Department of Statistics
University of Wisconsin, Madison
PhD Qualifying Exam Part II
Thursday, January 26, 2012
1:00-4:00pm, Room 133 SMI

- There are a total of FOUR (4) problems in this exam. Please do a total of TWO (2) problems.
- Each problem must be done in a separate exam book.
- Please turn in TWO (2) exam books.
- Please write your code name and **NOT** your real name on each exam book.

1. Suppose X and Y are non-negative random variables on a probability space $(\Omega, \mathcal{F}, P)$. Let $H(x,y)$ be a function on $[0,\infty)^2$ such that $E[|H(X,Y)|] < \infty$. Define function $\varphi(u) = u/(1+u)$ for $u \geq 0$. For integer $n = 0, 1, 2, \cdots$, $m = 2^n$, let

$$U_n = \sum_{j=1}^{\infty} \frac{j-1}{m} I\left(\frac{j-1}{m} \leq \varphi(X) < \frac{j}{m}\right), \qquad V_n = E[H(X,Y)|U_n],$$

where $I(\cdot)$ is the indictor function.

(a) Show

$$V_n = \sum_{j=1}^{m} c_{nj} I\left(\frac{j-1}{m} \leq \varphi(X) < \frac{j}{m}\right),$$

where

$$c_{nj} = \frac{E[H(X,Y)I\{(j-1)/(m-j+1) \leq X < j/(m-j)\}]}{P((j-1)/(m-j+1) \leq X < j/(m-j))}, \qquad j = 1, \cdots, m-1,$$

$$c_{nm} = \frac{E[H(X,Y)I(m-1 \leq X < \infty)]}{P(m-1 \leq X < \infty)}.$$

(b) Show $E[V_{n+k}|U_n] = V_n$ for any integers $k \geq 1$ and $n \geq 1$.

(c) Prove that there exists a random variable Z such that as $n \to \infty$, V_n converges to Z almost surely.

(d) Show that the random variable Z in (c) is almost surely equal to $E[H(X,Y)|X]$.

2. Let $X_1, X_2, \ldots$ be positive, i.i.d. random variables. Define

$$\bar{X}_n \equiv \frac{1}{n}\sum_{i=1}^{n} X_i, \quad G_n \equiv \left(\prod_{i=1}^{n} X_i\right)^{1/n}, \quad \text{and} \quad H_n \equiv \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}\right)^{-1}$$

to be the arithmetic, geometric, and harmonic means, respectively. Finally, let $\mathbf{M}_n = (\bar{X}_n, G_n, H_n)$.

(a) Prove that $\mathbf{M}_n \xrightarrow{a.s.} \mathbf{m}$. Determine $\mathbf{m}$ and state any additional assumptions used in the proof.

(b) Prove that $\sqrt{n}\,(\mathbf{M}_n - \mathbf{m})$ converges in distribution. Determine the limiting distribution and state any additional assumptions used in the proof.

(c) Prove that

$$\sqrt{n}\left(\frac{\bar{X}_n - G_n}{H_n} - c\right)$$

converges in distribution, for some constant c. Find the value of c and state any additional assumptions used in the proof.

(d) Additionally assume $\mathsf{E}X_1 = 1$, $\mathsf{Var}\,(X_1) = \sigma^2 < \infty$. Prove the following: if $k_n \to \infty$ as $n \to \infty$, then as $n \to \infty$,

$$\bar{X}_{k_n}^{n} \xrightarrow{d} \begin{cases} 1 & \text{if } n^2/k_n \to 0, \\ e^{cZ} & \text{if } n^2/k_n \to c^2, \\ \infty & \text{if } n^2/k_n \to \infty, \end{cases}$$

where $Z \sim N(0, \sigma^2)$.

3. In a school chemistry experiment, heated copper oxide is reduced to copper by passing a stream of hydrogen over it. The mass, x grams, of copper oxide and the resulting mass of copper, y grams, are noted. Three students perform this experiment four times each.

Let x_{ij} and y_{ij} denote the x and y values of student i in trial j $(i = 1, 2, 3;\ j = 1, 2, 3, 4)$. Let $\bar{x}_i$ denote the mean value of x_{ij} for student i and $\bar{x}$ the average of all the x_{ij} values, with similar definitions for $\bar{y}_i$ and $\bar{y}$. The following summary results are obtained.

$$
\begin{array}{ll}
\bar{x}_1 = 32.5,\ \bar{x}_2 = 36,\ \bar{x}_3 = 35, & \bar{y}_1 = 24.75,\ \bar{y}_2 = 27.75,\ \bar{y}_3 = 23.75 \\
\sum_{i=1}^{3}\sum_{j=1}^{4}(x_{ij} - \bar{x}_i)^2 = 93, & \sum_{i=1}^{3}\sum_{j=1}^{4}(x_{ij} - \bar{x})^2 = 119 \\
\sum_{i=1}^{3}\sum_{j=1}^{4}(y_{ij} - \bar{y}_i)^2 = 70.25, & \sum_{i=1}^{3}\sum_{j=1}^{4}(y_{ij} - \bar{y})^2 = 104.9167 \\
\sum_{i=1}^{3}\sum_{j=1}^{4}(x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i) = 74.5, & \sum_{i=1}^{3}\sum_{j=1}^{4}(x_{ij} - \bar{x})(y_{ij} - \bar{y}) = 90.5
\end{array}
$$

(a) R is used to fit the model $Y_{ij} = \beta_i + \gamma x_{ij} + \epsilon_{ij}$, $(i = 1, 2, 3;\ j = 1, 2, 3, 4)$, where the ϵ_{ij} are independent identically distributed normal with zero means. It gives the Type I ANOVA table:

	Df	Sum Sq	Mean Sq	F value
student	2	34.667	17.333	13.119
x	1	59.680	59.680	45.170
Residual	8	10.570	1.321	

Obtain the Type III ANOVA table and use it to test whether the students have equal ability to perform the experiment, at the 0.05 level of significance.

(b) Test the hypothesis $H_0 : \beta_1 = \beta_3$ at the 0.05 level of significance.

4. An experiment was carried out with five treatments in a randomized complete block design with ten blocks. Let y_{ik} denote the response for block i and treatment k.

(a) Complete the following partial ANOVA table for the model

$$y_{ik} = \eta + \beta_i + \tau_k + \epsilon_{ik} \quad (i = 1, \ldots, 10;\ k = 1, \ldots, 5)$$

where β_i and τ_k denote the block and treatment effects, respectively, with $\sum_{i=1}^{10} \beta_i = \sum_{k=1}^{5} \tau_k = 0$ and where the ϵ_{ik} are independent identically distributed normal with zero mean.

Source	d.f.	SS	MS
Blocks		135	
Treatments		100	
Residual			
Total (corr.)		307	

(b) Give an expression (in terms of y_{ik}) for the F-statistic for testing $H_0 : \tau_1 = \tau_2 = \ldots = \tau_5 = 0$ and report its numerical value.

(c) Let $\hat{\tau}_k$ denote the least-squares estimate of τ_k. Estimate the variance of $\hat{\tau}_1 - (\hat{\tau}_2 + \hat{\tau}_3)/2$.

(d) Suppose there is supplementary information in the form of a covariate x. Let x_{ik} denote the covariate value for block i and treatment k. Let $\bar{x}_{i\cdot}$, $\bar{x}_{\cdot k}$ and $\bar{x}_{\cdot\cdot}$ denote the means of x_{ik} over the dotted subscripts, with similar definitions for $\bar{y}_{i\cdot}$, $\bar{y}_{\cdot k}$ and $\bar{y}_{\cdot\cdot}$. Suppose that

$$\begin{aligned}
\sum_{i=1}^{10}\sum_{k=1}^{5}(y_{ik} - \bar{y}_{i\cdot} - \bar{y}_{\cdot k} + \bar{y}_{\cdot\cdot})(x_{ik} - \bar{x}_{i\cdot} - \bar{x}_{\cdot k} + \bar{x}_{\cdot\cdot}) &= -20 \\
\sum_{i=1}^{10}\sum_{k=1}^{5}(x_{ik} - \bar{x}_{i\cdot} - \bar{x}_{\cdot k} + \bar{x}_{\cdot\cdot})^2 &= 10.
\end{aligned}$$

Estimate γ in the ANCOVA model

$$y_{ik} = \eta + \beta_i + \tau_k + \gamma x_{ik} + \epsilon_{ik}.$$

(e) Test the hypothesis $H_0 : \gamma = 0$ at the 0.05 level.