STATISTICS Ph. D. COMPREHENSIVE EXAM
January, 2016

General Instructions: Write your Code Word on your answer sheets. Do not put your name or id number on any of your answer sheets. All solutions should be rigorously explained. Show all work. If you cannot solve a problem, at least explain what the problem is about and what approach may lead to its solution.

1. Consider predicting a random variable y on the real line using a random vector x. We want the predictor $f(x)$ that minimizes the expected weighted squared error loss $\mathrm{E}\{L[y, f(x)]\} = \mathrm{E}\{w(y)[y - f(x)]^2\}$ where $w(y) > 0$ for all y is a known function and the expectation is over both y and x. Show that the optimal predictor is $f(x) = \mathrm{E}[w(y)y|x]/\mathrm{E}[w(y)|x]$.

2. Let $X_1, X_2, \ldots, X_n$ be a random sample from a $Gamma(\alpha, \beta)$ population. Assume α is known and β is unknown. Consider testing $H_0 : \beta = \beta_0$.

   (a) What is the MLE of β?

   (b) Derive a Wald statistic for testing H_0, using the MLE in both the numerator and denominator of the statistic.

   (c) Repeat the previous part but using the sample standard deviation in the standard error.

   (d) Derive a score statistic for testing H_0.

3. Consider the family of distributions on $(0, \infty)$ with densities

$$f(x; \theta) \propto exp(-\theta x^a)$$

   $\theta > 0$, where $a > 1$ is fixed.

   (a) Let $Y \sim \exp(\theta)$. Find the density $f(x; \theta)$ for $x \equiv y^{1/a}$ for $a > 1$ fixed.

   (b) Calculate the maximum likelihood estimator $\hat{\theta}_n$ based on an independent sample of size n from this distribution.

   (c) Find the the asymptotic distribution of the maximum likelihood estimator and identify completely this distribution (i.e. identify all its parameters).

4. A married man who frequently talks on his mobile is well known to have conversations the lengths of which are independent, identically distributed random variables, distributed as exponential with mean $1/\lambda$. His wife has long been irritated by this behavior and knows, from infinitely many observations, the exact value of λ. In an argument with her husband, the woman produces $t_1, \ldots, t_n$, the times of n telephone conversations, to prove how excessive her husband is. He suspects that she has randomly chosen the observations, conditional on their all being longer than the expected length of a conversation. Assuming he is right in his suspicion, the husband want to use the data he has been given to infer the value of λ.

   (a) Find the (conditional) density of an observation.

(b) What is the minimal sufficient statistic he should use?

(c) Is the statistic from the previous item complete?

(d) Find the maximum likelihood estimator of λ and therefore, of the mean time the man spends of his mobile.

(e) Describe how you might find a $(1-\alpha)$ confidence interval on λ.

5. Consider the linear model

$$Y = X\beta + Z\gamma + e, \qquad E(e) = 0, \qquad \mathrm{Cov}(e) = V.$$

The matrix

$$A \equiv X[X'V^{-1}X]^{-}X'V^{-1}$$

is the oblique projection operator onto $C(X)$ along $C[V^{-1}X]^{\perp}$, i.e., if $v \in C(X)$, $Av = v$, and if $v \in C[V^{-1}X]^{\perp}$, $Av = 0$. Note that

$$(I-A)'V^{-1}(I-A) = (I-A)'V^{-1} = V^{-1}(I-A).$$

Let $\mathcal{A}_{X,Z}$ be the oblique projection operator onto $C(X,Z)$ along $C[V^{-1}(X,Z)]^{\perp}$. Assume that all matrix inverses exist.

(a) Show that

$$\mathcal{A}_{X,Z} = A + (I-A)Z[Z'(I-A)'V^{-1}(I-A)Z]^{-1}Z'(I-A)'V^{-1}.$$

(b) Show that

$$\hat{\gamma} = [Z'(I-A)'V^{-1}(I-A)Z]^{-1}Z'(I-A)'V^{-1}(I-A)Y$$
$$X\hat{\beta} = A(Y - Z\hat{\gamma})$$

provide generalized least squares estimates (BLUEs) for the linear model.