Project 1

Antonio Jurlina

2/1/2020

Time Series Econometrics

Dr. Wiesen

Project 1

1.) Let x_i and y_i (where $i = 1, 2, ..., N$) be iid observations. To help me prove the unbiasedness of the OLS estimator, I used the following expression in class.

$$\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{N}(x_i - \bar{x})y_i$$

Prove the above expression.

$$\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) =$$

$$\sum_{i=1}^{N}(x_i - \bar{x})y_i - \sum_{i=1}^{N}(x_i - \bar{x})\bar{y} =$$

$$\sum_{i=1}^{N}(x_i - \bar{x})y_i - \sum_{i=1}^{N}x_i\bar{y} + \sum_{i=1}^{N}\bar{x}\bar{y} =$$

$$\sum_{i=1}^{N}(x_i - \bar{x})y_i - N\bar{x}\bar{y} + N\bar{x}\bar{y} =$$

$$\sum_{i=1}^{N}(x_i - \bar{x})y_i$$

2.) Consider the following regression:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad , \quad i = 1, 2, ..., N$$

Assume that the conditional mean assumption is true. Namely, assume that $E(\varepsilon_i|x_1, x_2, ..., x_n) = E(\varepsilon_i|x_i) = 0$ Also assume that the OLS estimator $\hat{\beta}_1$ is unbiased (I showed a proof of this in class). Prove that the OLS estimator $\hat{\beta}_0$. Namely, prove that $E(\hat{\beta}_0) = \beta_0$. Hint: for the last step of the proof use the law of iterated expectations.

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$E(\hat{\beta}_0) = E(\bar{y}) - E(\hat{\beta}_1 \bar{x})$$

$$E(\hat{\beta}_0) = E[E(y)] - E(\hat{\beta}_1 \bar{x})$$

$$E(\hat{\beta}_0) = E[E(\beta_0 + \beta_1 x_i + \varepsilon_i)] - \bar{x}\beta_1$$

$$E(\hat{\beta}_0) = E[E(\beta_0) + E(\beta_1 x_i) + E(\varepsilon_i)] - \bar{x}\beta_1$$

$$E(\hat{\beta}_0) = E[E(\beta_0) + E(\beta_1 x_i)] - \bar{x}\beta_1$$

$$E(\hat{\beta}_0) = E(\beta_0 + \bar{x}\beta_1) - \bar{x}\beta_1$$

$$E(\hat{\beta}_0) = E(\beta_0) + E(\bar{x}\beta_1) - \bar{x}\beta_1$$

$$E(\hat{\beta}_0) = \beta_0 + \bar{x}\beta_1 - \bar{x}\beta_1$$

$$E(\hat{\beta}_0) = \beta_0$$

3.) For this problem, use the data set *Income_credit_card_data.xlsx*. The variable *AGE* is age, the variable *INCOME* is monthly income, the variable *INCPER* is income divided by number of dependents, and the variable *SPENDING* is average monthly credit card expenditures. Consider the following regression:

$$SPENDING_i = \beta_0 + \beta_1 AGE_i + \beta_2 AGE_i^2 + \beta_3 INCPER_i + \varepsilon_i \quad , \quad i = 1, 2, ..., N$$

Make the usual assumptions; the equation is true equation, the conditional mean assumption holds, the conditional errors are iid, and the observations are iid. Use R to estimate the above regression and report the results.

term	estimate	standard error	t statistic	p value
intercept	-23.7897	28.1753	-0.8443	0.3985
age	10.4214	1.4568	7.1536	0
age_sq	-0.1297	0.0182	-7.1134	0
incper	0.0025	2e-04	11.8099	0

Say you want to know at what age is SPENDING maximized. The age at which SPENDING reaches its maximum is a function of the β parameters, call this θ In R, use the bootstrapping procedure to calculate an estimate of the variance of $\hat{\theta}$. Let the number of bootstrapped samples be $M = 20000$.

```
bootstrap <- function(x, M) {

  theta <- c(1:M)

  for (j in 1:M){
    data <- sample_n(x, size = nrow(x), replace = TRUE)
    reg <- lm(data$spending ~ data$age + data$age_sq + data$incper)
    reg %>%
      summary() %>%
      coef() %>%
      .[2, 1] -> beta_1
    reg %>%
      summary() %>%
      coef() %>%
      .[3, 1] -> beta_2
    theta[j] <- -0.5 * beta_1 / beta_2
  }

  theta_var <- var(theta)

  return(theta_var)
}

theta_var <- bootstrap(x = income_credit_card_data, M = 1000)
```

The variance of $\hat{\theta}$ is 1.5750307 .

More at https://github.com/antoniojurlina/time_series_econometrics