

December 21, 2005

You have 180 minutes to complete this exam. 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 inside the front cover of your exam book whether to grade problem 4 or problem 5.**

x	0.0	0.5	1.0	1.5	2.0	2.5	3
$Q(x) = 1 - \Phi(x)$	0.500	0.3085	0.1587	0.0668	0.02275	0.00621	$1.35 \cdot 10^{-3}$

1. 70 points Please answer the following questions

- Write down a code \mathbf{c} , an integer from the set $\{1000, 1001, \dots, 9999\}$, 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.)
- Suppose an unending sequence of students enter this exam room. Student i , the i th student to enter, is asked to choose randomly and equiprobably among all possible codes (and independent of all prior choices) a code \mathbf{c}_i , just as you did in part (a). Let X_i denote the number of *unique* codes in the set $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_i\}$. Obviously $X_1 = 1$. Let $D_i = 1$ if code \mathbf{c}_i is different from all previously chosen code words $\mathbf{c}_1, \dots, \mathbf{c}_{i-1}$; otherwise $D_i = 0$. Find the conditional PMF $P_{D_i|X_{i-1}}(d|x)$.
- Find $E[X_i|X_{i-1} = x]$ and $E[X_i|X_{i-1}]$. Hint: $X_i = X_{i-1} + D_i$.
- Find $E[X_n]$. Hint: Let $\mu_n = E[X_n]$ and find a recursion for μ_n .
- What is $\lim_{n \rightarrow \infty} E[X_n]$? Explain your answer.
- 20 points Suppose we define $X_0 = 0$. Is the random sequence $\{X_n|n = 0, 1, 2, \dots\}$ 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.)

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 $\text{Var}[Z_n] = 1/2$.

- 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}, \dots . That is, \hat{X}_n is the output of an order 0 filter.
 - What is the mean square error e_0 of the estimator \hat{X}_n ?
 - What is the PDF $f_{\hat{X}}(x)$ of \hat{X} ?
 - Find the conditional expectation $E[X_n|Y_n]$.
3. 40 points The random sequence X_0, X_1, \dots 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 $Y(t) = X_{N(t)}$.
- Find $E[Y(t)]$ and the PDF $f_{Y(t)}(y)$.
 - Find the autocorrelation $R_Y(t, \tau)$. (Assume $|\tau| < t$.)

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

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 correctly 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 correctly¹. 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 .

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 $\hat{X} = \mathbf{R}'_{\mathbf{Y}X} \mathbf{R}^{-1}_{\mathbf{Y}} \mathbf{Y}$. If we want to form an optimal linear estimate $\hat{\mathbf{X}}$ of a vector $\mathbf{X} = [X_1 \ X_2 \ \cdots \ X_n]'$ using an observation \mathbf{Y} , we can form an optimal estimator for each X_i of the form $\hat{X}_i = \mathbf{R}'_{\mathbf{Y}X_i} \mathbf{R}^{-1}_{\mathbf{Y}} \mathbf{Y}$. A little bit of algebra will show that we can write the vector of estimates as $\hat{\mathbf{X}} = \mathbf{R}'_{\mathbf{Y}\mathbf{X}} \mathbf{R}^{-1}_{\mathbf{Y}} \mathbf{Y}$.

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 unknown 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 $\mathbf{Q} = [q_0 \ q_1 \ q_2]'$ given the observation $\mathbf{Y} = [Y_1 \ Y_1 \ \cdots \ Y_n]'$. Hint: The answer will be in terms of a certain $n \times 3$ matrix. You do not need to try to compute inverses.

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