STT 872, 867-868 Spring Preliminary Examination
Wednesday, January 4, 2017
12:30 - 5:30 pm

NOTE: This examination is closed book. Every statement you make must be substantiated. You may do this either by quoting a theorem/result and verifying its applicability or by proving things directly. You may use one part of a problem to solve the other part, even if you are unable to solve the part being used. A complete and clearly written solution of a problem will get a more favorable review than a partial solution.

You must start solution of each problem on the given page. Be sure to put the number assigned to you on the right corner top of every page of your solution.

1. Let $X_1, ..., X_n$ be i.i.d. from a Pareto distribution with the parameters $\lambda > 2$ and $x_0 > 0$, i.e. from a density

$$f(x; \lambda, x_0) = \frac{\lambda x_0^\lambda}{x^{\lambda+1}}, x \geq x_0,$$

and 0 elsewhere.

(a) Suppose x_0 is known. Find a complete and sufficient statistics for λ based on $X_1, ..., X_n$.

(b) Suppose x_0 is known. Let $g(\lambda) = \lambda^{-2}$. Formulate the two methods of finding UMVUE and use one of them to find it for $g(\lambda)$.

(c) Suppose x_0 is known. Find the MLE of $g(\lambda)$. Using delta method find its asymptotical variance, $n > 1$.

(d) Suppose x_0 is known. Let λ have a distribution with the density $f(\lambda) = \theta e^{-\theta\lambda}$, $\lambda > 0$. Determine the Bayes estimator of λ under the loss $\lambda(\delta(X_1, ..., X_n) - \lambda)^2$. Hint: first prove that the estimator has the form $E(\lambda^2|X)/E(\lambda|X)$.

(e) Suppose x_0 is known. Calculate the risk of the Bayes estimator in (d). Find conditions on θ that would make the estimator in (d) minimax. Are they satisfied, $n > 1$? Hint: first prove that the posterior risk has the form $E(\lambda^3|X) - (E(\lambda^2|X))^2/E(\lambda|X)$.

(f) Suppose x_0 is known. Check whether the Bayes estimator of λ in (d) dominates the MLE under the loss $\lambda(\delta(X_1, ..., X_n) - \lambda)^2$.

(g) Suppose x_0 is known. Derive a UMP unbiased test of size $\alpha \in (0, 1)$ for testing $H : \lambda \in [3, 4]$ vs $K : \lambda \in (2, 3) \cup (4, +\infty)$ in full detail.

(h) Find the m-th moment of X_1, $m \geq 1$.

(i) Suppose λ is known. Show that $\min(X_1, ..., X_n)\frac{n-3/\lambda}{n-2/\lambda}$ is the MRE estimator of x_0 under the loss $(\delta(X_1, ..., X_n)/x_0 - 1)^2, n > 1$.

(j) Is the estimator in (i) consistent in probability for x_0?

(k) Derive a UMP test of size $\alpha \in (0, 1)$ for testing $H : x_0 = 5$ vs $K : x_0 > 5$ in full detail.

(l) Derive the power of the test in (k).

2. Consider the linear regression model $\underline{y} = X\underline{\beta} + \underline{\varepsilon}$ where X is an $n \times p$ full rank design matrix. Suppose we partition the model as $\underline{y} = X_1\underline{\beta}_1 + X_2\underline{\beta}_2 + \underline{\varepsilon}$ where X_1 is $n \times p_1$ and X_2 is $n \times p_2$. Moreover, $\underline{\varepsilon}$ is normally distributed with mean 0 and variance $\sigma^2 I_n$.

(a) Show that the least squares estimate of $\underline{\beta}_2$ can be written as $\hat{\underline{\beta}}_2 = C'\underline{y}$ where $C = X_1C_{12} + X_2C_{22}$, C_{12} and C_{22} are matrices of orders $p_1 \times p_2$ and $p_2 \times p_2$, respectively, that depend on X_1 and X_2, and C_{22} is positive definite.

(b) Find a linear combination of $\underline{y}$ (i.e., $Q\underline{y}$) so that $Q\underline{y}$ follows the Gauss Markov model with only coefficients $\underline{\beta}_2$ in the model.

(c) Show that $[I_n - X_1(X_1'X_1)^{-1}X_1']C$ is of rank p_2, where C is defined in part (a).

(d) Let SS_{E_1} be the residual sum of squares for the regression of $\underline{y}$ on X_1 alone. Using part (c) to show that SS_{E_1} and $\hat{\beta}_2$ are not independent.

(e) Construct a $1 - \alpha$ level confidence region for $\underline{\beta}_2$.

(f) Assume $\underline{\beta}_2 = (\beta_{21}, \cdots, \beta_{2p_2})^T$. Using Scheffe's method to find $1 - \alpha$ simultaneous confidence intervals for β_{2j} for $j = 1, \cdots, p_2$.

3. Consider the case of a linear model $\mathrm{E}[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$ and $\mathrm{var}(\mathbf{y}) = \mathbf{V}$. Assume that $\mathbf{X}$ is of full column rank, and use the notation of $\hat{\boldsymbol{\beta}}_{\mathbf{W}^{-1}}$ to denote the weighted least squares estimator $(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}$. Thus $\hat{\boldsymbol{\beta}}_{\mathbf{I}}$ denote the ordinary least squares estimator and $\hat{\boldsymbol{\beta}}_{\mathbf{V}}$ denotes the MLE with a known variance-covariance matrix $\mathbf{V} = \mathrm{var}(\mathbf{y})$, i.e., $\hat{\boldsymbol{\beta}}_{\mathbf{V}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$.

(a) Show that $\hat{\boldsymbol{\beta}}_{\mathbf{I}}$ is unbiased no matter what the value of $\mathbf{V}$.

(b) Calculate the variance of $\hat{\boldsymbol{\beta}}_{\mathbf{I}}$.

(c) Show that $\hat{\boldsymbol{\beta}}_{\mathbf{V}}$ is unbiased.

(d) Calculate the variance of $\hat{\boldsymbol{\beta}}_{\mathbf{V}}$.

(e) Is $\hat{\boldsymbol{\beta}}_{\mathbf{V}}$ smaller than $\hat{\boldsymbol{\beta}}_{\mathbf{I}}$? Then why?

(f) Consider the simple situation where the mean of all the observations is μ and the data come in m equicorrelated clusters with correlation $\rho > 0$. Further assume that half the clusters are of size n and the other half are of size λn. We thus have

$$\mathbf{y} \sim (\mathbf{1}\mu, \mathbf{V}),$$

where

$$\mathbf{V} = \begin{bmatrix} \mathbf{I}_{m/2} \otimes \mathbf{V}_{0,n} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m/2} \otimes \mathbf{V}_{0,\lambda n} \end{bmatrix}$$

and $\mathbf{V}_{0,n} = \sigma^2[(1-\rho)\mathbf{I}_n + \rho\mathbf{J}_n]$.

Derive the relative efficiency of $\hat{\boldsymbol{\beta}}_{\mathbf{I}}$ and $\hat{\boldsymbol{\beta}}_{\mathbf{V}}$.