Problem Set #1

1. Let X be a vector in $\mathbf{R}^k$ with distribution P. Prove that $E_P[XX']$ has full rank if and only if there exists no $A \subseteq \mathbf{R}^k$ such that A is a proper linear subspace and $P\{X \in A\} = 1$.

2. Consider the binary choice model

$$Y = \mathbf{1}\{\beta_1 + \beta_2 X_2 - \epsilon \geq 0\} \ .$$

Suppose $\text{supp}(X_2) = \mathbf{R}$ and $\text{Med}(\epsilon|X) = 0$. Show that $\text{sign}(\beta_2)$ is identified; that is, determine from the distribution of the observed data whether $\beta_2 > 0$, $\beta_2 = 0$, or $\beta_2 < 0$. (Hint: Since β_2 may be equal to zero, this result does not follow immediately from the final theorem in the lecture on identification. Instead, use the lemma that was used in the proof of that theorem.)

3. Let D be an indicator variable for whether or not a patient was given a particular drug. Denote by Y_d the outcome of the patient if $D = d$. The observed outcome is given by $Y = DY_1 + (1-D)Y_0$. For example, if $D = 1$, then $Y = Y_1$ is observed and Y_0 is the outcome that would have been observed if, counter to fact, the patient had not been given the drug. Let F denote the distribution of (Y, D). Interest focuses on the average treatment effect, $E[Y_1 - Y_0]$. Suppose (Y_1, Y_0) is independent of D, as would be the case in a perfectly randomized trial. Show that

$$E[Y_1 - Y_0] = E[Y|D = 1] - E[Y|D = 0] \ .$$

Conclude that the average treatment effect is identified from F. What about the distribution of Y_1 or $Y_1 - Y_0$?

4. For each $\epsilon > 0$, let $A_n(\epsilon)$ be a sequence of numbers such that $A_n(\epsilon) \to 0$. Show that there exists a sequence $\epsilon_n \to 0$ such that $A_n(\epsilon_n) \to 0$. (Hint: First argue that for each k there exists n_k such that for $n > n_k$ we have $|A_n(\frac{1}{k})| < \frac{1}{k}$. Argue next that we may assume w.l.o.g. that $n_1 < n_2 < \cdots$. Finally, show that the sequence $\epsilon_n = \frac{1}{k}$ where $n_k \leq n \leq n_{k+1}$ works.)

5. Consider the problem of testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1 \neq \theta_0$. Suppose that for each $\alpha \in (0, 1)$ there exists a test of this null hypothesis that is asymptotically of level α and that is consistent (i.e., has power tending to one). Show that there exists a sequence of tests in which the probability of both Type 1 and Type 2 error go to zero simultaneously!

6. Recall the symmetric location model studied in class. Prove that the Wilcoxon Signed Rank Statistic, W_n, satisfies

$$W_n - \frac{1}{\sqrt{n}} \sum_{1 \leq i \leq n} G(|X_i|)\text{sign}(X_i) = o_{P_0}(1) \ ,$$

where G is the distribution of $|X_i|$.