Department of Statistics
University of Wisconsin, Madison
PhD Qualifying Exam Part II
Thursday, January 20, 2011
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 that X, Y, Y_n, $n \geq 1$, are random variables defined in a probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{A} = \sigma\{X\}$ be a σ-field generated by the random variable X. Denote by $\mathcal{B}$ the Borel field on $(-\infty, \infty)$. For $B \in \mathcal{B}$, denote by $\partial(B)$ the boundary of B.

   (a) Show that the statement:
   For all $-\infty < a_i < b_i < \infty$, $i = 1, 2$, such that $P(Y = a_2) = (Y = b_2) = 0$,

   $$\lim_{n\to\infty} P(a_1 < X \leq b_1, a_2 < Y_n \leq b_2) = P(a_1 < X \leq b_1)\, P(a_2 < Y \leq b_2)$$

   is equivalent to the statement:
   For all $A \in \mathcal{A}$ and all $B \in \mathcal{B}$ such that $P(Y \in \partial(B)) = 0$,

   $$\lim_{n\to\infty} P(A \cap \{Y_n \in B\}) = P(A)\, P(Y \in B). \tag{1}$$

   (b) Show that Statement (1) is equivalent to the statement:
   For all $B \in \mathcal{B}$ such that $P(Y \in \partial(B)) = 0$,

   $$\lim_{n\to\infty} P(Y_n \in B | \mathcal{A}) = P(Y \in B).$$

   (c) Show that Statement (1) is equivalent to the statement:

   $$\lim_{n\to\infty} E[f(Y_n)|X] = E[f(Y)]$$

   for all bounded, continuous functions f.

   (d) Show that Statement (1) is equivalent to the statement:

   $$\lim_{n\to\infty} E[W\, f(Y_n)] = E[W]\, E[f(Y)]$$

   for all $\mathcal{A}$-measurable, integrable random variables W and bounded, continuous functions f.

   (e) Suppose that Y_n are $\mathcal{A}$-measurable. Let Q be another probability on $(\Omega, \mathcal{A})$ and Q is absolutely continuous with respect to P. Show that statement (1) implies the statement:
   For all $\mathcal{A}$-measurable, integrable random variables W and bounded, continuous functions f,

   $$\lim_{n\to\infty} E_Q[W\, f(Y_n)] = E_Q[W] E[f(Y)],$$

   where E and E_Q denote expectations under P and Q, respectively.

2. Consider a positive random variable T with Laplace transform given by

$$Ee^{-\lambda T} = e^{-\sqrt{\lambda}}, \quad \lambda \geq 0.$$

Let $(T_i, i \geq 1)$ be a sequence of independent and identically distributed, random variables with the same distribution as T. Let $(a_n, n \geq 1)$ be a sequence of positive real numbers. In this problem we consider the sequence of random variables $(S_n, n \geq 1)$ defined by

$$S_n = \frac{1}{n} \sum_{i=1}^{n} a_i^2 T_i.$$

a. Give a necessary and sufficient condition on the sequence $(a_n, n \geq 1)$ which ensures that the sequence $(S_n, n \geq 1)$ converges in distribution. When this condition is satisfied, what is the limiting distribution?

For the remainder of the problem, set $a_n = \frac{1}{2\sqrt{n}}$.

b. Prove that the sequence of random variables $(S_n, n \geq 1)$ converges in distribution. What is the limiting distribution?

c. Let $k \in \mathbb{N}, k > 1$. Prove that the two-dimensional random sequence:

$$\left(S_{(k-1)n}, S_{kn}\right)$$

converges in distribution, as $n \to \infty$. Express the limiting distribution in terms of T_1 and T_2.

d. Does the sequence (S_n) converge in probability? Prove or disprove.

3. Factorial experiments are normally preferable to those in which successive treatment combinations are defined by changing only one factor at a time, as they permit estimation of interactions as well as of main effects. There may be cases, however, for example when it is very difficult to vary factor levels, where one-factor-at-a-time designs are needed. Consider fitting the model

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$$

to the following data from one such 2^3 experiment.

Run order	x_1	x_2	x_3	y
1	-1	-1	-1	519
2	1	-1	-1	558
3	1	1	-1	562
4	1	1	1	568
5	-1	1	1	560
6	-1	-1	1	567
7	-1	-1	-1	503

(a) Show that the following are estimable functions:

i. β_1

ii. β_2

iii. β_3

iv. $\beta_{12} + \beta_{13}$

v. $\beta_{23} - \beta_{12}$

vi. $\beta_{13} + \beta_{23}$.

(b) Find unbiased estimates of β_1, β_2, and β_3 that are not confounded by any linear drift in the experiment over the sequence of the seven observations.

4. A 1989 study investigated the effect of heredity and environment on intelligence. From adoption registers in France, researchers selected samples of adopted children whose biological parents and adoptive parents came from either the very highest or the very lowest socioeconomic status (SES) categories (based on years of education and occupation). They attempted to obtain samples of size 10 from each combination: (1) high adoptive SES and high biological SES, (2) high adoptive SES and low biological SES, (3) low adoptive SES and high biological SES, and (4) low SES for both sets of parents. It turned out, however, that only eight children belonged to the third combination. The 38 selected children were given intelligence quotient (IQ) tests and their scores are reported below.

SES of adoptive parents	SES of biological parents	IQ scores of adopted children									
High	High	136	99	121	133	125	131	103	115	116	117
High	Low	94	103	99	125	111	93	101	94	125	91
Low	High	98	99	91	124	100	116	113	119		
Low	Low	92	91	98	83	99	68	76	115	86	116

(a) Does the difference in mean scores for those with high and low SES biological parents depend on whether the adoptive parents were high or low SES? Test an appropriate hypothesis at the 5% level.

(b) How much is the mean IQ score affected by the SES of adoptive parents, and how much is it affected by the SES of the biological parents?

(c) At the 5% level of significance, is one of these effects significantly larger than the other?