

### Second Problem Set in STAT 700, F24

These 6 problems are due on Wednesday, September 18, 2024, 11:59pm, as a pdf-format upload to the course ELMS page. They include:

(I) Two problems on **identifiability**: # 1.1.3, 1.1.10 in Bickel and Doksum.

(II) One problem on expressing a statistic as a function(al) of the empirical distribution function:

**3.** Suppose that  $X_i$ ,  $i = 1, \dots, n$  are *iid* real-value observations with unknown distribution function  $F$  with density  $f$  which is positive on the whole real line and such that  $E(X_1^2) < \infty$ . We want to estimate the quantity

$$\theta \equiv E\left((X_1 - E(X_1))^2 \mid X_1 \geq 1\right)$$

Show that this parameter  $\theta$  can be expressed in the form

$$h(E(g_1(X_1)), E(g_2(X_1)), \dots, E(g_k(X_1)))$$

where  $h$  is a smooth real-valued function, and  $g_1, \dots, g_k$  are real-valued functions that need not be continuous, for a finite  $k$ . Use this representation to suggest a natural consistent estimator of  $\theta$  based on the data-sample  $\underline{X}_n = (X_1, X_2, \dots, X_n)$ .

(III) Three problems on **sufficiency**: problems # 1.5.5 and 1.5.16 in Bickel and Doksum, plus the following additional problem:

**6.** If  $Z_i \sim \mathcal{N}(\mu, 3\mu^2)$  are *iid* observations for  $i = 1, \dots, m$ , then find a minimal sufficient statistic  $(T_1, T_2)$  for  $\theta = \mu$ , and prove that it is minimal. Prove also that it is **not** complete.