Mathematical Statistics Qualifying Exam.

September 8, 2003

(1) For a random vector X $(k \times 1)$ with mean vector θ and covariance matrix Σ, show that $E(X^t A X) = tr[A\,E(XX^t)]$ in which A $(k \times 1)$ is a matrix of constants. Then show that $E(X^t A X) = tr[A\Sigma] + \theta^t A \theta$.

(2) Let $X_1, X_2, \ldots, X_n, \ldots$ be an i.i.d. sequence of random variables defined on a probability space in which the first four moments are finite. Now define

$$S_n^2 = \left(\frac{1}{n-1}\right)\sum_{i=1}^{n}(X_i - \overline{X})^2 \quad \text{and} \quad Q_n = \left[\frac{1}{2(n-1)}\right]\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2 .$$

a. Write quadratic forms for each of the statistics S_n^2 and Q_n.

b. Show that S_n^2 and Q_n are both unbiased estimators of the population variance.

c. Write out a complete statement of a theorem which establishes the consistency of each estimator. (Be certain to explain why its conditions are satisfied in this particular application.)

(3) Assume a sequence of random variables as put forth in Exercise (2). Additionally, assume the mean is zero. Also, set

$$u_n = \frac{\sum_{i=1}^{n} X_i}{n} \quad \text{and} \quad v_n = \frac{\sum_{i=1}^{n} X_i^2}{n}.$$

a. Write out the limiting distribution of $w_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$ (2×1).

b. Show that $\sqrt{n}\,(v_n - \sigma^2) \xrightarrow{L} N[0,\, \mathrm{var}(X^2)]$ where σ^2 is the population variance.

c. Find the asymptotic distribution of the coefficient of variation (= sample standard deviation divided by sample mean).

(4) a) State and prove the Lehmann-Scheffé theorem on UMVUE's in the presence of a complete sufficient statistic.

b) Write down the long form of the bivariate normal density in terms of two mean parameters, two variances, and the covariance.

c) Give with explanation a complete sufficient statistic from a random sample $(X_1, Y_1), \dots, (X_n, Y_n)$ from a bivariate normal.

d) Prove that $\frac{1}{n-1} \Sigma (X_i - \bar{X})(Y_i - \bar{Y})$ is unbiased for the covariance, and hence UMVU

e) This statistic is also UMVU for the "large" non-parametric problem assuming only $EX_1^2, EY_1^2 < \infty$. (assume that.) Why does it not follow that it is UMVU for any family nested between the normal family and the "large" nonparametric family?

f) What do you know about UMVU estimation of the correlation coefficient (in the normal problem)? How about efficient asymptotically unbiased estimation of the correlation coefficient?

(5) $X_1, \ldots, X_n \sim iid$ geom(p) $\left(f(x) = p^x(1-p)\,;\; x = 0, 1, 2, \ldots\right)$

Find the MLE of p and do a complete asymptotic analysis of its sampling distribution.