MIT 14.381: Fall 2019

Qualifying/Review Problem Set (due 10/31/19)

1. You're designing an RCT. Out of T individuals in the experiment, an untreated random sample of size n provides observations $\{y_{0i}\}_{i=1,\ldots,n}$, while the treated random sample of size $m = T - n$ provides observations $\{y_{1i}\}_{i=1,\ldots,m}$.

(a) Assuming observations are independent, derive the formula for the standard error of the estimated treatment effect, $\bar{y_1} - \bar{y_0}$.

(b) Show that if the outcome variable is homoskedastic, the optimal experimental design for estimating the average treatment effect sets the proportion treated equal to 1/2 (i.e., the standard error is minimized by setting $n = m$).

2. Consider two random variables, Y and X. Show that the variance of Y equals the variance of the conditional expectation function, $E[Y|X]$, plus the average conditional variance of Y given X. This decomposition is called the analysis of variance (ANOVA) formula. Note that the ratio of the variance of $E[Y|X]$ to the variance of Y is between zero and one. When does this "ANOVA R^2" equal the R^2 from a regression of Y on X?

3. You're interested in the linear regression $y_i = \beta x_i + \epsilon_i$ for a $k \cdot 1$ vector, x_i. Show that the population regression coefficient is $\beta = E[x_i x_i']^{-1} E[x_i y_i]$. Suppose $\hat{\beta}$ is an estimator of β. Define the *mean-squared error* (MSE) of $\hat{\beta}$ by $MSE(\hat{\beta}) \equiv E[(\hat{\beta} - \beta)^2]$.

(a) Show that $MSE(\hat{\beta}) = Bias^2(\hat{\beta}) + Var(\hat{\beta})$, where $Bias(\hat{\beta}) \equiv E[\hat{\beta} - \beta]$.

(b) The OLS estimator of β is the sample analog, $\hat{\beta}_{OLS}$. The Gauss-Markov Theorem says that, under certain conditions, the OLS estimator $\hat{\beta}_{OLS}$ is the "best" linear unbiased estimator of β, where "best" means the estimator that minimizes sampling variance. What are these conditions? Your classmate, a top performer in 14.382, claims she's found a linear estimator $\hat{\beta}_{Better}$ with lower MSE than $\hat{\beta}_{OLS}$. Does her claim contradict the theorem? Why or why not?

4. Consider the regression of Y on X_1, X_2, and X_3.

(a) Explain how to calculate the coefficient on X_1 using a two-step procedure in which the second step is a bivariate regression. Use the properties of regression residuals to explain why this works.

(b) What's the relationship between the coefficient on X_1 in a model that includes only the first two regressors and the coefficient on X_1 in a model that includes all three? Why is this important?

5. Suppose you use the vector $\{W_1, W_2, ..., W_K\}'$ as instruments to compute a two-stage least squares (2SLS) estimate of the effect of a scalar endogenous variable, X, on dependent variable, Y. Show that this 2SLS estimator is an IV estimator, that is, it can be written $\frac{Cov(Z,Y)}{Cov(Z,X)}$ for some scalar instrument, Z. What's Z?

6. Suppose the probability a woman works is described by a latent-index model:

$$y_i = 1(X_i'\beta > \varepsilon_i),$$

where y_i is employment status, X_i is a vector of personal characteristics, and ε_i is normally distributed and independent of X_i.

(a) Show that β is identified "up to scale". Explain why this is the best you can hope for from this model.

(b) Show that the marginal effects of regressors on employment rates are identified and derive a rule of thumb to compute them.

(c) Write down the likelihood function and FOC that generates the MLE for β.

(d) Propose a weighted nonlinear least squares estimator asymptotically equivalent to the MLE.