Last Updated: *February 8, 2019*

Overview of Useful Econometric Tools

By: Ariel Boyarsky (aboyarsky@uchicago.edu)

Preface

The following are a diverse set of econometric tools useful in economic analysis. Each section gives and overview of a different tool. The main purpose is to serve as a quick reference. Each section is written to be independent of the others. I will also try to add useful papers and book chapters to serve as more complete references. Please note this is a work and progress and I will add to it as time goes on. Please feel free to email me with any mistakes.

Contents

1 Hedonic Regression

This technique is common in evaluating real estate assets.

Suppose the price of a product are defined by,

$$p_i = f(z_{i,1}, \ldots, z_{i,k}, e) \tag{1}$$

Assumptions: $\mathbb{E}[e_i|z_i] = 0$ (Mean Independence - needed for unbiasedness), Homogeneity of Preferences, Information of Features, No Multico (A possible problem if there is a large number of categorical variables).

Where the z's define k characteristics and e is a random error term of unobservables.

In housing price models it is often common to assume these characteristics may be sorted into structural s, location n, and environmental v characteristics. As such one may often see the model,

$$P = f(s, n, v)$$

Our goal is to get the marginal contributions of each characteristic to price. As such, we need to make a parametric assumption to identify these contributions. In housing price analysis this is often a linearity assumption as in Equation 2. In other goods we can also specify a log-linear model as shown in Equation 3 - this is often done for high tech codes. This form is useful if we believe there is a multiplicative relationship between the features.

$$p_i = \beta_0 + \sum_{k=1}^{K} \beta_k z_{ik} + u \tag{2}$$

$$\ln p_i = \beta_0 + \sum_{k=1}^{K} \beta_k z_{ik} + u \tag{3}$$

In the second stage we estimate the willingness of the consumer to pay. As such,

$$p_i = g(X, Y, u)$$

Where X are the covariates of the product, Y is the consumer income, and u are individual level unobservables which we often are assume proxy for tastes.

1.1 References

Diewert (2003). Hedonic Regression: A Consumer Theory Approach.

2 Quantile Regression

Quantile regression is a tool used for retrieving the conditional quantile of an outcome variable. The problem is derived from the check loss function,

$$\rho_\tau(y) = y(\tau - \mathbb{1}\{y < 0\}) = \tau y_+ + (1-\tau)y_- \tag{4}$$

Where, $y_+ = max\{y, 0\}$ and $y_- = max\{-y, 0\}$ and $\tau \in [0, 1]$. Then given data $(Y_i, X_i)_{i=1,\ldots,n}$ quantile regression solves:

$$\min_\beta \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^T\beta) \tag{5}$$

Consider setting $\tau = 0.5$ then,

$$\sum_{i=1}^{n} \rho_{0.5}(Y_i - X_i^T\beta) = \sum_{i=1}^{n} |Y_i - X_i^T\beta|$$

Which is a median regression - under L_1 loss.

3 MCMC

Todo.

4 RCTs

ATE, ITT.

Todo.

5 Causal Inference in Observational Studies

Todo. Include the estimation procedures and interpretation for:

5.1 Selection on Observables and Matching

5.2 IV Approaches

5.2.1 LATE

5.2.2 MTE

5.3 RDD

5.4 DID