Math 281 Qualifying Exam. May 29, 2003

(1) Let {u, v, w} be a set of orthonormal vectors in R^d, and let Q (d×d) be given by $Q = uu^T + vv^T + ww^T$. If $X \sim N_d(0, Q)$, then show that $X^T X \sim \chi^2_3$.

(2) Assume that $g: R^d \to R^k$ where $g'(x)$ is continuous in a neighborhood of $\theta \in R^d$. Also, let X_n (d×1) be a sequence of random vectors for which

$$\sqrt{n}(X_n - \theta) \xrightarrow{L} X.$$

What is the limiting distribution of $\sqrt{n}[g(X_n) - g(\theta)]$? Write out the proof of the 1-dimensional version of this theorem.

(3) For the purpose of estimating p^2, suppose we have the choice between

(a) n binomial trials with probability p^2 of success, or

(b) n binomial trials with probability p of success,

and that as estimators of p^2 in the two cases, we would use respectively X/n and $(Y/n)^2$, where X and Y denote the number of successes in cases (a) and (b), respectively. For large n, what are the circumstances under which $(Y/n)^2$ would be preferred to X/n ?

(4) (a) Define a sufficient statistic (the definition needs a context).

(b) State and prove the Rao-Blackwell theorem.

$X_1, \ldots, X_n \sim_{iid} \text{Pois}(\lambda)$; $\lambda > 0$:

(c) Compute with explanation $E\left[\sum (X_i - \bar{X})^2 \,\middle|\, \sum X_i \right]$

(d) Find a positive lower bound for $\text{var}\left[\sum (X_i - \bar{X})^2 \right]$ in terms of n, λ.

(e) $Y_1, Y_2, \ldots \sim_{iid} \text{Pois}(\lambda)$; $Y^* = Y_{\min\{i : Y_i \geq 2\}}$

Find the UMVUE of λ based on the single datum Y^*.

(5) (a) Define minimaxity of an estimator (context needed).

(b) If $X_1, \ldots, X_n \sim_{iid} \mathcal{N}(\mu, \sigma^2)$; $\mu \in [0,1]$, $\sigma^2 \leq 1$,

Show that $\bar{X}$ is not minimax.

(c) Discuss the situation in (b) where instead of $\mu \in [0,1]$, $\mu \in [0, \infty)$. Pertinent theorems may be mentioned without proof, and you may touch on uniqueness and admissibility issues as well.

(6) $\binom{X_i}{Y_i}$, $i = 1, \ldots, n$ are a random sample from a bivariate normal distribution with correlation ρ. Three statistics are in common usage for assessing (monotone) dependency:

Pearson's r, Spearman's Rho, and Kendall's Tau. (They're all just statistics – don't confuse the written name Rho with the true parameter ρ.) Rho and Tau are "distribution-free" – their small sample distributions don't hinge on the parent normality assumption we are making here, so they have wider uses. All three however are asymptotically normal. With our assumptions, for fixed ρ it can be shown that as $n \to \infty$

$$\sqrt{n}\,(r - \rho) \xrightarrow{d} \mathcal{N}(0,\ 1 + o(1));$$

$$\sqrt{n}\left(\text{Rho} - \tfrac{3}{\pi}\rho\right) \xrightarrow{d} \mathcal{N}(0,\ 1 + o(1));$$

$$\sqrt{n}\left(\text{Tau} - \tfrac{2}{\pi}\rho\right) \xrightarrow{d} \mathcal{N}\left(0,\ \tfrac{4}{9} + o(1)\right).$$

Each appearance of $o(1)$ on the right is a term tending to 0 as $\rho \to 0$ (and free of n).

(a) Define the ARE (based on matching limit laws) for two estimators $\hat{\theta}_1, \hat{\theta}_2$ of a parameter θ.

(b) Basing three estimators of ρ above on each of r, Rho, and Tau in the obvious way, find ARE$(r, \tfrac{\pi}{3}\text{Rho})$, ARE$(r, \tfrac{\pi}{2}\text{Tau})$, and ARE$(\tfrac{\pi}{3}\text{Rho}, \tfrac{\pi}{2}\text{Tau})$ when $\rho = 0$. State any formulas that you use.

(c) Use the large-sample results for r and Tau to derive corresponding testing procedures for $H_0: \rho = 0$ vs. $H_1: \rho > 0$.

(d) Derive the approximate power functions for the procedures in (c).

(e) If a million observations gets you 70% power against a certain alternative $\rho > 0$ with the r-based procedure, approximately how many are needed to get the same power by the Tau-based procedure operating at the same significance level?