Assignment 5 -

Informed guesses on consistency (1)

After running a regression of y on x_1 and x_2, we test whether the coefficient on x_2 is significantly different from 0. That is:

- H_o: $\beta = 0$
- H_a: $\beta \neq 0$

The probability of selecting the true model is $P(|T| \leq c|\beta = 0)$ or $P(|T| > c|\beta \neq 0)$. Let test size $= \gamma$.

When $\beta = 0$, then

$$\begin{aligned}\lim_{n\to\infty} Pr(TrueModel) &= Pr(|T| < c|\beta = 0) \\ &= (1 - \gamma) \nrightarrow 1\end{aligned}$$

When $\beta \neq 0$, then

$$\lim_{n\to\infty} Pr(TrueModel) = Pr(|T| > c|\beta \neq 0)$$

$$TestPower \xrightarrow{p} 1$$

Informed guesses on consistency (2)

When $\beta = 0$, the true model is $y = \alpha X_1 + e$. There are 2 possibilities.

1. T-test accept $\hat{\beta} = 0$, the fitted line is $\hat{y} = \hat{\alpha} X_1 + \hat{e}$. $\hat{\alpha}$ is consistent.
2. T-test reject $\hat{\beta} = 0$, the fitted line is $\hat{y} = \hat{\alpha} X_1 + \hat{\beta} X_2 + \hat{e}$. Since $\hat{\beta} \xrightarrow{p} \beta = 0$, thus $\hat{\alpha}$ is consistent.

When $\beta \neq 0$, the true model is $y = \alpha X_1 + \beta X_2 + e$. From part 1 we showed that when $\beta \neq 0$, the probability of of choosing the true model (Here is $y = \alpha X_1 + \beta X_2 + e$) converges in probability to 1. Thus $\hat{\alpha}$ is consistent.

2. Compare the variance of the three estimators

Table 1:

	n	alpha_UR	alpha_R	alpha_E	n.1	var(UR)	var(
1	50	0.194	0.197	0.195	50	0.210	0.14
2	100	0.199	0.199	0.200	100	0.143	0.10
3	150	0.200	0.198	0.199	150	0.118	0.08
4	200	0.202	0.200	0.200	200	0.102	0.07

The variance of the unrestricted estimator is significantly larger than that of the restricted. This is consistent with what we saw in class. When $\beta = 0$, adding additional (useless) variable X2 to the model will increase the variance of the estimator of α. The variances of the post-test OLS estimator lie in between the variance of the restricted and the unrestricted ones.

3. Compare the bias of the three estimators

Table 2:

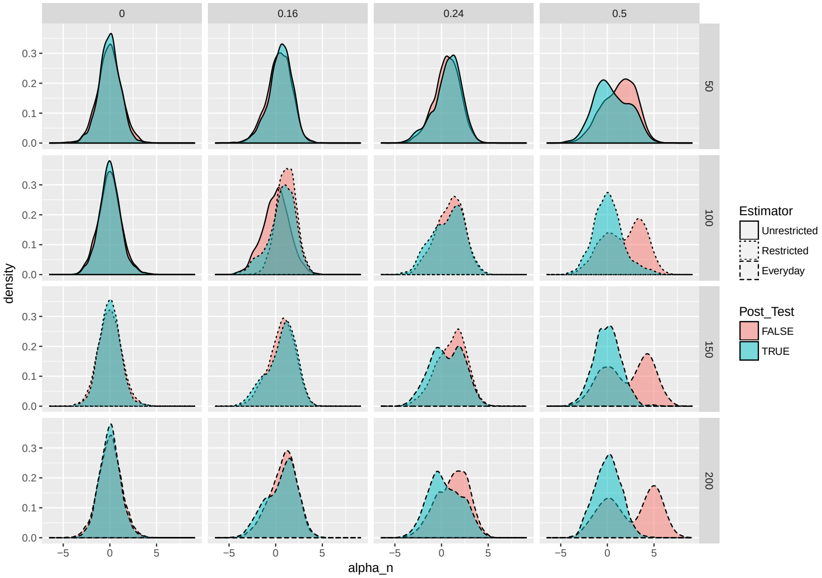
	n	beta	bias_ur	bias_r	bias_e	bias_omitted
1	50	0.160	0.001	0.111	0.071	0.112
2	100	0.160	0.003	0.114	0.066	0.168
3	150	0.160	-0.005	0.110	0.051	0.350
4	200	0.160	-0.001	0.110	0.044	0.112
5	50	0.240	-0.001	0.167	0.094	0.168
6	100	0.240	0.001	0.170	0.068	0.350
7	150	0.240	0.002	0.168	0.047	0.112
8	200	0.240	-0.001	0.166	0.030	0.168
9	50	0.500	0.006	0.353	0.072	0.350
10	100	0.500	0.0005	0.352	0.012	0.112
11	150	0.500	-0.001	0.351	0.0001	0.168
12	200	0.500	0.004	0.351	0.004	0.350

4. Ratio of alpha in the confidence interval

Table 3:

	n	beta	CI_ratio_ur	CI_ratio_r
1	50	0	0.950	0.950
2	100	0	0.951	0.949
3	150	0	0.944	0.948
4	200	0	0.949	0.950
5	50	0.160	0.962	0.888
6	100	0.160	0.943	0.792
7	150	0.160	0.951	0.726
8	200	0.160	0.954	0.665
9	50	0.240	0.949	0.804
10	100	0.240	0.946	0.621
11	150	0.240	0.946	0.477
12	200	0.240	0.956	0.378
13	50	0.500	0.955	0.378
14	100	0.500	0.947	0.102

5. Density plot of standardized distribution



6. Summary:

1. Consistency: Consistency holds for the everyday OLS.
2. Unbiaseness: When beta is zero, the everyday estimator is unbiased. When beta is nonzero, the bias of the estimator decreases as the sample size increases and as the true beta increases.
3. Efficiency: Variance of the everyday estimator gets larger when beta gets larger.

When the model is correctly specified, the usual OLS distribution well approximates the distribution of the post-test OLS.