Department of Statistics
University of Wisconsin, Madison
PhD Qualifying Exam Part I
Tuesday, January 24, 2012
12:30-4:30pm, Room 133 SMI

- There are a total of FOUR (4) problems in this exam. Please do a total of THREE (3) problems.
- Each problem must be done in a separate exam book.
- Please turn in THREE (3) exam books.
- Please write your code name and **NOT** your real name on each exam book.

1. Four statisticians I, II, III and IV play a sequence of games. For each game, the winning probabilities of I, II, III and IV are $\frac{1-\theta}{2}$, $\frac{1-\theta}{2}$, $\frac{\theta}{2}$ and $\frac{\theta}{2}$, respectively, where $0 \le \theta \le 1$. There is only one winner in each game and no tie is allowed. Assume that outcomes of games are independent of each other. For a fixed integer $r \ge 2$, the stopping rule is to terminate as soon as one of the following conditions holds: (1) I and II together win $r$ games; (2) III and IV together win $r+1$ games. At the time of termination, let $X_1$, $X_2$, $X_3$ and $X_4$ denote the numbers of games won by I, II, III and IV, respectively.

   (a) Prove or disprove the statistic $T = (X_1 + X_2, X_3 + X_4)$ is complete. Explain your reasoning clearly.

   (b) Find a UMVU estimator of $\theta$. Give details.

   (c) Now in the problem description before part (a), let $\frac{1-\sqrt{\phi}}{2}$, $\frac{1-\sqrt{\phi}}{2}$, $\frac{\sqrt{\phi}}{2}$ and $\frac{\sqrt{\phi}}{2}$, respectively, denote the winning probabilities of I, II, III and IV for each game, where $0 \le \phi \le 1$. Find a UMVU estimator of $\phi$. Give details.

2. Let $(X_0, X_1, X_2)$ be a 3-dimensional random vector having a multinomial distribution with size $n$ and probability vector $(p^2, 2p(1-p), (1-p)^2)$, where $p \in (0,1)$ is unknown. Consider the problem of testing

$$H_0 : p = p_0 \quad \text{versus} \quad H_1 : p \neq p_0$$

where $p_0$ is a known constant in (0,1).

(a) Derive the likelihood ratio test statistic $\lambda_n$.

(b) Derive Wald's asymptotic test statistic $W_n$ and Rao's score test statistic $R_n$.

(c) Assume that $H_0$ holds. Show that, as $n \to \infty$, $W_n$ converges in distribution to the chi-square distribution with one degree of freedom and $W_n/R_n$ converges in probability to 1.

(d) Assume that $H_0$ holds. Show that $-2\log\lambda_n/W_n$ converges in probability to 1 as $n \to \infty$.

(e) Show that the likelihood ratio test is equivalent to a uniformly most powerful unbiased test.

3. A system of interest involves three random variables, $X$, $Y$, and $Z$, where $X$ and $Z$ are independent, with $X \sim \text{Normal}(\mu_1, 1)$ and $Z \sim \text{Normal}(\mu_2, 1)$, and where, given $X$ and $Z$, $Y$ is normal with mean $Z$ and known variance $\sigma^2 > 0$. We have in mind a test of $H_0 : \mu_1 = \mu_2$ against a one-sided alternative. For the statistic $X - Z$ we could reject if the p-value $p(x, z) = 1 - \Phi[(x - z)/\sqrt{2}]$ is sufficiently small, where $\Phi$ is the standard normal cumulative distribution. However, $Z$ is unobserved. Rather than base a test directly on $X - Y$, as would be natural, we consider instead the statistic $T(x, y) = E\left[p(x, Z)|Y = y\right]$.

   (a) What is the distribution of $Z$ given $Y = y$?

   (b) Show that under the null hypothesis, $T(X, Y)$ has probability density

$$f(t) = \frac{a\phi\left[a\Phi^{-1}(1-t)\right]}{\phi\left[\Phi^{-1}(1-t)\right]}$$

   for a constant $a > 1$ and for $t \in (0, 1)$, where $\Phi^{-1}$ and $\phi$ are the quantile function and density function, respectively, of a standard normal distribution. Express $a$ in terms of $\sigma^2$.

   (c) If we use $T(X, Y)$ as our p-value and (incorrectly) treat it as uniformly distributed on the null, how is the significance level of our test affected?

4. Consider a one-way ANOVA with $p$ groups and $m$ observations per group. Let $n = mp$ be the number of total observations. Here, the model can be written as

$$Y_{ij} = \beta_j + \epsilon_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, p,$$

where $\epsilon_{ij}$'s are i.i.d. with zero mean and finite variance. The distribution of $\epsilon_{ij}$ is unknown. Let $\hat{\beta}_j$ be the solution to the following estimating equation:

$$0 = \sum_{i=1}^{m} \Psi(y_{ij} - \hat{\beta}_j),\ j = 1, \ldots, p, \tag{1}$$

where $\Psi$ is a given continuously differentiable function. Let $\beta_j^*$ be the true value of $\beta_j$.

Assume

(1). $E(\Psi(\epsilon_{ij})) = 0$ and $E(\Psi'(\epsilon_{ij})) = d \neq 0$.

(2). $p \to \infty, n \to \infty, p(\log p)/n \to 0$.

(a) Show that there are solutions $\{\hat{\beta}_j\}$ of (1) and there exist a constant $B > 0$ and a sequence $\{\delta_n > 0\}$ such that for $j = 1, \ldots, p$, and $0 < u \leq \delta_n$,

$$P\{|\hat{\beta}_j - \beta_j^*| \geq u\} \leq 2\exp\{-Bu^2 n/p\}. \tag{2}$$

(b) Show that

$$\max_j\{|\hat{\beta}_j - \beta_j^*|\} \xrightarrow{P} 0 \tag{3}$$