December 21, 2005

You have 180 minutes to complete this exam. Items with unspecified point values are worth ten points. Put your name and your Rutgers netid (but no part of your SSN) on each exam book *(10 points)*. Please read both sides of the exam carefully and ask the instructor if you have any questions. **You must complete problems 1, 2 and 3 and EITHER problem 4 or problem 5. You must indicate in the front cover of your exam book whether you wish to have problem 4 or problem 5 graded.**

1. *70 points* Please answer the following questions

   (a) Write down a code $\mathbf{c}$, an integer from the set $\{1000, 1001, \ldots, 9999\}$, that you can remember but will seem random to other students in the class. (This secret code will be used to return your exam scores by email so do not reveal your code to others. Note this item has no wrong answers.)

   (b) Suppose an unending sequence of students enter this exam room. Student n, the nth student to enter, is asked to choose randomly and equiprobably among all possible codes (and independent of all prior choices) a code $\mathbf{c}_n$, just as you did in part (a). Let X_n denote the number of *unique* codes in the set $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$. Obviously $X_1 = 1$. Let $D_n = 1$ if code $\mathbf{c}_n$ is different from all prevously chosen code words $\mathbf{c}_1, \ldots, \mathbf{c}_{n-1}$; otherwise $D_n = 0$. Find the conditional PMF $P_{D_n|X_{n-1}}(d|x)$.

   Given $X_{n-1} = x$, x unqiue codes have already been chosen, Thus $D_n = 1$ iff student n chooses one of the $9000 - x$ unchosen codes. This occurs with probability $1 - x/9000$. The conditional PMF of D_n is

$$P_{D_n|X_{n-1}}(d|x) = \begin{cases} x/9000 & d = 0, \\ 1 - x/9000 & d = 1, \\ 0 & \text{otherwise.} \end{cases}$$

   (c) Using the fact that $X_n = X_{n-1} + D_n$, find the conditional expected values $E[X_n|X_{n-1} = x]$ and $E[X_n|X_{n-1}]$.

$$\begin{aligned} E[X_n|X_{n-1} = x] &= E[X_{n-1} + D_n|X_{n-1} = x] \\ &= E[X_{n-1}|X_{n-1} = x] + E[D_n|X_{n-1} = x] \\ &= x + 1 - \frac{x}{9000} = 1 + \frac{8999}{9000}x \end{aligned}$$

   It follows directly that

$$E[X_n|X_{n-1}] = 1 + \frac{8999}{9000}X_{n-1}.$$

   (d) Find $E[X_n]$. Hint: Let $\mu_n = E[X_n]$ and find a recursion for μ_n.

$$E[X_n] = E[E[X_n|X_{n-1}]] = E[1 + \alpha X_{n-1}] = 1 + \alpha E[X_{n-1}].$$

   Let $\alpha = 8999/9000$. By the iterated expectation,

$$E[X_n] = E[E[X_n|X_{n-1}]] = E[1 + \alpha X_{n-1}] = 1 + \alpha E[X_{n-1}].$$

   Following the hint, we define $\mu_n = E[X_n]$ so that

$$\mu_n = 1 + \alpha\mu_{n-1}.$$

Since $\mu_1 = 1$, we observe that

$$\mu_2 = 1 + \alpha\mu_1 = 1 + \alpha$$
$$\mu_3 = 1 + \alpha\mu_2 = 1 + \alpha + \alpha^2$$

and so on. In fact, it follows that

$$\mu_n = 1 + \alpha + \cdots + \alpha^{n-1} = \frac{1-\alpha^n}{1-\alpha}.$$

(e) What is $\lim_{n\to\infty} E[X_n]$? Explain your answer.
From the previous answer,

$$\lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \mu_n = \frac{1}{1-\alpha} = 9000.$$

In fact, this answer could be derived from first principles, Eventually for for some sufficently large $n = n*$, all 9000 unique codes are produced, yielding $X_{n^*} = 9000$. Moreoever, for all $n > n^*$, $X_n = 9000$. It follows that for $n > n^*$ that $E[X_n] = 9000$.

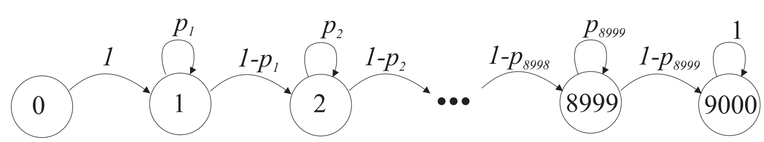
(f) *20 points* Suppose we define $X_0 = 0$. Is the random sequence $\{X_n | n = 0, 1, 2, \ldots\}$ a Markov chain? If not, explain why. If so, explain why, find and sketch the Markov chain, and find the limiting state probabilities (or explain why they don't exist.)
The sequence $X_0, X_1, \ldots$ is a Markov chain. To show this, we frist observe that

$$P[X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_0 = 0] = P[D_n = x_n - x_{n-1} | X_{n-1} = x_{n-1}, \ldots, X_0 = 0].$$

Given the past history $X_{n-1} = x_{n-1}, \ldots, X_0 = 0$, $D_n = 0$ with probability $x_{n-1}/9000$ independent of X_i for $i < n-1$. All that matters is that x_{n-1} unique codes have been chosen by time $n-1$. Thus,

$$\begin{aligned} P[X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_0 = 0] &= P[D_n = x_n - x_{n-1} | X_{n-1} = x_{n-1}] \\ &= P[X_n = x_n | X_{n-1} = x_{n-1}]. \end{aligned}$$

Using $p_n = n/9000$, the corresponding Markov chain is



This chain has transient states $0, 1, \ldots, 9000$ and the sole absorbing state 9000 which is by itself the single recurrent communicating class. No matter what state the system starts in, the system eventually ends up in state 9000. The limiting state probabilities are $\pi_i = 0$ for $i = 0, 1, \ldots, 8999$ and $\pi_{9000} = 1$.

2. *40 points* Let X_n be a wide sense stationary Gaussian random sequence with expected value $E[X_n] = 0$ and autocorrelation function

$$R_X[k] = E[X_n X_{n+k}] = 2^{-|k|}.$$

We observe the noisy random sequence $Y_n = X_n + Z_n$ where Z_n is an iid Gaussian noise sequence, independent of X_n, with $E[Z_n] = 0$ and $\mathrm{Var}[Z_n] = 1/2$.

(a) Find the LMSE estimate $\hat{X}_n$ of X_n given *only* the observation Y_n. Note that $\hat{X}_n$ cannot use prior observations $Y_{n-1}, Y_{n-2}, \ldots$ That is, $\hat{X}_n$ is the output of an order 0 filter.

Since both X_n and Z_n have zero expected value, $Xhat_n = aY_n$ for the optimal choice of a. We could derive the optimal a from first principles, or we could use apply our LMSE Estimator formula

$$\hat{X}_n = R_{Y_n}^{-1} R_{Y_n X_n} Y_n.$$

In this case, all terms are simply scalars. In particular,

$$R_{Y_n} = \text{Var}[Y_n] = \text{Var}[X_n] + \text{Var}[Z_n] = 1 + \frac{1}{2} = \frac{3}{2}$$

and

$$R_{Y_n X_n} = E[Y_n X_n] = E[(X_n + Z_n)X_n] = E\left[X_n^2\right] = 1.$$

It follows that

$$\hat{X}_n = \frac{2}{3} Y_n.$$

(b) What is the mean square error e_0 of the estimator $\hat{X}_n$?

From first principles, the mean squared error is

$$e_0 = E\left[(\hat{X}_n - X_n)^2\right] = E\left[\left(\frac{2}{3}Y_n - X_n\right)^2\right] = E\left[\left(\frac{2}{3}(X_n + Z_n) - X_n\right)^2\right].$$

Simplifying and then expanding the square, we find that

$$e_0 = E\left[\left(-\frac{1}{3}X_n + \frac{2}{3}Z_n\right)^2\right] = E\left[\frac{1}{9}X_n^2 - \frac{4}{9}X_n Z_n + \frac{4}{9}Z_n^2\right] = \frac{1}{9} + \frac{4}{9}\frac{1}{2} = \frac{1}{3}.$$

(c) What is the PDF $f_{\hat{X}_n}(x)$ of $\hat{X}_n$?

Since $\hat{X}_n = (2/3)X_n + (2/3)Z_n$ is the sum of independent Gaussian random variables, we know that $\hat{X}_n$ is Gaussian. Moreover,

$$E\left[\hat{X}_n\right] = \frac{2}{3}(E[X_n] + E[Z_n]) = 0, \qquad \text{Var}[\hat{X}_n] = \frac{4}{9}\text{Var}[X_n] + \frac{4}{9}\text{Var}[Z_n] = \frac{2}{3}.$$

Since $\hat{X}_n$ is Gaussian, it has PDF

$$f_{\hat{X}_n}(x) = \frac{1}{\sqrt{2\pi(2/3)}} e^{-x^2/[2(2/3)]} = \sqrt{\frac{3}{4\pi}} e^{-3x^2/4}.$$

(d) Find the conditional expectation $E[X_n|Y_n]$.

Because X_n, Z_n, Y_n and $\hat{X}_n$ are all jointly Gaussian, the LMSE estimate $\hat{X}_n$ equals the conditional expected value $E[X_n|Y_n]$. That is $E[X_n|Y_n] = \hat{X}_n = 2Y_n/3$.

3. *40 points* The random sequence $X_0, X_1, \ldots$ is an iid random sequence such that each X_n is a Gaussian $(0,1)$ random variable. $N(t)$ is a Poisson process of rate λ that is independent of X_n. Let $\{Y(t)|t \geq 0\}$ denote a random process defined by $Y(t) = X_{N(t)}$.

(a) Find $E[Y(t)]$ and the PDF $f_{Y(t)}(y)$.

First we condition on $N(t) = n$ and find that $Y(t)$ has conditional PDF

$$f_{Y(t)|N(t)=n}(y) = f_{X_n|N(t)=n}(y) = f_{X_n}(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

Note that the second step required that X_n be independent of $N(t)$. The conditional PDF of $Y(t)$ given $N(t) = n$ is independent of $N(t)$. Thus

$$f_{Y(t)}(y) = f_{Y(t)|N(t)=n}(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

We see that $Y(t)$ is a Gaussian $(0,1)$ random variable and thus $E[Y(t)] = 0$.

(b) Find the autocorrelation $R_Y(t,\tau)$. (Assume $|\tau| < t$.)

We will find $R_Y(t,\tau) = E[Y(t)Y(t+\tau)] = E\left[X_{N(t)}X_{N(t+\tau)}\right]$ using conditional expectation. The key issue in calculating the expectation is whether $N(t) = N(t+\tau)$. If $N(t) = N(t+\tau) = n$, then $X_{N(t)} = X_{N(t+\tau)} = X_n$ and $R_Y(t,\tau) = E\left[X_n^2\right] = 1$. If $N(t) = n$ and $N(t+\tau) = n' \neq n$, then

$$R_Y(t,\tau) = E[X_n X_{n'}] = E[X_n]\, E[X_{n'}] = 0.$$

To finish the problem, let $N = N(t+\tau) - N(t)$. Note that $N = 0$ if there are zero arrivals of the Poisson process in the interval $[t, t+\tau]$ (or $[t+\tau, t]$ if $\tau < 0$). Thus $P[N=0] = e^{-\lambda|\tau|}$. Finally,

$$\begin{aligned} R_Y(t,\tau) &= P[N=0]\, E[Y(t)Y(t+\tau)|N=0] + P[N\neq 0] \underbrace{E[Y(t)Y(t+\tau)|N\neq 0]}_{=0} \\ &= P[N=0] = e^{-\lambda|\tau|} \end{aligned}$$

(c) *20 points* Is $Y(t)$ a Gaussian random process? Hint: Find the PDF of $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 \end{bmatrix}' = \begin{bmatrix} Y(t_1) & Y(t_2) \end{bmatrix}'$ where $t_2 = t_1 + \tau > t_1$.

For the process to be Gaussian, Y_1 and Y_2 must have a bivariate Gaussian PDF. Since we found in part (a) that Y_1 has a Gaussian PDF, we can just find whther the condtional PDF of Y_2 given Y_1 is Gaussian. As in previous parts, what matters is if $N = N(t+\tau) - N(t) = 0$. If $N = 0$, then $Y_2 = Y_1$ so that

$$f_{Y_2)|Y_1,N=0}(y_2|y_1) = \delta(y_2 - y_1).$$

However, if $N = n \neq 0$, then $Y_2 = X_{N(t)+n}$ and $Y_1 = X_{N(t)}$ are independent and

$$f_{Y_2)|Y_1,N=n}(y_2|y_1) = f_{Y_2}(y_2) = \frac{1}{\sqrt{2\pi}} e^{-y_2^2/2}.$$

In fact, whenever $N \neq 0$,

$$f_{Y_2)|Y_1,N\neq 0}(y_2|y_1) = f_{Y_2}(y_2) = \frac{1}{\sqrt{2\pi}} e^{-y_2^2/2}.$$

Finally,

$$\begin{aligned} f_{Y_2|Y_1}(y_2|y_1) &= P[N=0]\, f_{Y_2)|Y_1,N=0}(y_2|y_1) + P[N\neq 0]\, f_{Y_2)|Y_1,N\neq 0}(y_2|y_1) \\ &= e^{-\lambda|\tau|}\delta(y_2 - y_1) + (1 - e^{-\lambda|\tau|}) \frac{1}{\sqrt{2\pi}} e^{-y_2^2/2}. \end{aligned}$$

Since this conditional PDF is not Gaussian, Y_1 and Y_2 are not bivariate Gaussian and thus Y_n is not a Gaussian random sequence.

4. *40 points* A group of n people form a football pool. The rules of this pool are simple: 16 football games are played each week. Each contestant must pick the winner of each game against a point *spread* specified by Las Vegas oddsmakers. The contestant who picks the most games corectly over a 16 week season wins the pool. The spread is a point difference d such that picking the favored team is a winning pick only if that team wins by more than d points; otherwise, the pick of the opposing team is a winner. Setting the spread is a tricky task. For our purposes, we will assume that the oddsmakers use their expertise to attempt to set a spread such that a random pick of either team will be a winner with probability 1/2, independent of the past history of the football season and any other games played. On the other hand, each contestant in the pool can devote himself to the study of team's past histories, performance trends, official injury reports, the coach's weekly press conference, chat room gossip and any other wisdom that might help in placing a winning bet.

After m weeks, each pool contestant will have picked $16m$ games. Each contestant i will have picked W_i games correctly where $0 \leq W_i \leq 16m$. For example, after $m = 14$ weeks, $16(14) = 224$ games have been played and in my pool the leader (call him Jake) has picked 119 games correctly while the worst contestant (call him Fred) has picked 85 games correctly[1] The interesting probability question is whether the contestants actually have any ability to pick football games, or are the contestants essentially just picking games at random as the oddsmakers intend? In particular, does the pool leader have skills or is he just lucky? To address this question, we wish to design a significance test to determine whether the pool leader actually has any skill at picking games. Let H_0 denote the null hypothesis that *all* players, including the leader, pick winners in each game with probability $p = 1/2$, independent of the outcome of any other game. In the following, you may use a central limit theorem approximation for binomial PMFs as needed.

Suppose the pool has $n = 38$ contestants. Based on the observation of W, the number of winning picks by the pool leader after m weeks of the season, design a one-sided significance test for hypothesis H_0 at significance level $\alpha = 0.05$. *You must justify your choice of the rejection region.* Given that Jake is the leader with 119 winning picks in $m = 14$ weeks, do you reject or accept the hypothesis H_0.

This problem has a lot of words, but is not all that hard. The pool leader has picked $W = \max(W_1, \ldots, W_n)$ *games correctly. Under hypothesis* H_0*, the* W_i *are iid with PDF*

$$P_{W_i|H_0}(w) = \binom{16m}{w} \left(\frac{1}{2}\right)^w \left(\frac{1}{2}\right)^{16m-w}.$$

Since $E[W_i|H_0] = 8m$ *and* $\mathrm{Var}[W_i|H_0] = 16m(1/2)(1/2) = 4m$*, we can use a Central Limit theorem approximation to write*

$$P[W_i \leq w|H_0] = P\left[\frac{W_i - 8m}{2\sqrt{m}} \leq \frac{w - 8m}{2\sqrt{m}}\right] = \Phi\left(\frac{w - 8m}{s\sqrt{m}}\right).$$

Givern H_0*, the conditional CDF for* W *is*

$$\begin{aligned} P[W \leq w|H_0] &= P[\max(W_1, \ldots, W_n) \leq w|H_0] \\ &= P[W_1 \leq w, \ldots, W_n \leq w|H_0] \\ &= P[W_1 \leq w|H_0] \cdots P[W_n \leq w|H_0] \\ &= (P[W_i \leq w|H_0])^n = \Phi^n\left(\frac{w - 8m}{2\sqrt{m}}\right) \end{aligned}$$

[1] If you're curious, I personally have picked 100 games correctly. That's lousy but I am pleased to be well ahead of Fred.

We choose the rejection region such that $W > w*$ because we want to reject the hypothesis H_0 that everyone is merely guessing if the leader does exceptionally well. Thus,

$$\alpha = P[R] = P[W > w^*|H_0] = 1 - \Phi^n\left(\frac{w^* - 8m}{2\sqrt{m}}\right).$$

For $\alpha = 0.05$, we find that

$$\Phi\left(\frac{w^* - 8m}{2\sqrt{m}}\right) = (0.95)^{1/n}.$$

For $n = 38$,

$$Q\left(\frac{w^* - 8m}{2\sqrt{m}}\right) = 1 - (0.95)^{1/38} = 1.35 \times 10^{-3}.$$

It follows that

$$\frac{w^* - 8m}{2\sqrt{m}} = 3,$$

or

$$w^* = 8m + 6\sqrt{m}.$$

After $m = 14$, we require $w^* = 134.5$. Thus, if the leader Jake has $w \geq 135$ winning picks after 14 weeks, then we accept the hypothesis that Jake has an ability to pick winners better than random selection.

5. *40 points* Recall that when we use an observation vector $\mathbf{Y}$ to form a linear estimate of a random variable X, the optimal estimator is $\hat{X} = \mathbf{R}'_{\mathbf{Y}X}\mathbf{R}_{\mathbf{Y}}^{-1}\mathbf{Y}$. If we want to form a optimal linear estimate $\hat{\mathbf{X}}$ of a vector $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}'$ using an observation $\mathbf{Y}$, we can form an optimal estimator for each X_i of the form $\hat{X}_i = \mathbf{R}'_{\mathbf{Y}X_i}\mathbf{R}_{\mathbf{Y}}^{-1}\mathbf{Y}$. A little bit of algebra will show that we can write the vector of estimates as $\hat{\mathbf{X}} = \mathbf{R}'_{\mathbf{YX}}\mathbf{R}_{\mathbf{Y}}^{-1}\mathbf{Y}$.

   Now suppose Y_k is a noisy version of a quadratic function. That is,

   $$Y_k = q_0 + q_1 k + q_2 k^2 + Z_k$$

   where $q_0 + q_1 k + q_2 k^2$ is an unknown quadratic function of k and Z_k is a sequence of iid Gaussian $(0, 1)$ noise random variables. We wish to estimate the unkown parameters q_0, q_1 and q_2 of the quadratic function. Suppose we assume q_0, q_1 and q_2 are iid Gaussian $(0, 1)$ random variables. Find the optimal linear estimator $\hat{\mathbf{Q}}(\mathbf{Y})$ of $\mathbf{Q} = \begin{bmatrix} q_0 & q_1 & q_2 \end{bmatrix}'$ given the observation $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_1 & \cdots & Y_n \end{bmatrix}'$. Hint: The answer will be in terms of a certain $n \times 3$ matrix. You do not need to try to compute inverses.

   The key to this problem is to write $\mathbf{Y}$ in terms of $\mathbf{Q}$. First we observe that

   $$\begin{aligned} Y_1 &= q_0 + 1q_1 + 1^2 q_2 + Z_1 \\ Y_2 &= q_0 + 2q_1 + 2^2 q_2 + Z_2 \\ \vdots &\quad \vdots \\ Y_n &= q_0 + nq_1 + n^2 q_2 + Z_n \end{aligned}$$

   In terms of the vector $\mathbf{Q}$, we can write

   $$\mathbf{Y} = \underbrace{\begin{bmatrix} 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ \vdots & \vdots & \vdots \\ 1 & n & n^2 \end{bmatrix}}_{\mathbf{K}} \mathbf{Q} + \mathbf{Z}$$

Thus we see that $\mathbf{Y} = \mathbf{KQ} + \mathbf{Z}$.

From the problem statement we know that

$$E[\mathbf{Q}] = E[\mathbf{Z}] = 0, \qquad \mathbf{R_Q} = \mathbf{R_Z} = \mathbf{I}.$$

Following the vector version of optimal linear estimator as given, we wish to find

$$\hat{\mathbf{Q}} = \mathbf{R}'_{\mathbf{YQ}}\mathbf{R}_{\mathbf{Y}}^{-1}\mathbf{Y}.$$

We need to find

$$\begin{aligned}\mathbf{R_{YQ}} &= E\left[\mathbf{YQ}'\right] \\ &= E\left[(\mathbf{KQ} + \mathbf{Z})\mathbf{Q}'\right] = \mathbf{K}E\left[\mathbf{QQ}'\right] = \mathbf{K}.\end{aligned}$$

We also need

$$\begin{aligned}\mathbf{R_Y} &= E\left[\mathbf{YY}'\right] \\ &= E\left[(\mathbf{KQ} + \mathbf{Z})(\mathbf{KQ} + \mathbf{Z})'\right] \\ &= E\left[(\mathbf{KQ} + \mathbf{Z})(\mathbf{Q}'\mathbf{K}' + \mathbf{Z}')\right] \\ &= \mathbf{K}E\left[\mathbf{QQ}'\right]\mathbf{K}' + \mathbf{K}E\left[\mathbf{QZ}'\right] + E\left[\mathbf{ZQ}'\right]\mathbf{K}' + E\left[\mathbf{ZZ}'\right] \\ &= \mathbf{KK}' + \mathbf{I}\end{aligned}$$

It follows that

$$\hat{\mathbf{Q}} = \mathbf{R}'_{\mathbf{YQ}}\mathbf{R}_{\mathbf{Y}}^{-1}\mathbf{Y} = \mathbf{K}'(\mathbf{KK}' + \mathbf{I})^{-1}\mathbf{Y}.$$