Short 10–15 min   Stat learning

Math 583   Quiz 6 Fall 2017   Name____________________

e

3.15

1) The table below shows simulation results for bootstrapping OLS (reg) and lasso and ridge regression (RR) with 10-fold CV when $\boldsymbol{\beta} = (1, 1, 0, 0)^T$. The β_i columns give coverage = the proportion of CIs that contained β_i and the average length of the CI. The test is for $H_0 : (\beta_3, \beta_4)^T = \mathbf{0}$ and H_0 is true. The "coverage" is the proportion of times the prediction region method bootstrap test failed to reject H_0. OLS used 1000 runs while 100 runs were used for lasso and ridge regression. Since 100 runs were used, a cov in [0.89, 1] is reasonable for a nominal value of 0.95. If the coverage for both methods ≥ 0.89, the method with the shorter average CI length was more precise. (If one method had coverage ≥ 0.89 and the other had coverage < 0.89, we will say the method with coverage ≥ 0.89 was more precise.) The results for the lasso test were omitted since sometimes $\boldsymbol{S}^*_T$ was singular. (Lengths for the test column are not comparable unless the statistics have the same asymptotic distribution.)

Table 3.5

Table 1: Bootstrapping lasso and RR, $n = 100, \psi = 0.9, p = 4, B = 250$

		β_1	β_2	β_3	β_4	test
reg	cov	0.942	0.951	0.949	0.943	0.943
	len	0.658	5.447	5.444	5.438	2.490
RR	cov	0.97	0.02	0.11	~~0.992~~ 0.10	0.05
	len	0.681	0.329	0.334	0.334	2.546
reg	cov	0.947	0.955	0.950	0.951	0.952
	len	0.658	5.511	5.497	5.500	2.491
lasso	cov	0.93	0.91	0.92	0.99	
	len	0.698	3.765	3.922	3.803	

a) For β_3 and β_4 which method, ridge regression or the OLS full model, was better?

OLS   (RR cov is too low)

b) For β_3 and β_4 which method, lasso or the OLS full model, was more precise?

lasso   (3.8 < 5.5)

40

SL Q5

3.13 e 2) Suppose the MLR model $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{e}$, and the regression method fits $\boldsymbol{Z} = \boldsymbol{W\eta} + \boldsymbol{e}$. Suppose $\hat{Z} = 245.63$ and $\overline{Y} = 105.37$. What is $\hat{Y}$?

rob 79

$$\hat{Y} = \hat{Z} + \overline{Y} = \boxed{351.00}$$

25

3) To get a large sample 90% PI for a future value Y_f of the response variable, find a large sample 90% PI for a future residual and add $\hat{Y}_f$ to the endpoints of the of that PI. Suppose forward selection is used and the large sample 90% PI for a future residual is $[-778.28, 1336.44]$. What is the large sample 90% PI for Y_f if $\hat{\boldsymbol{\beta}}_{I_{min}} = (241.545, 1.001)^T$ used a constant and the predictor *mmen* with corresponding $\boldsymbol{x}_{I_{min},f} = (1, 75000)^T$?

marry data

3.14

$$\hat{Y}_f = 241.545 + 1.001(75000) = 75316.545$$

so 95% PI $= [-778.28 + \hat{Y}_f, 1336.44 + \hat{Y}_f]$

$$= \boxed{[74538.265, 76652.985]}$$

rob 7/10

25