Qualifying Exam for
Statistics and Biostatistics

M.A. Theory Exam

April 1, 2000

10:00am – 1:00pm

Instructions: This exam consists of a total of six questions – two probability questions (MA 581), two statistics questions (MA 582), and two stochastic processes questions (MA 583). Statistics M.A. candidates are to answer a total of four questions, with at least one question from each category. Biostatistics M.A. candidates are to answer a total of three questions, consisting of one MA 581 question and two MA 582 questions.

Write your name on the front of the bluebooks. Use a separate blue book for each major topic (MA 581, MA 582, and MA 583). Number the questions clearly in your work and start each question on a new page. You must show your work to make it clear how you obtained your answers. Answers without any work may lose credit even if they are correct, and will receive no credit if incorrect.

This is a closed-book three-hour exam. You may not refer to your notes or textbooks.

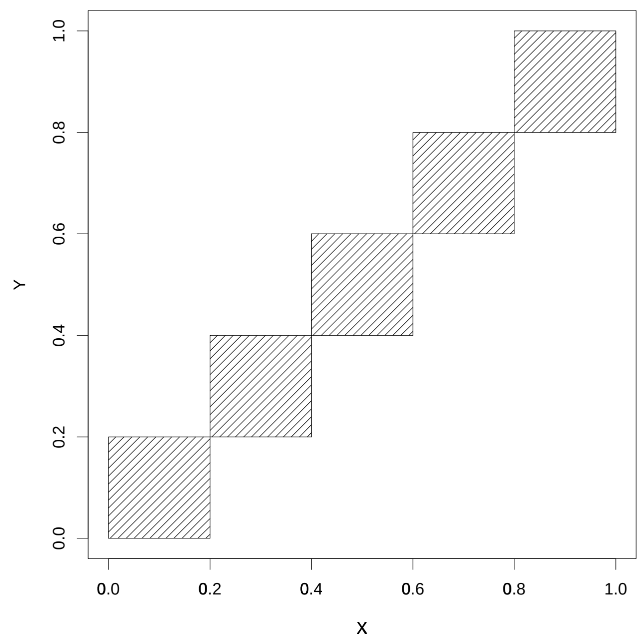
Probability (MA 581) Problems

April 2000

1. For fixed positive K, suppose X and Y have the joint probability density function

$$f(x,y) = \begin{cases} c & \text{if both } \frac{k}{K} < x \leq \frac{k+1}{K} \text{ and} \\ & \frac{k}{K} < y \leq \frac{k+1}{K},\ k = 0, 1, \ldots, K-1 \\ 0 & \text{otherwise} \end{cases}$$

For example, below is the region over which the joint density is non-zero when $K = 5$.



(a) Find the value of c (as a function of K) that makes $f(x,y)$ a proper density function.

(b) Determine the marginal distribution of X.

(c) Find the correlation between X and Y as a function of K. What happens to the correlation as K becomes large?

2. Suppose X and Y are jointly distributed random variables with mass function

		X		
		1	2	3
	2	1/12	1/6	1/12
Y	3	1/6	0	1/6
	4	0	1/3	0

(a) Find $P(X > 7/Y)$.

(b) Show that X and Y are not independent.

(c) Determine the joint distribution of two new random variables, U and V, that are independent, but that the marginal distribution of U is the same as X, and the marginal distribution of V is the same as Y.

(d) Determine the distribution of the random variable $X + Y$.

Statistics (MA 582) Problems

April 2000

1. Let X be a random variable with cumulative distribution function (cdf) F_X, and suppose $X_1, \ldots, X_n$ are a random sample of size n.

   (a) How might you estimate $\mathrm{E}(e^{-2X})$ from the sample if you did not know the specific form of F_X? Determine an expression for the estimate.

   (b) Suppose $Y = \min(X_1, \ldots, X_n)$. Find an expression for the cdf of Y in terms of F_X.

   (c) Determine the maximum likelihood estimate of $\mathrm{E}(e^{-2X})$ if X has the cdf $F_X(x) = 1 - e^{-\lambda x}$, for $x \geq 0$ and $\lambda > 0$. *Hint: First determine $E(e^{-2X})$ as a function of λ.*

   (d) Suppose again that $F_X(x) = 1 - e^{-\lambda x}$, for $x \geq 0$ and $\lambda > 0$, and you are interested in testing

   $\mathbf{H}_o$: $\mathrm{E}(e^{-2X}) = 1/3$
   $\mathbf{H}_a$: $\mathrm{E}(e^{-2X}) = 1/2$

   Derive the procedure for carrying out a most powerful test. That is, determine the test statistic (in its simplest form), state or derive its distribution, and briefly explain how to compute a p-value for the test.

2. Suppose a random variable, X, has a Poisson distribution with mean μ. Suppose $X_1, \ldots, X_n$ is a random sample of size n.

   (a) Find the Bayes estimator of μ from the sample under squared-error loss, that is, the posterior mean of μ, with prior distribution

   $$f(\mu) = e^{-\mu}$$

   (b) Consider the estimator

   $$T = \frac{1}{n} \sum_{i=1}^{n} X_i^2$$

   Show that this estimator is never unbiased for μ.

   (c) Suppose you are interested in testing the hypotheses

   $\mathbf{H}_o$: $\mu = 10$
   $\mathbf{H}_a$: $\mu = 9$

   Derive an expression for the necessary sample size, n, to guarantee a significance level of $\alpha = 0.10$ and power of 0.90, assuming that n is large enough to permit a normal approximation to $\sum_{i=1}^{n} X_i$. You do not need to perform the computation, though your answer should be an expression that only involves numbers. You may use the fact that $\mathrm{P}(Z < 1.645) = 0.90$ for $Z \sim \mathrm{N}(0, 1)$.

Stochastic Processes (MA 583) Problems

April 2000

1. A spider hunting a fly moves between locations A and B, according to a Markov chain with transition matrix
$$\begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} ,$$
starting at location A. Meanwhile, the fly is unaware of the spider and moves between the same two locations according to a Markov chain with transition matrix
$$\begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{bmatrix} ,$$
starting at location B. If the two insects meet at the same location, the spider catches the fly and the hunt ends.

   (a) Explain how the progress of the hunt (except for knowing the location of where the hunt ends) can be described by a Markov chain with three states.
   [HINT: One state should be simply the hunt ended.]

   (b) Set up the transition probability matrix for this Markov chain.

   (c) Classify each state of the Markov chain as being recurrent or transient.

   (d) For transient states i and j, define
   $$s_{i,j} = E[\text{ Number of Time Periods Chain is in State } j \,|\, \text{Begin in state } i\,] \ .$$
   Also, let $\mathbf{T}$ be the set of transient states and let $\mathbf{P_T}$ be the matrix of transition probabilities *only* between transient states (i.e., states in the set $\mathbf{T}$).
   The values in the matrix $\mathbf{S} \equiv [s_{i,j}]$ are known to be related to the transition probabilities in $\mathbf{P_T}$ by the expression $\mathbf{S} = (\mathbf{I} - \mathbf{P_T})^{-1}$.
   Using this formula and your results from parts (a) – (c), calculate the expected duration of the hunt.

2. Let $\{N(t), t \geq 0\}$ be a Poisson process with rate λ, that is independent of a nonnegative random variable T, where T has mean μ and variance σ^2. Find

   (a) $\text{Cov}(T, N(T))$

   (b) $\text{Var}(N(T))$