = -n(\lambda_1 - \lambda_0) + ln\left(\frac{\lambda_1}{\lambda_0}\right) \sum_{j=1}^{n} y_j$$

$\hat{X} = 1$ if

$$l(y_1, \ldots, y_n) \geq 0 \Leftrightarrow \frac{1}{n} \sum_{j=1}^{n} y_j \geq \frac{\lambda_1 - \lambda_0}{ln\left(\frac{\lambda_1}{\lambda_0}\right)}$$

Note: here we see that we do not need to observe the individual y_j's for the detection; only the sample mean $\bar{m} = \frac{1}{n} \sum_{j=1}^{n} y_j$ is enough. We say that $\bar{m}$ is a sufficient statistic for this detection.

6. (15%)

Assume that the pair (X, Y) is equally likely to be any one of the eight values indicated in the figure below.

1. Specify the random variable $E[X|Y]$ as a function of Y;
2. What is the p.m.f. of $E[X|Y]$?
3. Are X and Y uncorrelated? Explain.

---

1. To specify $E[X|Y]$, we compute it for each value of Y. Using the fact that (X, Y) is equally likely to be any one of the eight values, we get

$$E[X|Y] = \begin{cases} 1 & \text{if } Y = 2 \\ 1 & \text{if } Y = 0 \\ 1 & \text{if } Y = 1 \\ 4/3 & \text{if } Y = -1 \end{cases}$$

   Notice that $E[X|Y]$ takes only two values.

2. The pmf of $E[X|Y]$ is given by

$$E[X|Y] = \begin{cases} 1 & w.p \;\; 5/8 \\ 4/3 & w.p \;\; 3/8 \end{cases}$$

3. We can easily verify that $E[XY] = 2/8 \neq E[X]E[Y] = 27/80$. Thus X and Y are uncorrelated.

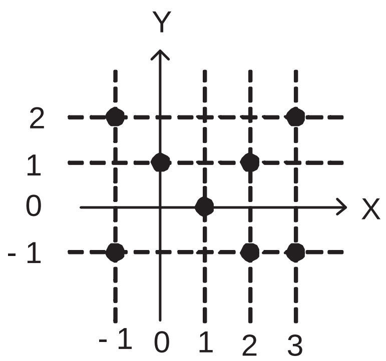

Figure 1: Sample space of the random variables X, Y defined in Problem 6.