Dept Copy

# Department of Statistics
# University of Wisconsin, Madison
# PhD Qualifying Exam Part I
## January 19, 2010
## 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 $X_1, \ldots, X_n$ are i.i.d. random variables from a uniform distribution on the interval $(\sigma, \tau)$.

   (a) Find the asymptotic limit of $\frac{1}{n^2}\sum_{i=1}^n e^{1/U_i}$ in probability as $n \to \infty$, where $U_i = (X_i - \sigma)/(\tau - \sigma)$.

   (b) Find the maximum likelihood estimators of parameters $(\sigma, \tau)$.

   (c) Find the asymptotic joint distribution of $(X_{(1)}, X_{(n)})$, where $X_{(1)} = \min(X_1, \cdots, X_n)$ and $X_{(n)} = \max(X_1, \cdots, X_n)$.

   (d) We want to test the hypothesis $H_0 : \sigma = -\tau$. Find the maximum likelihood estimators of parameters $(\sigma, \tau)$ under $H_0$.

   (e) Find the likelihood ratio statistic $\Lambda_n$ for testing $H_0$.

   (f) Find the limit distribution of $\Lambda_n$ under $H_0$.

2. *Background.* Let $\mathbf{x} = (x_1, x_2, \ldots, x_p)'$ be a random $p$-dimensional column vector with known covariance matrix $\Sigma$. In *principal component analysis*, $\mathbf{x}$ is linearly transformed to a $p$-vector $\mathbf{z} = \mathbf{A}'\mathbf{x}$, where $\mathbf{A}$ is a constant $p \times p$ matrix, such that

   (i) the column vectors of $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p)$ have length 1:

$$\mathbf{a}_j'\mathbf{a}_j = 1, \quad j = 1, 2 \ldots, p,$$

   (ii) the elements of $\mathbf{z} = (z_1, z_2, \ldots, z_p)'$ are uncorrelated, i.e.,

$$\text{cov}(z_j, z_k) = 0 \text{ for } j \neq k,$$

   (iii) the variance of $z_1$ is maximum over all linear combinations $\mathbf{b}'\mathbf{x}$, subject to $\mathbf{b}'\mathbf{b} = 1$; the variance of $z_2$ is next largest, and so on.

   It is well known that the columns of $\mathbf{A}$ are the solutions (called eigenvectors) $\mathbf{a}$ of the equations

$$(\Sigma - \lambda\mathbf{I})\mathbf{a} = 0, \quad \mathbf{a}'\mathbf{a} = 1.$$

   *Questions.*

   (a) Let $\lambda_j$ be the value of $\lambda$ associated with $\mathbf{a}_j$. Show that

$$\text{var}(z_j) = \lambda_j \text{ for } j = 1, 2, \ldots, p.$$

   (b) Show that $\mathbf{a}_j'\mathbf{a}_k = 0$ for $j \neq k$.

   (c) Show that $\sum_{j=1}^{p} \text{var}(x_j) = \sum_{j=1}^{p} \text{var}(z_j)$.

   (d) Principal components are sometimes used in regression where the unknown $\Sigma$ is estimated by the sample covariance matrix $\widehat{\Sigma}$ of the $\mathbf{x}$ variables. Suppose that in one such application with $p = 2$, a simple linear least-squares regression model fitted to each principal component separately gave the results:

$$\begin{aligned} \widehat{y} &= 3 + 2z_1 \\ \widehat{y} &= 1 + z_2. \end{aligned}$$

   Explain, with full justification, what you can or cannot say about the least-squares estimates of $(\beta_0, \beta_1, \beta_2)$ in the full model

$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \epsilon.$$

3. Consider the following linear mixed model:

$$y_{ij} = \boldsymbol{x}_{ij}\boldsymbol{\beta} + Z_{ij}\boldsymbol{b}_i + \epsilon_{ij},\ i = 1, \ldots, n;\ j = 1, \ldots, m,$$

where $\boldsymbol{x}_{ij}$ is a $m \times p$ fixed effect design matrix for subject $i$, $z_{ij}$ is a $m \times q$ random effect design matrix for subject $i$, $\boldsymbol{\beta}$ is a $p$-length vector of regression coefficients, $\boldsymbol{b}_i = (b_{i1}, \ldots, b_{iq}) \overset{iid}{\sim} N(0, \text{diag}\{\tau_1^2, \ldots, \tau_q^2\})$, $i = 1, \ldots, n$, are random effects, and $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, are random errors. $\epsilon_{ij}$'s and $b_i$ 's are independent.

**For Questions (a)-(b),** assume $q = 1$.

Consider the following testing problem:

$$H_0 : \tau_1^2 = 0\ vs.\ H_1 : \tau_1^2 > 0 \tag{1}$$

(a) Show the asymptotic null distribution of likelihood ratio test for the testing problem (1).

(b) Define a one-sided score test for the testing problem (1), and show its asymptotic null distribution.

**For Question (c),** assume $q > 1$.

Consider the following testing problem:

$$H_0 : \tau_1^2 = 0, \ldots, \tau_q^2 = 0\ vs.\ H_1 : \tau_1^2 > 0, \ldots, \tau_q^2 > 0 \tag{2}$$

(c) Show the asymptotic null distribution of likelihood ratio test for the testing problem (2).

4. Suppose $(Z_1, T_1), \cdots, (Z_n, T_n)$ are i.i.d. as $(Z, T)$, where $T > 0$ is a survival time and $Z \in \mathbb{R}$ with $0 < E[Z^2] < \infty$ and $P(Z = 0) = 0$ is a predictor random variable. Let $H(z, t; \beta)$, $\beta \in \mathbb{R}$, denote the distribution of $(Z, T)$ and assume that the marginal distribution of $Z$ does not involve $\beta$. Consider the estimating equation

$$\frac{1}{n}\sum_{i=1}^{n} \psi_0(z_i, t_i; \beta) = 0,$$

where $\psi_0(z, t; \beta) = \frac{\dot{r}(z;\beta)}{r(z;\beta)} - \dot{r}(z;\beta)t$, where $r(.;\beta) > 0$ is a known 1-1 function of $\beta$ for almost all $z$ with derivative $\dot{r}(z;\beta) = \partial r(z;\beta)/\partial\beta$.

Suppose $H$ is such that $T$ given $Z = z$ has the Weibull distribution with c.d.f.

$$1 - \exp\{-r(z;\beta_0)t^\alpha\} \quad \alpha > 0.$$

*Hint:* You may use the fact that if $W$ has the Weibull distribution with c.d.f. $1 - \exp(-w^\alpha/\theta)$, then $E[W] = \theta^{1/\alpha}c_0(\alpha)$ and $\text{var}(W) = \theta^{2/\alpha}c(\alpha)$, where $c_0(\alpha) = \Gamma(\alpha^{-1} + 1)$ and $c(\alpha) = \Gamma(2\alpha^{-1} + 1) - [\Gamma(\alpha^{-1} + 1)]^2$.

(a) For what values of $\alpha$, does $E_H[\psi_0(Z, T; b)] = 0$ have the solution $b = \beta_0$? Show your work.

(b) Let $r(z;\beta) = \exp(-\beta z)$ and let $\alpha$ be the answer(s) to part (a). Is the solution $b = \beta_0$ in part (a) unique? Justify your answer.

(c) Let $D(b) = \frac{1}{n}\sum_{i=1}^{n} \psi_0(Z_i, T_i; b)$. Suppose $r(z, \beta) = \exp(-\beta z)$. Show that $D(b) = 0$ has a unique solution with probability one.

(d) Let $\beta_1$ be the solution to $E_H[\psi_0(Z, T; b)] = 0$ where $H$ is the distribution for which $(T \mid z)$ has the Weibull distribution given above. Let $r(z, \beta) = \exp(-\beta z)$ and let $\widehat{\beta}$ be the solution to $D(b) = 0$ as detailed in (c). Derive the asymptotic distribution of $\sqrt{n}(\widehat{\beta} - \beta_1)$ assuming that $(Z, T)$ has the distribution $H$. Simplify your answer as much as possible using the given distributional assumptions.

(e) Suppose $P(Z = 1) = P(Z = -1) = 1/2$. Give the asymptotic distribution in (d) for this case. What is the value of the variance in the asymptotic distribution of $\sqrt{n}(\widehat{\beta} - \beta_1)$ when $\alpha = 1$?

# Department of Statistics
# University of Wisconsin, Madison
# PhD Qualifying Exam Part II
# January 21, 2010
# 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.