Patrik Guggenberger, Department of Economics, Penn State, Econ 501, 8/17/2018

Comprehensive Exam

This is a closed-book exam. The total number of points is 100. The number of points per question is indicated in round brackets. Please provide complete explanations for your answers. Write clearly.

1. **(20 total)** Suppose the true performance of student i on this exam, X_i, satisfies $X_i \sim N(\mu, \sigma^2)$. Suppose, however, that the grading process records Y_i, where

$$Y_i = X_i + u_i + v_i,$$

$u_i \sim N(0,1)$ is the measurement error from the first grader, $v_i \sim N(0,1)$ is the error from the second grader, and X_i, u_i, v_i are mutually independent.

8 a) (10) Calculate $X_i^* = E(X_i|Y_i)$ and $var(X_i|Y_i)$.

b) (10) Assume $i = 1, ..., n$. Derive the probability limits of $\overline{X}^* = n^{-1} \sum_{i=1}^n X_i^*$ and of $T_n = \log(n^{-1} \sum_{i=1}^n (X_i^* - \overline{X}^*)^2)$ as $n \to \infty$.

You can apply any theorem we proved in class.

2. **(30 total)** Suppose that $X_i \sim N(\mu_i, 1)$ for $i = 1, ..., n$ are independent. Consider the following five tests of the null hypothesis $H_0 : \mu_1 = ... = \mu_n = 0$ versus the alternative H_1 : "not H_0" with acceptance regions

$$\begin{aligned}
A_1 &= \{(x_1, ..., x_n) : \textstyle\sum_{i=1}^n x_i \leq k_1\}, \\
A_2 &= \{(x_1, ..., x_n) : |\textstyle\sum_{i=1}^n x_i| \leq k_2\}, \\
A_3 &= \{(x_1, ..., x_n) : x_i^2 \leq k_3 \text{ for all } i = 1, ..., n\}, \\
A_4 &= \{(x_1, ..., x_n) : x_1^2 \leq k_4\}, \\
A_5 &= \{(x_1, ..., x_n) : \textstyle\sum_{i=1}^n x_i^2 \leq k_5\},
\end{aligned}$$

where k_j, $j = 1, ..., 5$ are constants (that possibly depend on the sample size n) chosen to ensure that the tests have size equal to α.

a) (10) Determine k_j, $j = 1, ..., 5$.

b) (10) Let $n = 2$ from now on. Determine the power functions of the first and the fourth test.

c) (10) Show that none of the five tests is (weakly) more powerful than all the other tests considered here uniformly over the alternative space $\{(\mu_1, \mu_2) : -\infty < \mu_1, \mu_2 < \infty\}$.

3. **(15 total)** Assume f and g are measurable functions $\Omega \to [0, +\infty]$, where $(\Omega, \mathcal{F}, \mu)$ is a measure space. Define $v(A) = \int_A f d\mu$ for $A \in \mathcal{F}$. Show that

a) (6) v is a measure,

b) (9) $\int g dv = \int g f d\mu$. For b) a sketch of the proof is sufficient.

4. **(25 total)** This question is about MLE and GMM.

a) (5) Assume X_i are i.i.d. uniform $[-\theta, \theta]$, $i = 1, ..., n$ for some $\theta > 0$. Find the MLE of θ.

b) (5) Suppose $Y_i = \theta + \varepsilon_i$, $i = 1, ..., n$, where ε_i are i.i.d uniform $[-1, 1]$ and $\theta \in (0, 1)$. What can you say about the MLE of θ?

c) (15) In b) compute $E|Y_i|$. Use this calculation to suggest a GMM estimator (with nonrandom weighting "matrix" equal to 1) for the unknown parameter θ. Derive the probability limit and the limiting distribution of this estimator. (It is easier to work from first principles here rather than using the theory of extremum estimators developed in class.)

5. (10 total) Let $\{X_i : 1 \leq i \leq n\}$ be i.i.d. random variables on a probability space $(\Omega, \mathcal{F}, P)$. Let $M > 0$ be a finite constant.

a) (6) Show that

$$P\left[\frac{1}{n}\sum_{i=1}^{n}|X_i| \leq M\right] = 1 \Rightarrow P[|X_i| \leq M] = 1.$$

b) (4) Show that the implication in a) is in general not true if the independence assumption is dropped.

Andres Aradillas-Lopez, Department of Economics, Penn State, Econ 510, 8/17/2018

Comprehensive Exam

This is a closed-book exam. The total number of points is 100. The number of points per question is indicated in round brackets. Please provide complete explanations for your answers. Write clearly.

Consider the following model

$$Y_i = X_i'\beta_0 + \varepsilon_i$$

where $X_i \in \mathbb{R}^{k_1}$ (assume X_i includes an intercept). ε_i is an unobserved latent variable. Suppose there exists an observable $Z_i \in \mathbb{R}^{k_2}$ such that the distribution of ε_i conditional on Z_i satisfies

$$Pr\left(\varepsilon_i \leq c - \epsilon | Z_i\right) = Pr\left(\varepsilon_i \geq c - \epsilon | Z_i\right) \quad \forall\, \epsilon \in \mathbb{R},$$

almost surely in Z_i, where $c \in \mathbb{R}$ is an unknown constant. Note that the previous condition simply states that, almost surely in Z_i, the conditional distribution of ε_i given Z_i is symmetric around the unknown constant c. For simplicity, assume that the distribution of ε_i conditional on Z_i is absolutely continuous with respect to Lebesgue measure (i.e, ε_i does not have point-masses conditional on Z_i).

Assume throughout that the econometrician observes an iid sample $(Y_i, X_i, Z_i)_{i=1}^n$, and that $k_2 \geq k_1$. In what follows, feel free to describe any additional assumptions you think are necessary. However, incorrect assumptions will subtract points from your answers.

(i) **[20 points]** Show that the previous assumptions imply a conditional mean-zero restriction. From here, characterize a system of unconditional (in Z_i) mean-zero restrictions that can be used to identify and estimate β_0. Can you identify the intercept in β_0 under the assumptions described above?

(ii) **[40 points]** Describe a GMM estimator based on the unconditional moment restrictions you described in part (i). Characterize the asymptotic distribution of this GMM estimator. What is the probability-limit of the estimated intercept?

(iii) **[40 points]** Let us limit attention to instrument-functions of the type $A(Z)$, where $A(Z)$ is a $k_1 \times 1$ vector of functions of Z (recall that k_1 denotes the dimension of β_0). Describe how to find the efficient choice of $A(Z)$ given the conditional mean-zero restriction you found in part (i).