D. McFadden

## **PROBLEM SET 4 (Estimation)**

(Due Monday, March 2, with discussion in section on Feb. 25)

1. If h(t) is a convex function and  $t = T(\mathbf{x})$  is a statistic, then  $Eh(T) \ge h(ET)$ , with the inequality strict when h is not linear over the support of T. When h is a concave function,  $Eh(T) \le h(ET)$ . If T is an unbiased estimator of a parameter  $\sigma^2$ , what can you say about  $\sqrt{h(T)}$  as an estimator of  $\sigma$  and exp(h(T)) as an estimator of  $exp(\sigma^2)$ ?

2. A simple random sample i = 1,...,n is drawn from a binomial distribution b(K,1,p); i.e.,  $K_i = 1$  with probability p and  $K_i = 0$  with probability 1-p. Which of the following statistics are sufficient: a.  $(K_1,...,K_n)$  b.  $(K_1^2,[K_2+...+K_n]^2)$  c.  $\bar{K}$  d.  $(\bar{K},[K_1^2+...+K_n^2])$  e.  $[K_1^2+...+K_n^2]$ ?

3. You want to estimate mean consumption from a random sample of households i = 1,...,n. You have two alternative income measures,  $C_{1i}$  which includes the value of in-kind transfers and  $C_{2i}$  which excludes these transfers. You believe that  $\tilde{C}_1$  will overstate consumption because in-kind transfers are not fully fungible, but  $\tilde{C}_2$  will understate consumption because these transfers do have value. After some investigation, you conclude that  $0.7\cdot\tilde{C}_1 + 0.3\cdot\tilde{C}_2$  is an unbiased estimator of mean consumption. Your friend Dufus proposes instead the following estimator: Draw a random number between 0 and 1, report the estimate  $\tilde{C}_2$  if this random number is less than 0.3, and report the estimate  $\tilde{C}_1$  otherwise. Is the Dufus estimator unbiased? Does it pass the test of ancillarity?

4. Suppose  $T(\mathbf{x})$  is an unbiased estimator of a parameter  $\theta$ , and that T has a finite variance. Show that T is *inadmissible* by demonstrating that  $(1-\lambda)\cdot T(\mathbf{x}) + \lambda \cdot 17$  for  $\lambda$  some small positive constant has a smaller mean square error. (This is called a Stein shrinkage estimator. The constant 17 is obviously immaterial, zero is often used.)