Master's Exam - Spring 2010

Solution Sketches

March 18, 2010
1:00 pm – 5:00 pm

NUMBER:________________________________

A. The number of points for each problem is given.

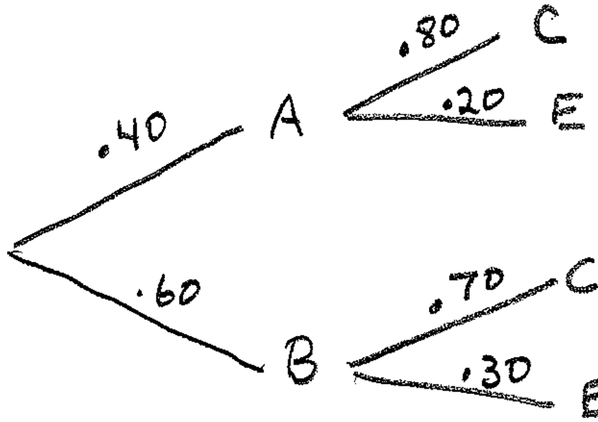
B. There are 8 problems with varying numbers of parts.

1 – 4	Probability (60 points)
5 – 8	Statistics (60 points)

C. Write your answers on the exam paper itself. If you need more room you may use the extra sheets provided. Answer as many questions as you can on each part. Tables are provided.

PROBABILITY PART. NUMBER:________________

1. (10 points) Roads A and B are the only escape routes from a state prison. Prison records show that, of the prisoners who tried to escape, 40% used road A, and 60% used road B. These records also show that 80% of those who tried to escape via A, and 70% of those who tried to escape via B were captured. Suppose that a prisoner has successfully escaped from the prison. What is the probability that he/she used Road A?



$$P(A \mid E) = \frac{P(A \cap E)}{P(E)} = \frac{(.40)(.20)}{(.40)(.20) + (.60)(.30)}$$

$$= \frac{.08}{.08 + .18} = \frac{8}{26} = .3077$$

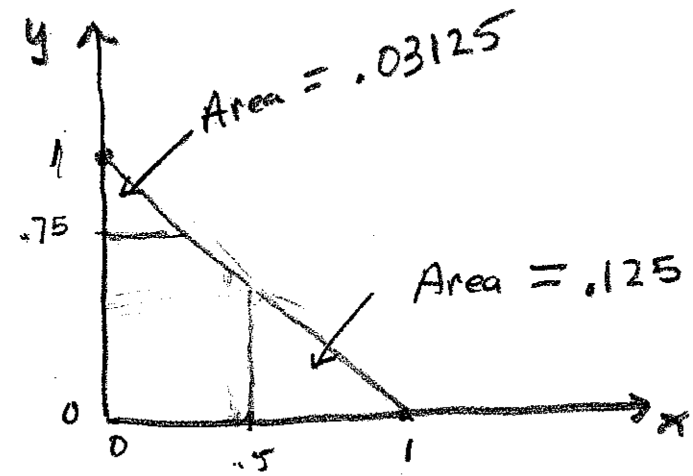
2. A continuous random vector (X, Y) has a joint probability density function given by

$$f_{X,Y}(x,y) = \begin{cases} 2 & \text{if } x > 0, y > 0, x + y < 1, \\ 0 & \text{otherwise.} \end{cases}$$

(a) (10 points) Compute $P(X \leq 0.5\,,\ Y \leq 0.75)$.

(b) (10 points) Calculate $E(X|Y = 0.5)$.

(a)



$P(X \leq 0.5, Y \leq 0.75) = 2[.5 - .03125 - .125]$

$= 0.6875$

(b) The cond dist of $X \mid Y = y$ is uniform on $[0, 1-y]$ so $E(X \mid Y = y) = \frac{1-y}{2}$

$E(X \mid Y = .5) = .25$

3. (10 points) Suppose that 10 components in a system are connected in series (this means that the system fails if at least one of component fails) and the lifetimes of the components, $T_1, \ldots, T_{10}$ are independent random variables that are exponentially distributed with parameter $\lambda = 1$. Find the expected lifetime of the system.

X = Lifetime

$$P(X > x) = P(T_1 > x, T_2 > x, \ldots, T_{10} > x)$$

$$= \prod_{i=1}^{10} P(T_i > x)$$

$$= \prod_{i=1}^{10} e^{-x} = e^{-10x}, \quad x > 0$$

$$P(X \le x) = 1 - e^{-10x}, \quad x > 0$$

$$f_X(x) = 10e^{-10x}, \quad x > 0$$

$$E(X) = \int_0^{\infty} x \, 10e^{-10x} \, dx = \frac{1}{10}$$

(use integration by parts)

4. (a) (5 points) Let X_n has the cumulative distribution function (cdf) $F_n(x)$ for $n = 1, 2, \cdots$ and X has the cdf $F(x)$. Define X_n converges in distribution to X as $n \to \infty$ and X_n converges in probability to X as $n \to \infty$, respectively.
(b) (8 points) Consider a random variable $X_n = n\, I_{(-\frac{1}{n}, \frac{1}{n})}(U)$, where U is a uniform random variable on $(-1, 1)$ and $I_A(x)$ is an indicator function such that $I_A(x) = 1$ for $x \in A$ and $I_A(x) = 0$ for $x \notin A$. Show that X_n converges in distribution to 0 as $n \to \infty$.
(c) (7 points) For X_n and X defined in (b), show that X_n converges in probability to 0 as $n \to \infty$.

(a) $X_n \xrightarrow{d} X$ if $F_n(x) \to F(x)$ as $n \to \infty$ at all points x where F is cont. where F_n is the cdf of X_n, F is the cdf of X.

$X_n \xrightarrow{P} X$ if for every $\epsilon > 0$, $P(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$

(b) $X_n = n$ with prob $1/n$
$= 0$ with prob $1 - \frac{1}{n}$

$$F_n(x) = \begin{cases} 0, & x < 0 \\ 1 - \frac{1}{n}, & 0 \le x < n \\ 1, & x \ge n \end{cases}$$

Fix $x < 0$, $F_n(x) \to 0$
For $x > 0$, $F_n(x) = 1 - \frac{1}{n}$ for all $n > x$ so $F_n(x) \to 1$

The cdf of degenerate at 0 is $F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \ge 0 \end{cases}$

so $F_n(x) \to F(x)$ for all cont. pts of F, i.e., $x \neq 0$

(c) $P(|X_n - 0| > \varepsilon) = \frac{1}{n} \to 0$ so $X_n \to 0$ in pr.

5. Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x|\theta) = \frac{\theta}{(1+x)^{\theta+1}}$ with support $0 \le x < \infty$ and with $0 < \theta < \infty$.

(a) (8 points) Find the method of moments estimator of θ.

(b) (8 points) Find the maximum likelihood estimator of θ and state explicitly the asymptotic distribution of $\sqrt{n}(\hat{\theta}_{mle} - \theta)$ when n is large, where $\hat{\theta}_{mle}$ is the maximum likelihood estimator of θ.

(a) Calculate $\int_0^\infty \frac{\theta x}{(1+x)^{\theta+1}} dx = E_\theta(X)$ using integration by parts. (Need $\theta > 1$ for the integral to be proper.

Set $\bar{X} = E_\theta(X)$ and solve for θ.

(b) $\hat{\theta}_{mle} = \dfrac{n}{\sum_{i=1}^{n} \ln(1+X_i)}$

6. Let $X_1, \dots, X_{100}$ be a random sample from $\Gamma(4, \lambda)$ distribution with density $f(x, \theta) = \frac{1}{6}\lambda^4 x^3 e^{-\lambda x}$, $0 < x < \infty$.

(a) (10 points) Consider testing $H_0 : \lambda = 1$ versus $H_1 : \lambda > 1$. Prove that the rejection region of the uniformly most powerful (UMP) test has the form $\{\sum_{i=1}^{100} X_i < c\}$.

(b) (8 points) Use normal approximation to find the critical value c such that the test in part (a) has approximate significance level $\alpha = 0.05$.

(a) Calculate N-P test for $\lambda = 1$ vs. $\lambda = \lambda_1$ where $\lambda_1 > 1$. The test does not depend on λ_1 so it is UMP. The test has rejection region of the form $\sum_1^{100} X_i < c$.

(b) Use normal approximation to distribution of $\sum_1^{100} X_i$ with $\lambda = 1$.

7. (10 points) A Professor wondered if students tended to make better scores on his test depending if the test were taken in the morning or afternoon. From a group of 14 similarly talented students, he randomly selected some to take a test in the morning and some to take it in the afternoon. The scores by groups were:

Morning	89.8	90.2	98.1	91.2	88.9	90.3	99.2
Afternoon	87.3	87.6	87.3	91.8	86.4	89.2	93.1

Does it appear that time of day makes a difference in performance on a test? Use a nonparametric method, State the model, define hypotheses, and carry out a test at level $\alpha = 0.1$.

For example, use Wilcoxon two-sample test.

8. The numbers of accidents experienced by taxi drivers in a certain city were observed for a fixed period of time with results as shown in the table for total 353 taxi drivers.

Accidents per Taxi driver	0	1	2	3	4
Frequency	254	64	21	10	4

(a) (8 points) Suppose that the data provided in the table come from a Poisson distribution with a parameter λ. What is the maximum likelihood estimator of λ for this data?
(b) (8 points) Test the null hypothesis that the data come from a Poisson distribution at level $\alpha = 0.05$ using Pearson's chi-square statistics.

(a) mle is sample average $= \frac{152}{353} = \hat{\lambda}_{mle}$

(b) Determine expected counts under fitted model Poisson $\hat{\lambda} = \frac{152}{353}$. Calculate χ^2 and compare to critical value

$5-1-1 = 3$ d.f.