

Pillar 1 - Conformal Prediction

DARS

2019-06-14

1 Set up

2 Course selection

I select courses with large and small n, and large and small CV error. I do not select courses whose lasso model contains less than 3 non-zero coefficients (df).

- COR1002: large n
- COR1004: large n and smaller CV error (mean absolute error) than COR1002
- SCI3003: small n and large CV error
- SCI2040: small n and small CV erro
- SSC3044: small CV error
- SSC3038: small CV error and n twice as large as SSC3044
- SCI2018: large CV error
- SCI2010: large CV error and n twice as large as SCI2018

3 Conformal Prediction setp by step

3.1 Cross-validation

3.2 Training, calibration and test sets

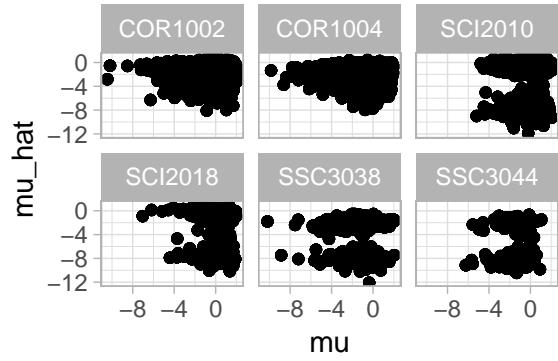
3.3 Fit lasso on training set

3.4 Compute standardized non-conformity scores on calibration test

3.5 Evaluate error rate on test set

4 Loop through all courses

Illustration of the effect of error_hat on interval's width (for significance of 99%):



Correlation between the actual errors and the predicted errors (error_hat):

5 Timing

```
## Time difference of 29.34084 mins
```