Department of Statistics
University of Wisconsin, Madison
PhD Qualifying Exam Part I
Monday, August 27, 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. Suppose that $Y_1, \ldots, Y_n$ are independent observations following the distribution $N(\mu, \sigma^2)$, where $\sigma^2 \in (0, \infty)$. Define

$$T_n = \frac{1}{n}\sum_{i=1}^{n} \frac{(Y_i - \mu)^3}{\sigma^3}, \qquad \hat{T}_n = \frac{1}{n}\sum_{i=1}^{n} \frac{(Y_i - \hat{\mu})^3}{\hat{\sigma}^3}$$

where $(\hat{\mu}, \hat{\sigma}^2)$ are the maximum likelihood estimates of (μ, σ^2).

(a) Derive a non-degenerate limiting distribution of $\sqrt{n}(T_n - a_n)$ for some suitable a_n as $n \to \infty$.

(b) Derive a non-degenerate limiting distribution of $\sqrt{n}(\hat{T}_n - b_n)$ for some suitable b_n as $n \to \infty$.

(c) Compare the variances of the limiting distributions for T_n in part (a) and $\hat{T}_n$ in part (b).

Hint: For integers $k \geq 1$ *and* $Z \sim N(0, 1)$, $E(Z^{2k}) = (2k-1) \times (2k-3) \times \cdots \times 5 \times 3 \times 1$.

2. A random variable Y is modeled as normally distributed with mean x and variance σ^2 conditional upon the event that a second random variable X satisfies $X = x$. For $\theta \in (0, 1)$, the density of X is

$$f(x|\theta) = \begin{cases} 2(1-\theta)\phi(x) & \text{if } x < 0 \\ 2\theta\phi(x) & \text{if } x \geq 0 \end{cases}$$

where $\phi(x) = (1/\sqrt{2\pi})\exp\{-x^2/2\}$ is the probability density function of a standard normal random variable.

(a) Derive expressions for the marginal probability density of Y and for the conditional density of Y given $X \geq 0$, both in terms of $\phi(\cdot)$ and the standard normal cumulative distribution function $\Phi(\cdot)$.

(b) A random sample $\{(X_i, Y_i) : i = 1, 2, \ldots, n\}$ is available where all pairs have the same distribution as in Part (a). Construct the likelihood function for $(\theta. \sigma^2)$ and derive the maximum likelihood estimator.

3. Consider a series of n trials, $X_1, \ldots, X_n$, with possible outcomes 0 and 1, where $n \geq 4$. The first three trials are independent with $P(X_i = 0) = 1/2$, $i = 1, 2, 3$. For $i = 4, \ldots, n$, the result of trial X_i depends on those of the previous trials through the following conditional probability,

$$P(X_i = x_i | X_1 = x_1, \ldots, X_{i-1} = x_{i-1}) = \begin{cases} p, & x_i \neq x_{i-1}, \\ 1 - p, & x_i = x_{i-1}, \end{cases}$$

where $0 < p < 1$, and $x_1, \cdots, x_i$ take values 0 and 1.

Let

$$W = \sum_{i=4}^{n} |X_i - X_{i-1}|.$$

(a) Prove or disprove that W is a complete statistic for p. Explain your reasoning clearly.

(b) For $n = 5$, find a UMVU estimator of p. Provide details to justify your answer.

(c) For $p = 1/2$, compute $E(W^2)$ in details.

4. Consider generalized linear models. Assume that we have n observations: $(Y_i, \boldsymbol{x}_i)$, $i = 1, \ldots, n$, where Y_i is the random response and $\boldsymbol{x}_i = (x_{i1}, \cdots, x_{ip})^T$ is a vector of p fixed covariates for the ith observation. Denote by $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)^T$ an unknown p-dimensional vector of regression coefficients. Let $\theta_i = \sum_{j=1}^p x_{ij}\beta_j, \mu_i = E(Y_i)$ and $\sigma_i^2 = Var(Y_i) \in (0, \infty)$. Assume that the density of Y_i belongs to the following exponential family:

$$f(y_i; \theta_i) = \exp\{\theta_i y_i - b(\theta_i)\},$$

where $b(\cdot)$ has a continuous second derivative, $b'(\theta_i) = \mu_i$ and $b''(\theta_i) = \sigma_i^2$. Suppose that all θ_i's are contained in a compact subset of a space Θ.

(a) Write down the likelihood function. Let $\hat{\boldsymbol{\beta}}$ be the MLE of $\boldsymbol{\beta}$. State some suitable conditions and derive a non-degenerate limiting distribution of $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ under the conditions as $n \to \infty$.

(b) For any $M > 0$, let a_{ni}, $i = 1, \cdots, n$, be real numbers such that $\sum_{i=1}^n a_{ni}^2 \sigma_i^2 = 1, \sum_{i=1}^n a_{ni}^2 \leq M$ and m_n's are real numbers satisfying $\max_{1 \leq i \leq n} a_{ni}^2 m_n = o(1)$.

(i) Derive the moment generating function of $\sum_{i=1}^n a_{ni}(Y_i - \mu_i)$.

(ii) For any $\epsilon > 0$, prove that

$$P\left(\sum_{i=1}^n a_{ni}(Y_i - \mu_i) > \sqrt{2m_n}\right) \leq \exp\{-m_n(1 - \epsilon)\}.$$