Master's Exam - Spring 2011
March 17, 2011
1:00 pm - 5:00 pm

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

B. There are problems with varying numbers of parts.

Problems 1 - 4 Probability (60 points)

Problems 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.

Good Luck!

Sketchs of Solutions Gilliland

Probability Test Number 4/10/11

1. (15 points) Two boys play basketball in the following way. They take turns shooting and stop when the first basket is made. Player A goes first and has probability 0.2 of making a basket on any throw. Player B, who shoots second, has probability 0.1 of making a basket. The outcomes of the successive trials are assumed to be independent. What is the probability that Player B is the first to make a basket?

The event Player B is the first to make a basket is $C = \{ \bar{A}B, \bar{A}\bar{B}\bar{A}B, \bar{A}\bar{B}\bar{A}\bar{B}\bar{A}B, \ldots \}$ where, for example, $\bar{A}\bar{B}\bar{A}B$ denotes the outcome (Player A misses, Player B misses, Player A misses, Player B makes basket) on trials 1, 2, 3, 4. Note that

$$P(C) = (.8)(.1) + (.8)(.9)(.8)(.1) + (.8)(.9)(.8)(.9)(.8)(.1) + \cdots$$

$$= (.08)[1 + .72 + .72^2 + \cdots] \quad \text{Geometric Series}$$

$$= .08 \frac{1}{1-.72} = \frac{.08}{.28} = \frac{2}{7}.$$

2. (15 points) Let X and Y be independent random variables with probability density functions f and g, respectively, where $f(x) = 2x$ for $0 < x < 1$, and $g(x) = 1$ for $0 < x < 1$. Let $Z := \min\{X, Y\}$. Find the probability density function of Z.

Let $Z = \min\{X, Y\}$.

$$P(Z > z) = P(X > z)P(Y > z)$$
$$= (1 - z^2)(1 - z)$$

Z has cdf $F(z) = 1 - (1-z^2)(1-z)$, $0 < z < 1$

and pdf

$$f(z) = F'(z) = \frac{d}{dz}\left[1 - z^2 - z + z^3\right]$$
$$= 2z + 1 - 3z^2, \quad 0 < z < 1$$

3. (15 points) Let X be a random variable with the density $f(x) = 2x$ for $0 < x < 1$. Conditionally on $X = x$, let U be uniformly distributed on $[0, x]$. Calculate $E(X|U)$.

The joint density of (X, U) is $f(x,u) = 2x \cdot \frac{1}{x} = 2$, on $0 \leq u \leq x < 1$.

The marginal density of U is $g(u) = \int_u^1 2\,dx = 2(1-u)$, $0 \leq u < 1$.

The conditional density of $X|U=u$ is

$$h(x|u) = \frac{2}{2(1-u)}, \quad u \leq x < 1$$

that is, uniform on the interval $[u, 1]$.

Thus, $E(X|U) = \frac{1+U}{2}$.

4. (15 points) Suppose there are two boxes. The first box has 1 red ball and 39 white balls and the second box has 1 red ball and 49 white balls. The following experiment is performed. An unbiased coin is tossed. If head appears, a ball is drawn randomly from the first box, and, if tail appears, a ball is drawn randomly from the second box. This experiment is repeated 100 times with replacement. Using an approximation, calculate the approximate probability that at least 3 red balls are drawn in these 100 repetitions.

p = Prob(Red ball drawn on a single trial)

$$= \frac{1}{2}\cdot\frac{1}{40} + \frac{1}{2}\cdot\frac{1}{50} = \frac{9}{400} \ *$$

X = No. of Red balls drawn in 100 independent trials.

$$X \sim B(100, \tfrac{9}{400}).$$

Using Poisson approximation, $\lambda = np = 2.25$,

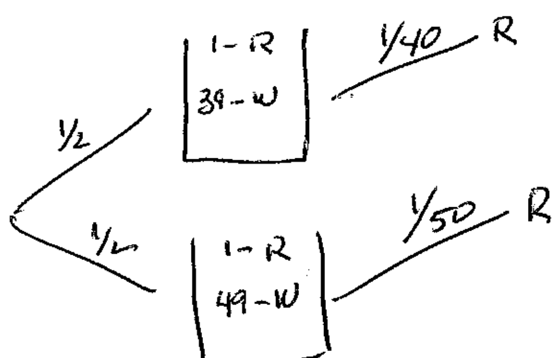
$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - e^{-2.25}\left[1 + 2.25 + \frac{2.25^2}{2}\right]$$

$$= 1 - .10540\,[5.78125]$$

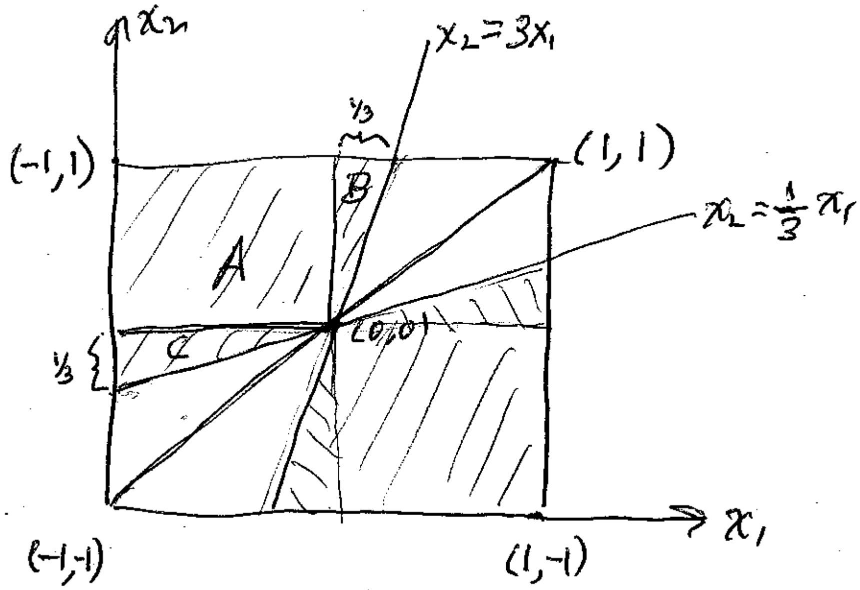
$$= 1 - .6093$$

$$= .3907$$

*



5. (15 points) Let $\{X_1, X_2\}$ be a random sample from a distribution F_θ, $-\infty < \theta < \infty$, where F_θ is the continuous uniform distribution $U[\theta - 1, \theta + 1]$. Someone proposes $\bar{X} \pm 2s/\sqrt{2}$ as a confidence interval estimator for θ. This confidence interval estimator can be written $\bar{X} \pm |X_1 - X_2|$. Determine $P_\theta(\bar{X} \pm |X_1 - X_2|\ \textit{captures}\ \theta)$. (Hint: You may make the calculation assuming $\theta = 0$.)



Consider the region $x_2 > x_1$ where $|x_1 - x_2| = x_2 - x_1$. We see that $\bar{x} \pm |x_1 x_2|$ covers 0 if $\frac{x_1+x_2}{2} - (x_2 - x_1) \le 0 \le \frac{x_1+x_2}{2} + (x_2 - x_1)$ from which $(x_2 \ge 3x_1,\ x_2 \ge \frac{1}{3}x_1) = A \cup B \cup C$. The density is uniform 1/4 on the square so

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) = 1 \cdot \frac{1}{4} + \overbrace{\frac{1}{2} \cdot 1 \cdot \frac{1}{3}}^{\text{area of } \Delta} \cdot \frac{1}{4}$$
$$+ \frac{1}{2} \cdot 1 \cdot \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{3}$$

Answer is $2 \times \frac{1}{3} = \frac{2}{3}$.

6. Let $X_1, X_2, \ldots, X_n$ be *iid* with probability density $f_\theta(x) = \theta x^{\theta-1}, 0 < x < 1, \theta > 0$.

(a) (9 points) Find the maximum likelihood estimator (*mle*) of θ.

$$\ln f_\theta(x_1, x_2, \cdots x_n) = n \ln\theta + (\theta-1)\ln y \quad \text{where } y = \prod_{i=1}^{n} x_i.$$

$$\frac{\partial \ln f_\theta}{\partial \theta} = \frac{n}{\theta} + \ln y \qquad \text{mle } \hat{\theta} = -\frac{n}{\ln y}$$

$$\frac{\partial^2 \ln f_\theta}{\partial \theta^2} = -\frac{n}{\theta^2} < 0$$

(b) (6 points) Find the asymptotic distribution of the maximum likelihood estimator.

$$\text{Information } I(\theta) = -E_\theta\left[\frac{\partial^2}{\partial \theta^2} \log f_\theta(X)\right]$$

$$\log f_\theta = \log\theta + (\theta-1)\log X$$

$$\frac{\partial \log f_\theta(X)}{\partial \theta^2} = \frac{1}{\theta} + \log X \qquad \frac{\partial^2 \log f_\theta(X)}{\partial \theta^2} = -\frac{1}{\theta^2}$$

$$\therefore I(\theta) = -E_\theta\left(-\frac{1}{\theta^2}\right) = \frac{1}{\theta^2}$$

$$\text{Asymptotically, } \hat{\theta} \sim N\left(\theta, \frac{1}{nI(\theta)}\right) = N\left(\theta, \frac{\theta^2}{n}\right)$$

7. Suppose that we are given a random sample $X_1, X_2, \ldots, X_n$ from the p.d.f. $f_\theta(x) = \theta x^{\theta-1}, 0 < x < 1, \theta > 0$, where $\theta > 0$ is an unknown parameter. The null hypothesis $H_0: \theta = 1$ is to be tested against the alternative $H_1: \theta > 1$.

(a) (8 points) Determine the family of uniformly most powerful tests.

Let $\theta > 1$. $\dfrac{f_\theta(x_1,\ldots x_n)}{f_1(x_1,\ldots x_n)} = \dfrac{\theta^n (x_1 \cdots x_n)^{\theta-1}}{1^n (x_1 \cdots x_n)^0} = \theta^n \left(\prod_1^n x_i\right)^{\theta-1}$

$$r(\theta) = \log\left(\frac{f_\theta(x_1,\ldots x_n)}{f_1(x_1,\ldots,x_n)}\right) = n \log\theta + (\theta-1)\log(y)$$

where $y = \prod_{i=1}^n x_i$. Note that $r(y)$ is a $\uparrow$ function of y.

$$\frac{dr(\theta)}{d\theta} = \frac{n}{\theta} + \log y$$

By N-P the most powerful test rejects H_0 when y is large. True for each $\theta > 1$ so a test that rejects for large y is UMP.

(b) (6 points) Assuming that the sample size n is sufficiently large, use the central limit theorem to find a uniformly most powerful test of approximate significance level $\alpha = 0.05$. (Hint: negative logarithm of a uniformly distributed random variable follows an exponential distribution.)

Find c such that $P_1[Y > c] = .05$.

$-\log Y = \sum_1^n -\log(X_i)$ where under $\theta = 1$, $X_i \sim$ Uniform $(0,1)$.

But $E_1(-\log X_i) = 1$ and $Var_1(-\log X_i) = 1$ so

$$-\log Y \sim N(n, n) \text{ approximately}$$

and, for large n, rejection region is in the form $-\log Y < d$ with $d = -n - 1.645\sqrt{n}$. (Rejection regions $Y > c$ are equivalent to rejection regions $-\log Y < d$, monotone transformation.)

(c) (6 points) For sample size $n = 1$ find the uniformly most powerful test of exact significance level $\alpha=0.05$.

With $n = 1$, $Y = X_1 \sim U(0,1)$ under $H_0: \theta = 1$. Reject for large values of Y, i.e. $X_1 > c$ where $P_1(X_1 > c) = .05$.

Take $c = .95$.

8. A student wants to test if the probabilities of getting a head are the same for tossing a quarter and tossing a dime. In each experiment, he tossed a quarter and a dime at the same time and recorded the number of heads. The procedure was repeated 500 times. The results are shown in the following table.

Number of heads	0	1	2
Frequency	149	242	109

(a) (5 points) Does the data suggest that the probability of getting a head while tossing a quarter is different from tossing a dime? Carry-out a chi-square test at level .05.

Let $P_Q(H) = p_1$, $P_D(H) = p_2$. If $p_1 = p_2$, then the number of H's observed is $B(1000, p)$ where $p = p_1 = p_2$. $\hat{p} = \dfrac{242 + 2(109)}{1000} = .46$ and the expected cell counts are

$$E_0 = 500(.54)^2 = 145.8$$
$$E_1 = 500(2(.46)(.54)) = 248.4$$
$$E_2 = 500(.46)^2 = 105.8$$

$$\chi^2 = \frac{(149-145.8)^2}{145.8} + \frac{(242-248.4)^2}{248.4} + \frac{(109-105.8)^2}{105.8}$$

$= 0.332$

Critical value is $\chi^2_{1,.95} = 3.84$

Retain H_0: $p_1 = p_2$

(b) (5 points) Does the data suggest that for both coins the probability of getting a head is .5? Carry-out a chi-square test at level .05.

If $p_1 = p_2 = .5$,

$$E_0 = 500(.5)^2 = 125$$

$$E_1 = 500(.25) = 250$$

$$E_2 = 500(.5)^2 = 125$$

$$\chi^2 = \frac{(149-125)^2}{125} + \frac{(242-250)^2}{250} + \frac{(109-125)^2}{125}$$

$$= 6.912$$

Critical value is $\chi^2_{2,.95} = 5.99$

Reject $H_0: p_1 = p_2 = .5$