Homework

Verbeek exercise 3.2

Fall 2005

Instructions

Using the 400 observations on 400 Dutch men's clothing stores, the following results are obtained.

Exercise 3.2

Consider the following linear models of sales:

$$\text{A:} \quad sales_i = \beta_1 + \beta_2 hoursw_i + \beta_3 ssize_i + u_i \tag{1}$$

$$\text{B:} \quad sales_i = \gamma_1 + \gamma_2 ssize_i + \gamma_3 nfull_i + \gamma_4 npart_i + u_i \tag{2}$$

$$\text{C:} \quad sales_i = \delta_1 + \delta_2 hoursw_i + \delta_3 ssize_i + \delta_4 nfull_i + \delta_5 npart_i + u_i \tag{3}$$

where $i = 1, 2, \ldots, 400$.

Results in tabular form

Least Least Squares (LS) estimates using the 400 observations
Dependent variable: sales

Variable	A	B	C
hoursw	37.52	-	32.76
	(2.83)		(3.06)
ssize	-22.14	-15.41	-23.91
	(1.62)	(1.63)	(1.65)
nfull	-	1348.9	557.30
		(176.4)	(172.2)
npart	-	605.0	685.0
		(255.6)	(225.4)
constant	5133.59	4924	3751
	(321)	(469)	(427)
$\bar{R}^2$	0.3626	0.2230	0.3961
AIC	6404.1722	6484.3746	6384.5446
SC	6416.1466	6500.3404	6404.502
RESET2	2.0913		2.1598
RESET3	16.9793		17.4638

Comments on results

The model selection rules all point to model C. Model C minimizes AIC and SC and maximizes $\bar{R}^2$. The p-values associated with both RESET2 tests are larger than .05, which indicates that neither model is misspecified. However, RESET3 *is significant* at the 5% level in both models A and C and therefore I believe it is safe to conclude that that both functional forms are misspecified.

Reformulating the model

$$sales_i = \beta_1 + \beta_2 hoursw_i + \beta_3 ssize_i + \beta_4 nown_i + u_i \tag{4}$$

To determine whether number of owners improves the model one can use a t-test on its coefficient (i.e., Ho: $\beta_4 = 0 \quad Ha$: Not Ho). The $t_{396} = -1.47$ which is not in the rejection region for a 10% test of this hypothesis. I would conclude that it does not improve the model.

You could also use the AIC or SC criteria. The results here are ambiguous. AIC decreases when adding *nown*, but SC does not.

$$sales_i = \beta_1 + \beta_2 hoursw_i + \beta_3 ssize_i + \beta_4 nown_i + \beta_5 npart_i + u_i \tag{5}$$

To determine whether number of part time workers improves the model one can use a t-test on its coefficient (i.e., Ho: $\beta_5 = 0 \quad Ha$: Not Ho). I tested this hypothesis conditional on the number of owners, but you don't have to

since this variable was not found to be an improvement to the model. The $t_{395} = 3.65$ which is in the rejection region for a 10% test of this hypothesis. I would conclude that it does improve the model. In addition, both AIC and SC criteria suggest a better overall fit when the number of part time workers is added to the model.

Non nested tests

The nonnested F test basically uses model C as the base model and tests two sets of hypotheses as subset models.

$$\text{Ho:} \quad sales_i = \beta_1 + \beta_2 hoursw_i + \beta_3 ssize_i + u_i \tag{6}$$

$$\text{Ha:} \quad sales_i = \gamma_1 + \gamma_2 npart_i + \gamma_3 ssize_i + \gamma_4 nfull_i + u_i \tag{7}$$

The first leg of the test augments (6) with $npart_i$ and $nfull_i$ and the joint hypothesis that the coefficients on these two variables is zero is tested using a Wald statistic. The $F_{2,395} = 12.01$ which has a pvalue of 0. Model A is not adequate at the 5% level of significance. Reversing Ho and Ha means augmenting (7) with $hoursw_i$ and results in a t-test of the hypothesis that $hoursw_i$ in the augmented model is significant. The $t_{395} = \sqrt{114.5156} = 10.71$ which is greater than 1.96, the 5% critical value. Model B is not adequate either.

The J-test is conducted with similar results. Model A against model B produces a t-ratio on $\hat{y}$ of 4.57; Model B against model A produces a t-ratio on $\hat{y}$ of 10.7. Both are greater than the 5% critical value of 1.96 and are deemed inadequate relative to each other.