Logistic Regression

Predictive Power and ROC

Demetris Athienitis

R^2 analogy

A naive way of summarizing predictive power is to calculate the correlation between observed responses and fitted responses.

Example (Horseshoe crab continued)

Correlation between the observed $y = 0, 1$ and fitted probabilities.

```
> cor(y,fitted(fit)) # weight
[1] 0.3955277
> cor(y,fitted(fit2)) # weight and color
[1] 0.4476282
> cor(y,fitted(fit2.2)) # weight and binary dark
[1] 0.3958138
> cor(y,fitted(fit2.3)) # weight and linear color
[1] 0.4385387
```

Cross-validation

A more sophisticated method is *leave-one-out cross-validation*, and producing classification tables.

1. Fit the model to the data leaving out i$^{\text{th}}$ observation
2. Use fitted model and the predictor settings of the i$^{\text{th}}$ observation to compute response $\hat{\pi}(\boldsymbol{x}_i)$
3. Predict

$$\hat{y}_i = \begin{cases} 1 & \hat{\pi}(\boldsymbol{x}_i) > 0.50 =: \pi_0 \quad \text{(cutoff probability)} \\ 0 & \hat{\pi}(\boldsymbol{x}_i) \leq 0.50 \end{cases}$$

where the cutoff of 0.50 can be altered.

Example (Horseshoe crab continued)

Using the model with weight and (qualitative) color we obtain the *confusion matrix*.

Actual	Predicted $\hat{y} = 0$	Predicted $\hat{y} = 1$	Total
$y = 0$	27	35	62
$y = 1$	17	94	111

$$\text{Sensitivity} = P(\hat{Y} = 1 | Y = 1) = \frac{94}{111} \approx 0.847$$

$$\text{Specificity} = P(\hat{Y} = 0 | Y = 0) = \frac{27}{62} \approx 0.435$$

$$P(\text{correct classification}) = \frac{94 + 27}{173} \approx 0.699$$

Section 2

Receiver Operating Characteristic Curve

The *receiver operating characteristic* (ROC) curve plots the true positive rate, sensitivity, against false positive rate, 1-specificity, as the cutoff value π_0 varies from 0 to 1. It can also be thought of as a plot of the Power as a function of the Type I Error of the decision rule.

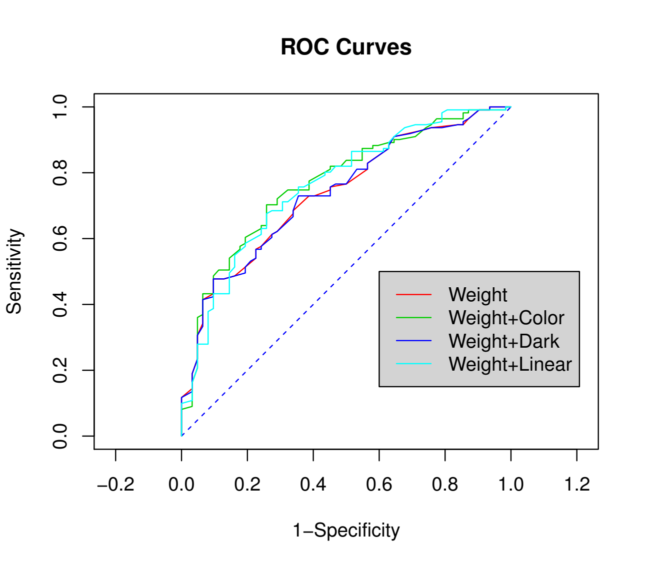
- The higher the sensitivity for a given specificity, the better, so a model with a higher ROC curve is preferred to one with a lower ROC curve
- The area under the ROC curve is a measure of predictive power, called the concordance index, c
  - Models with larger c have better predictive power
  - When $c = 1/2$ it is no better than random guessing
- If feasible, use cross-validation
- ROC curves should not be used with random predictors

Example (Horseshoe crab continued)

Concordance indexes for some models seen thus far

Model	Concordance
Weight	0.738
Weight and Color	0.769
Weight and Dark	0.738
Weight and Linear Color	0.761

Example (continued)



We learned

- Correlation of response to fitted values
- Cross-validation and the “confusion” matrix
- ROC