

## 3.4: Model Validation

Dr. Bean - Stat 5100

An extreme example of cross validation is leave one out cross validation (LOOCV), which is cross validation where only one observation is withheld on each iteration. Please describe the potential advantages and disadvantages of the LOOCV approach .

### Advantages

- The difference between the fitted model and the final model will be very small since only one observation is removed (small model bias).
- Unlike k-fold cross validation, there is no need to set a random seed to generate reproducible results.

### Disadvantages

- If your dataset is large, LOOCV can be very computationally expensive.
- An estimate of the model error based on a single point will have high variance.

Suppose that you use cross validation as a means of selecting variables for your final regression model. What issues, if any, are there with reporting cross-validated error from the final model resulting from that selection process?

The variable selection procedure is using all of the data. Reporting the error resulting from the final model incorrectly assumes that the variable selection process was “blind” to a subset of the data. The consequence is that the reported cross-validated error will be an underestimate of the true model error.

Suppose that you use cross validation as a means of selecting variables for your final regression model. What issues, if any, are there with reporting cross-validated error from the final model resulting from that selection process?

The variable selection procedure is using all of the data. Reporting the error resulting from the final model incorrectly assumes that the variable selection process was “blind” to a subset of the data. The consequence is that the reported cross-validated error will be an underestimate of the true model error.

What might be some reasons for selecting a final model that does NOT minimize the test or validation set error?

If there is a much simpler/more interpret-able method with roughly equal accuracy to a more complex method, the simpler method may be preferred. A similar reason would be to select a model that provides more “stable” predictions (small changes in Y for small changes in X) if the sacrifice in accuracy is small.

Our discussion of accuracy has revolved around squared errors  $(Y - \hat{Y})^2$ . What are some other ways to measure error and why might we be interested in them?

- Mean Absolute Error ( $|Y - \hat{Y}|$ ) gives less influence to the poorest predictions.
- Median Absolute Error gives even less influence to the poorest predictions.
- Mean Absolute Percentage Error ( $|\frac{Y - \hat{Y}}{Y}|$ ) accounts for differences in scale.

These alternatives place a different emphasis on good and poor predictions relative to the data scale.