ECON 251

Discussion Section

Week 3 Solutions

1. Recap basic regression definitions and results.

   $E[Y|X=x] = g(x)$

   $E[Y|X=x] = \beta_0 + \beta_1 x$ whenever $g(\cdot)$ is linear

   $Y = \beta_0 + \beta_1 X + U$

   $E[Y|X] = E[\beta_0 + \beta_1 X + U|X] = \beta_0 + \beta_1 X + E[U|X] = \beta_0 + \beta_1 X$ if $E[U|X] = 0$

$$\min_{\beta_0,\beta_1} \sum_{i=1}^{N} U_i^2 \Leftrightarrow \min_{\beta_0,\beta_1} \sum_{i=1}^{N} (Y_i - \beta_0 - \beta_1 X_i)^2$$

$$\Rightarrow \qquad \hat{\beta}_1^{OLS} := \frac{\sum(X_i-\overline{X})(Y_i-\overline{Y})}{\sum(X_i-\overline{X})^2} \qquad \hat{\beta}_0^{OLS} := \overline{Y} - \hat{\beta}_1^{OLS} \cdot \overline{X}$$

2. Decompose total variance and derive R^2 in a simple linear regression model.

   Total variation in our outcome variable $SST := \sum_{i=1}^{N}(Y_i - \overline{Y})^2$

   Amount of variation explained by model $SSE := \sum_{i=1}^{N}(\hat{Y}_i - \overline{Y})^2$

   Amount of UNexplained (residual) variation $SSR := \sum_{i=1}^{N} \hat{U}_i^2 = \sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2$

   $\Rightarrow$ fraction of total variation explained by model $R^2 := \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$

3. Explain the difference between the conditional independence ($E(u|x) = E(u) = 0$) which is stronger and gives unbiasedness, versus exogeneity ($\text{Cov}(u, x) = 0$) which is weaker (since it's implied by but doesn't imply independence) and gives consistency.

$E(u|x) = E(u) = 0$ requires that the average value of unobserved variables does not vary with the treatment. In other words, the conditional expectation function of the error term does not depend on the value of x, and always equals its unconditional expectation $E(u)$ which we can normalize to $= 0$ since we usually include an intercept.

Conditional independence implies exogeneity, but exogeneity does not imply independence. Hence, exogeneity is a weaker assumption. That $\text{Cov}(u, x) = 0$ means there is no linear relationship between these variables (though their population relationship via the conditional expectation function is free to vary possibly nonlinearly). Exogeneity implies that $\hat{\beta}_1^{OLS}$ is consistent for β_1 (as $N \to \infty$), while independence implies $\hat{\beta}_1^{OLS}$ is unbiased for β_1 (for any sample size N).

4. Let $Y = \beta_0 + \beta_1 X + e$ denote a linear population regression function. Prove that whenever $Cov(X, e) = 0$ we can write the values of $\{\beta_0, \beta_1\}$ in terms of $E(X)$, $E(Y)$, $Var(Y)$, and $Cov(X, Y)$. What is the economic meaning behind the assertion that the value of the parameter $Cov(X, e)$ must be $= 0$ in the population? (Bonus: if it fails, why would an infinite sample, or the whole population, be useless for establishing causality?)

$\text{Cov}(X, e) = 0$

$\text{Cov}(X, Y - \beta_0 - \beta_1 X) = 0$

$\text{Cov}(X, Y) - \text{Cov}(X, \beta_0) - \text{Cov}(X, \beta_1 X) = 0$

$\text{Cov}(X, Y) - 0 - \beta_1 \text{Cov}(X, X) = 0$

$\text{Cov}(X, Y) - \beta_1 \text{Var}(X) = 0$

$$\beta_1 = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{E[XY] - E[X]E[Y]}{E[X^2] - E[X]^2}$$

$$\Rightarrow \hat{\beta}_1^{MOM} := \frac{\widehat{\text{Cov}}(X, Y)}{\widehat{\text{Var}}(X)} = \hat{\beta}_1^{OLS}$$

$E[Y] = E[\beta_0 + \beta_1 X + e]$

$E[Y] = \beta_0 + \beta_1 E[X] + E[e]$

$E[Y] = \beta_0 + \beta_1 E[X] + 0$

$$\beta_0 = E[Y] - \beta_1 E[X]$$
$$\Rightarrow \hat{\beta}_0^{MOM} := \hat{E}[Y] - \hat{\beta}_1^{MOM} \cdot \hat{E}[X]$$
$$= \overline{Y} - \hat{\beta}_1^{MOM} \overline{X} = \hat{\beta}_1^{OLS}$$

5. Nonlinearity in linear regressions. Interpreting regression coefficients when the outcome or response has been transformed. (Recall elasticity definition from microeconomics!)

$Y = \beta_0 + \beta_1 X + U$ — $\hat{\beta}_1^{OLS}$ gives change in Y given +1 unit increase in X

$\log Y = \beta_0 + \beta_1 \log X + U$ — $\hat{\beta}_1^{OLS}$ gives % change in Y given +1% increase in X

$\log Y = \beta_0 + \beta_1 X + U$ — $100 \cdot \hat{\beta}_1^{OLS}$ gives % change in Y given +1 unit increase in X

$Y = \beta_0 + \beta_1 \log X + U$ — $\frac{1}{100} \hat{\beta}_1^{OLS}$ gives change in Y given +1% increase in X