Logistic Regression

Probability and Statistical Inference

Bojan Božić

TU Dublin

Winter 2020

When And Why

- To predict an outcome variable that is categorical from one or more categorical or continuous predictor variables.
  - There are many important research topics for which the dependent variable is "limited".
  - E.g. Purchase or not, live or die, employ or not, commit crime or not, pass exam or not.
  - Participation data is not continuous or distributed normally.
- Used because having a categorical outcome variable violates the assumption of linearity in normal regression.
- Binary logistic regression is a type of regression analysis where the depenent variable is a dummy variable: coded 0 (did not vote) or 1 (did vote).

Logistic Regression

Predictive model - what does it allow you do?

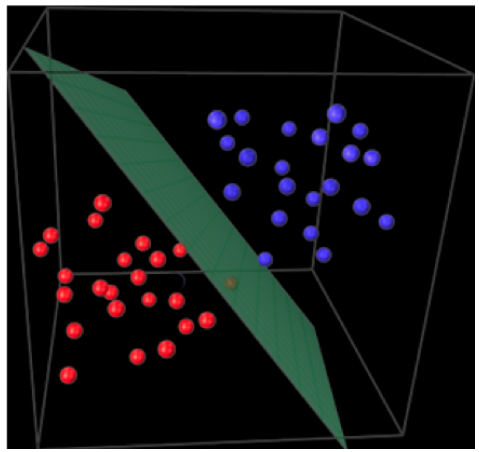
- Logistic regression does not try to predict the value of a numeric variable given a set of inputs.
- Instead, the output is a **probability** that the given input point belongs to a certain *class* or *category*.

Logistic Regression

- Suppose that we have only two classes (positive or negative).
- The probability in question is P+ =the probability that a certain data point belongs to the +ive class.
- $P- = 1 - P+$
- Thus, the output of Logistic Regression always lies in [0, 1].

Logistic Regression

- Assumption is that your input space can be separated into two nice 'regions', one for each class, by a **linear** (read: straight) **boundary**.
- For two dimensions, this boundary is a straight line-no curving.
- For three dimensions, its a plane. And so on.
- If your data points do satisfy this constraint, they are said to be **linear-separable**.



Logistic Regression

- Suppose we have two input variables the function corresponding to the boundary will be something like.
  - $b_0 + b_1 x_1 + b_2 x_2$
  - Our output variable is not part of our space unlike linear regression.
- Consider a point in the space (a, b) where a is the value of x_1 and b is the value of x_2.
  - $b_0 + b_1 a + b_2 b$
  - Depending on the location of (a, b), there are three possibilities to consider...

Logistic Regression - Location of (a,b)

- It lies in the region defined by points of the +ive class.
  - The result of our equation will be positive, lying somewhere in (0,∞).
  - Mathematically, the higher the magnitude of this value, the greater is the distance between the point and the boundary.
  - Intuitively then it means there is a greater probability that (a,b) belongs to the +ive class - probability between (0.5, 1].
- It lies in the region defined by points of the -ive class.
  - The result of our equation will be negative, lying in (-∞, 0).
  - As for the positive case, the higher the magnitude of the value, the greater the probability that (a,b) belongs to the -ive class - probability within [0, 0.5).
- It lies ON the linear boundary.
  - The result of our equation = 0.
  - Means that the model cannot really say whether it belongs to the +ive or -ive class.
  - As a result, the probability will be exactly 0.5.

Logistic Regression

- So how do we map our outcome of the equation (values -∞ to ∞) to an estimate of the likelihood of our point being in the +ive class to the probability?
- Use the **ODDS RATIO**.

Logistic Regression

- Let $P(X)$ denote the probability of an event X occurring. In that case, the odds ratio $OR(X)$ is defined by

$$OR(X) = \frac{P(X)}{1 - P(X)'}$$

- The ratio of the probability of the event happening, vs. it not happening.
- Probability and odds convey the exact same information but P goes from 0 to 1, OR goes from 0 to ∞.

Logistic Regression

- To work in the same space as our boundary equation which goes from (-∞ to ∞), we need to take the **logarithm** of OR, called the **log-odds function**.
- This is e to the power of the value of our boundary equation.

Logistic Regression

1. Compute the boundary equation (alternatively, the log-odds function) value, $b_0 + b_1 a + b_2 b$. Lets call this value t.
2. Compute the Odds Ratio, by doing $OR+ = e^t$ (Since t is the logarithm of $OR+$) - e is the base of the natural logarithm.
3. Knowing $OR+$, it is possible to compute $P+$.

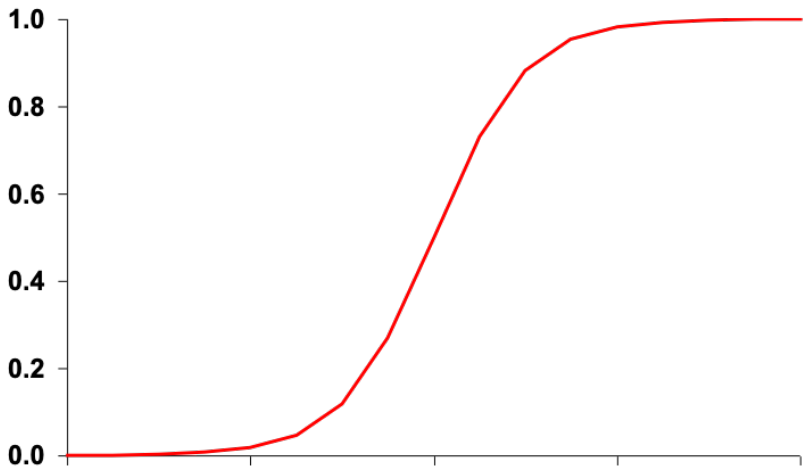
Logistic Regression With One Predictor

- Outcome:
  - We predict the probability of the outcome occurring.
- b_0 and b_1:
  - Can be thought of in much the same way as regression.
  - Note the normal regression equation forms part of the logistic regression equation.

$$P(Y) = \frac{1}{1 + e^{-(b_0 + b_1 X_1 i)}}$$

Logistic Function

$$P(Success|X) = \frac{e^{\beta_0+\beta_1 X}}{1 + e^{\beta_0+\beta_1 X}}$$



Binary Logistic Regression

- Binary outcome variable.
- Code 0 when event does not occur.
- Code 1 when it does.
- We are interested in the Odds of something happening.

Odds ratio

- Odds = probability of an event occurring / probability of event not occurring = $\frac{p}{1-p}$.
- Lets consider the relationship between sentence received in court for a crime and gender.
  - Looking at example from Tarling Statistical Modelling for Social Researchers.
  - Using sample from Offenders Index in the UK.

Odds ratio

Table 3.1 Type of sentence awarded by sex: adults aged 21 and over convicted of shop theft

Sentence	*Males*		*Females*		*Total*	
	Number	*%*	*Number*	*%*	*Number*	*%*
Custody	650	29.4	90	13.9	740	25.9
Community penalty	537	24.3	181	27.9	718	25.1
Fine	595	26.9	151	23.3	746	26.1
Discharge	429	19.4	227	35.0	656	22.9
Total	2211	100	649	100	2860	100

Pearson chi-squared: 104.975; df, 3; $p < .001$
Likelihood ratio: 106.887; df, 3; $p < .001$

Note: It is not necessary to quote these statistics to three decimal places. They are only reported here in order to make comparisons with subsequent computer output.

- Custody - some form of imprisonment.
- Community penalty – probation etc.
- Fine - monetary penalty or compensation to victim.
- Discharge - either not guilty or guilty and no sentence applied (or sentence conditional on no reoffending for a period).

Odds ratio

Table 3.1 Type of sentence awarded by sex: adults aged 21 and over convicted of shop theft

Sentence	*Males*		*Females*		*Total*	
	Number	*%*	*Number*	*%*	*Number*	*%*
Custody	650	29.4	90	13.9	740	25.9
Community penalty	537	24.3	181	27.9	718	25.1
Fine	595	26.9	151	23.3	746	26.1
Discharge	429	19.4	227	35.0	656	22.9
Total	2211	100	649	100	2860	100

Pearson chi-squared: 104.975; df, 3; $p < .001$
Likelihood ratio: 106.887; df, 3; $p < .001$

Note: It is not necessary to quote these statistics to three decimal places. They are only reported here in order to make comparisons with subsequent computer output.

- There is an association between gender and sentencing from the table.
- Females:
  - Less likely to receive custodial sentence13.9% receive custodial sentence, 35% are discharged.
  - Slightly more likely to receive probation (27.9%) than be fined (23.3%).
- Males:
  - More likely to receive custodial sentence (29.4%) and less likely to be discharged (19.4%).
  - Similar proportions receiving probation (24.3%) and fined (26.9%).

Odds ratio

$$\frac{p}{1-p} = \frac{.294}{.706} = .416$$

The odds of being a male and not sentenced to custody is the inverse of the odds of being sentenced to custody $= 1/.416 = 2.403$.

Thus males are 2.4 times more likely to receive a sentence other than custody than they are to receive a custodial sentence.

Odds ratios

- For females, probability of being given a custodial sentence is $90/649 = .139$.
- Probability of not being given a custodial sentence is 1 - .139 = .861.
- Odds of being given a custodial sentence is $.139/.861 = .161$.
- Odds of not being given a custodial sentence is 6.211.

Odds ratios

To compare males and females:

- Odds ratio of males to females receiving a custodial sentence $= .416/.161 = 2.58$.
  - Males are 2.6 times more likely to receive a custodial sentence than females.
- Odds ratio of females to males receiving a custodial sentence $= .161/.416 = .387$.
  - Females are 38% as likely to receive a custodial sentence as males.

Logistic regression

- Odds ratios are used to calculate probabilities in a regression.
- Predictor variables can be either continuous or categorical.
- There is no assumption of linearity between variables.
- It is however sensitive to high correlation between predictor variables (multicollinearity).

Example - Youthcohort Dataset

- Taken from Quantitative Data Analysis in Education, Paul Connolly.
- Dataset Descriptor: `http://cw.routledge.com/textbooks/9780415372985/pdfs/youthcohort.pdf`.
- Research question: What factors predict the likelihood that a student will answer yes when asked if they sat their maths GCSE?

Example - Variables

- Sat Maths (satmath): 0 = no, 1 = yes.
- Respondent gender (s1gender): 1 = male, 2 = female.
- Highest parental qualification: 1 at least one parent with a degree, 2 at least one parent with an A Level, 3 neither parent with an A Level.

Example

You need:

- One categorical dependent variable - satmath.
- Two or more continuous or categorical variables: s1gender, s1pared (both categorical).

Basic Model

```
# Assume we have read the data into a
# data frame mydata

logmodel1 <- glm(satmath ~ s1gender+s1pared,
     data = mydata, na.action = na.exclude,
     family = binomial())
```

Interpreting the output

- The baseline/null model is the baseline comparator to which any model is compared.
- Predictions of this baseline model are made purely on whichever category occurred most often in our dataset.
- Our aim is to improve this prediction by including our additional variables.

Baseline/Null Model

$logmodel1 < -glm(satmath \sim 1, data = mydata, family = binomial())$

Interpreting the output

- **Omnibus** test of model is used to check that the new model (with explanatory variables included) is an improvement over the baseline model.
- It uses chi-square tests to see if there is a significant difference between the baseline model (null) and the model created.
- The Chi-square statistic is 39.34, df = 3, which is statistically significant ($p < 0.001$).

```
> lmtest::lrtest(logmodel1)
Likelihood ratio test

Model 1: satmath ~ s1gender + s1pared
Model 2: satmath ~ 1
  #Df  LogLik Df Chisq Pr(>Chisq)
1   4 -1199.8
2   1 -1219.5 -3 39.34   1.47e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpreting the output

Deviance residuals, are a measure of model fit.
Used as part of the calculation of Chi square:

```
#Chi-square
modelChi <- logmodel1$null.deviance - logmodel1$deviance
#p-value
chidf <- logmodel1$df.null - logmodel1$df.residual
chisq.prob <- 1 - pchisq(modelChi, chidf)
chisq.prob
```

```
> summary(logmodel1)

Call:
glm(formula = satmath ~ s1gender + s1pared, family = binomial(),
    data = mydata, na.action = na.exclude)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.0419   0.1540   0.1708   0.2188   0.2401

Coefficients:
                                        Estimate Std. Error z value Pr(>|z|)
(Intercept)                               4.4287     0.1685  26.277  < 2e-16 ***
s1genderFemale                            0.1879     0.1292   1.454    0.146
s1paredAt least one parent with A-level  -0.2087     0.2268  -0.920    0.358
s1paredNeither parent with A-level       -0.8967     0.1739  -5.155 2.53e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2439.0  on 13200  degrees of freedom
Residual deviance: 2399.6  on 13197  degrees of freedom
AIC: 2407.6
```

Interpreting the output

```
> summary(logmodel1)

Call:
glm(formula = satmath ~ s1gender + s1pared, family = binomial(),
    data = mydata, na.action = na.exclude)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.0419   0.1540   0.1708   0.2188   0.2401

Coefficients:
                                    Estimate Std. Error z value Pr(>|z|)
(Intercept)                           4.4287     0.1685  26.277  < 2e-16 ***
s1genderFemale                        0.1879     0.1292   1.454    0.146
s1paredAt least one parent with A-level  -0.2087     0.2268  -0.920    0.358
s1paredNeither parent with A-level       -0.8967     0.1739  -5.155 2.53e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2439.0  on 13200  degrees of freedom
Residual deviance: 2399.6  on 13197  degrees of freedom
AIC: 2407.6

Number of Fisher Scoring iterations: 7
```

The AIC loss function $(2k - 2 * log(L))$ tries handle the bias when fitting a model. When you fit a model if you increase the number of parameters you will improve the log likelihood but will run into the danger of over fitting. AIC adjusts for increasing the number of parameters. Minimizing the AIC selects the model where the improvement in log likelihood is not worth the penalty for increasing the number of parameters.

Interpreting the output

```
> summary(logmodel1)

Call:
glm(formula = satmath ~ s1gender + s1pared, family = binomial(),
    data = mydata, na.action = na.exclude)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.0419   0.1540   0.1708   0.2188   0.2401

Coefficients:
                                  Estimate Std. Erro z value Pr(>|z|)
(Intercept)                         4.4287     0.168  26.277  < 2e-16 ***
s1genderFemale                      0.1879     0.129   1.454    0.146
s1paredAt least one parent with A-level  -0.2087     0.226  -0.920    0.358
s1paredNeither parent with A-level       -0.8967     0.173  -5.155 2.53e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2439.0  on 13200  degrees of freedom
Residual deviance: 2399.6  on 13197  degrees of freedom
AIC: 2407.6

Number of Fisher Scoring iterations: 7
```

The z statistic or Wald statistic is calculated to get the significance of the predictor.

We can see that s1gender is not statistically significant and only one level of s1parend is significant.

Interpreting the output

```
> summary(logmodel1)

Call:
glm(formula = satmath ~ s1gender + s1pared, family = binomial(),
    data = mydata, na.action = na.exclude)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.0419   0.1540   0.1708   0.2188   0.2401

Coefficients:
                                    Estimate Std. Error z value Pr(>|z|)
(Intercept)                          4.4287     0.1685  26.277  < 2e-16 ***
s1genderFemale                       0.1879     0.1292   1.454    0.146
s1paredAt least one parent with A-level  -0.2087     0.2268  -0.920    0.358
s1paredNeither parent with A-level       -0.8967     0.1739  -5.155 2.53e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2439.0  on 13200  degrees of freedom
Residual deviance: 2399.6  on 13197  degrees of freedom
AIC: 2407.6

Number of Fisher Scoring iterations: 7
```

What does this mean?

Being female, the log odds (estimate) of answering yes to whether you sat maths at GCSE increase by 0.1879 in comparison to being male but this NOT statistically significant.

Interpreting the output

```
> summary(logmodel1)

Call:
glm(formula = satmath ~ s1gender + s1pared, family = binomial(),
    data = mydata, na.action = na.exclude)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.0419   0.1540   0.1708   0.2188   0.2401

Coefficients:
                                    Estimate Std. Error z value Pr(>|z|)
(Intercept)                           4.4287     0.1685  26.277  < 2e-16 ***
s1genderFemale                        0.1879     0.1292   1.454    0.146
s1paredAt least one parent with A-level -0.2087     0.2268  -0.920    0.358
s1paredNeither parent with A-level   -0.8967     0.1739  -5.155 2.53e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2439.0  on 13200  degrees of freedom
Residual deviance: 2399.6  on 13197  degrees of freedom
AIC: 2407.6

Number of Fisher Scoring iterations: 7
```

What does this mean?

Having neither parent with an A level, the log odds of answering yes to whether you sat maths at GCSE maths decrease by -0.8967 in comparison to having at least one parent with a degree and this is statistically significant.

Interpreting the output

```
> summary(logmodel1)

Call:
glm(formula = satmath ~ s1gender + s1pared, family = binomial(),
    data = mydata, na.action = na.exclude)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.0419   0.1540   0.1708   0.2188   0.2401

Coefficients:
                                    Estimate Std. Error z value Pr(>|z|)
(Intercept)                           4.4287     0.1685  26.277  < 2e-16 ***
s1genderFemale                        0.1879     0.1292   1.454    0.146
s1paredAt least one parent with A-level  -0.2087     0.2268  -0.920    0.358
s1paredNeither parent with A-level    -0.8967     0.1739  -5.155 2.53e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2439.0  on 13200  degrees of freedom
Residual deviance: 2399.6  on 13197  degrees of freedom
AIC: 2407.6

Number of Fisher Scoring iterations: 7
```

What does this mean?

Having at least one parent with an A level , the log odds of answering yes to whether you sat maths GCSE decrease by -0.2087 in comparison to having at least one parent with a degree and this is NOT statistically significant

Interpreting the output

$$P(Y) = \frac{1}{1 + e^{b_0 + b_1 X_1 + b_2 X_2}}$$

- You can generate the equation.
- b_0 is the intercept.
- b_1 is the coefficient of the first predictor etc ...

The odds ratio: exp(B)

$$Odds\ ratio = \frac{Odds\ after\ a\ unit\ change\ in\ the\ predictor}{Original\ odds}$$

Indicates the change in odds resulting from a unit change in the predictor.

- $OR > 1$: Predictor $\uparrow$, Probability of outcome occurring $\uparrow$.
- $OR < 1$: Predictor $\uparrow$, Probability of outcome occurring $\downarrow$.

Interpreting the output

```
exp(coefficients(logmodel1))
(Intercept) 83.8207195
s1genderFemale 1.2067584
s1paredAt least one parent with A-level 0.8116561
s1paredNeither parent with A-level 0.4079235
```

You can exponentiate the coefficients and interpret them as odds-ratios.

Now we can say that being female, the odds of answering yes to sitting maths are 1.20 when compared to being than being male.

And having neither parent with an a level, the odds of answering yes to sitting maths are 0.4079 when compared to having at lease one parent with a degree.

Interpreting the output

```
exp(coefficients(logmodel1))
(Intercept) 83.8207195
s1genderFemale 1.2067584
s1paredAt least one parent with A-level 0.8116561
s1paredNeither parent with A-level 0.4079235
```

The intercept represents the odds of answering yes to having sat maths for a male respondent having at least one parent with a degree.

Interpreting the output

```
#Pseudo Rsquared plus Chi-square of the model
rcompanion::nagelkerke(logmodel1,restrictNobs=TRUE)
```

```
## $Models
##
## Model: "glm, satmath ~ s1gender + s1pared, binomial(), mydata, na.exclude"
## Null:  "glm, satmath ~ 1, binomial(), fit$model, na.exclude"
##
## $Pseudo.R.squared.for.model.vs.null
##                              Pseudo.R.squared
## McFadden                           0.01613010
## Cox and Snell (ML)                 0.00297568
## Nagelkerke (Cragg and Uhler)       0.01763970
##
## $Likelihood.ratio.test
##  Df.diff LogLik.diff Chisq  p.value
##       -3      -19.67 39.34 1.47e-08
##
## $Number.of.observations
##
## Model: 13201
## Null:  13201
##
## $Messages
## [1] "Note: For models fit with REML, these statistics are based on refitting
 with ML"
##
## $Warnings
## [1] "None"
```

How useful is the model?

Cox and Snell R^2 and **Nagelkerke** R^2 are pseudo R^2 statistics - in this case they indicate that between 0.29% and 1.76% of the variance in the variability of responses to they sat maths at GCSE.

Interpreting the model

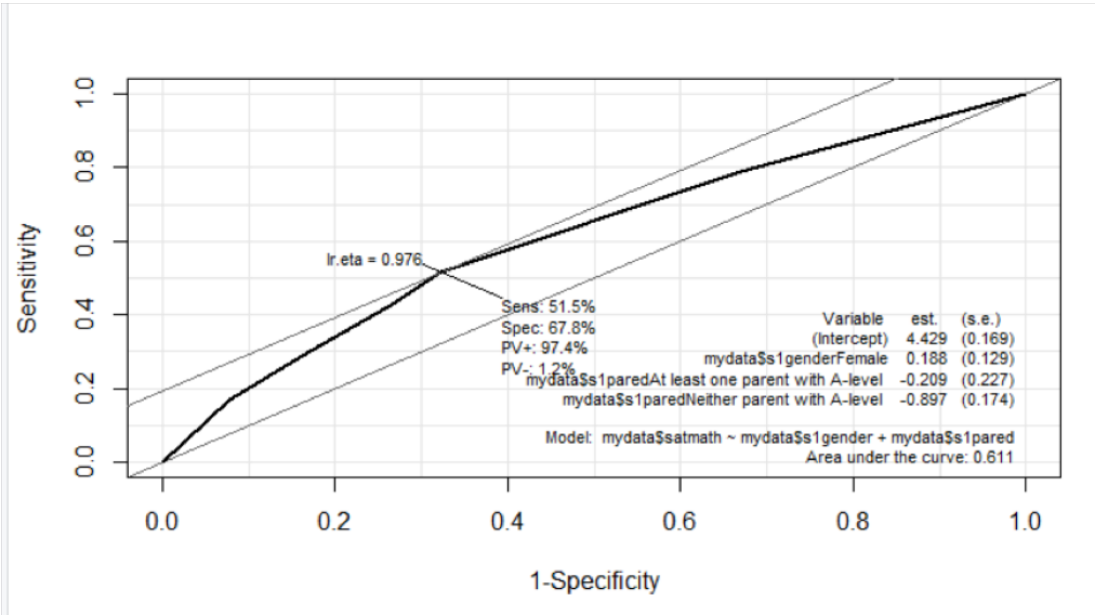
A confusion matrix/classification matrix is a table that is often used to **describe the performance of a classification model** on a set of test data for which the true values are known.

Interpreting the output

- **Classification Table/Confusion Matrix** provides us with an indication of how well the model can predict the correct category for each case.
- Compare this to the classification for null to see the improvement.

Sensitivity

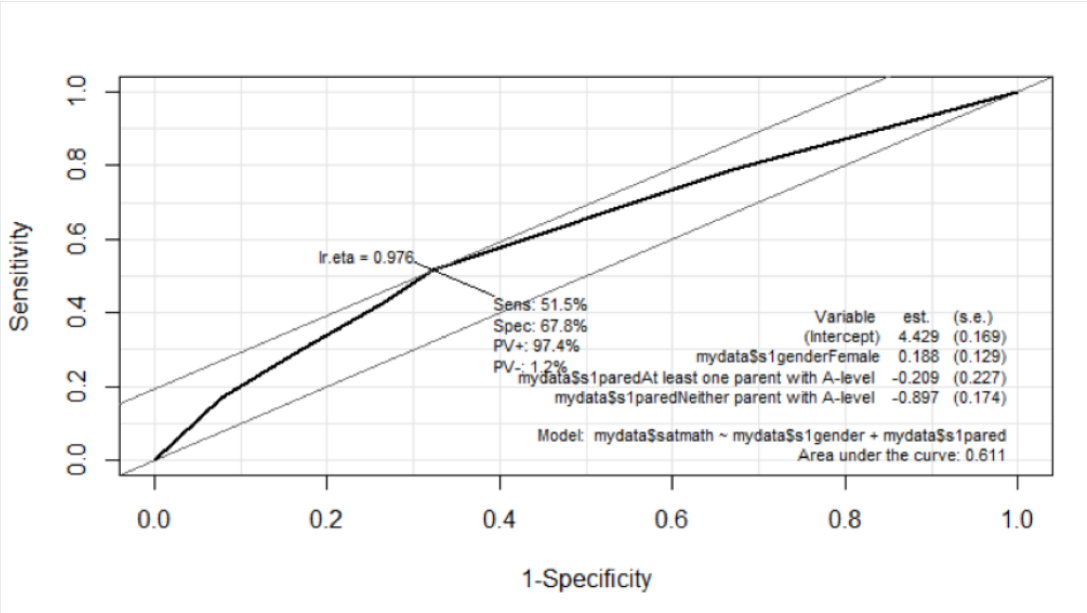
```
#Output the sensitivity, specificity, and ROC plot
Epi::ROC(form=mydata$satmath ~ mydata$s1gender+mydata$s1pared,
    plot="ROC")
```



- The **sensitivity** of the model is the percentage of the group with the characteristic of interest that has been accurately identified by the model.
- Those with "yes" for sitting mathematics correctly predicted by the model (i.e., true positives).
- In this case the model was able to identify 51.5% of people who said they sat

Specificity

```
#Output the sensitivity, specificity, and ROC plot
Epi::ROC(form=mydata$satmath ~ mydata$s1gender+mydata$s1pared,
    plot="ROC")
```



- The **specificity** of the model is the percentage of the group without the characteristic of interest.
- Those with "no" for sitting mathematics and were also correctly predicted as not having the observed characteristic (i.e., true negatives).
- In this case 67.8%.

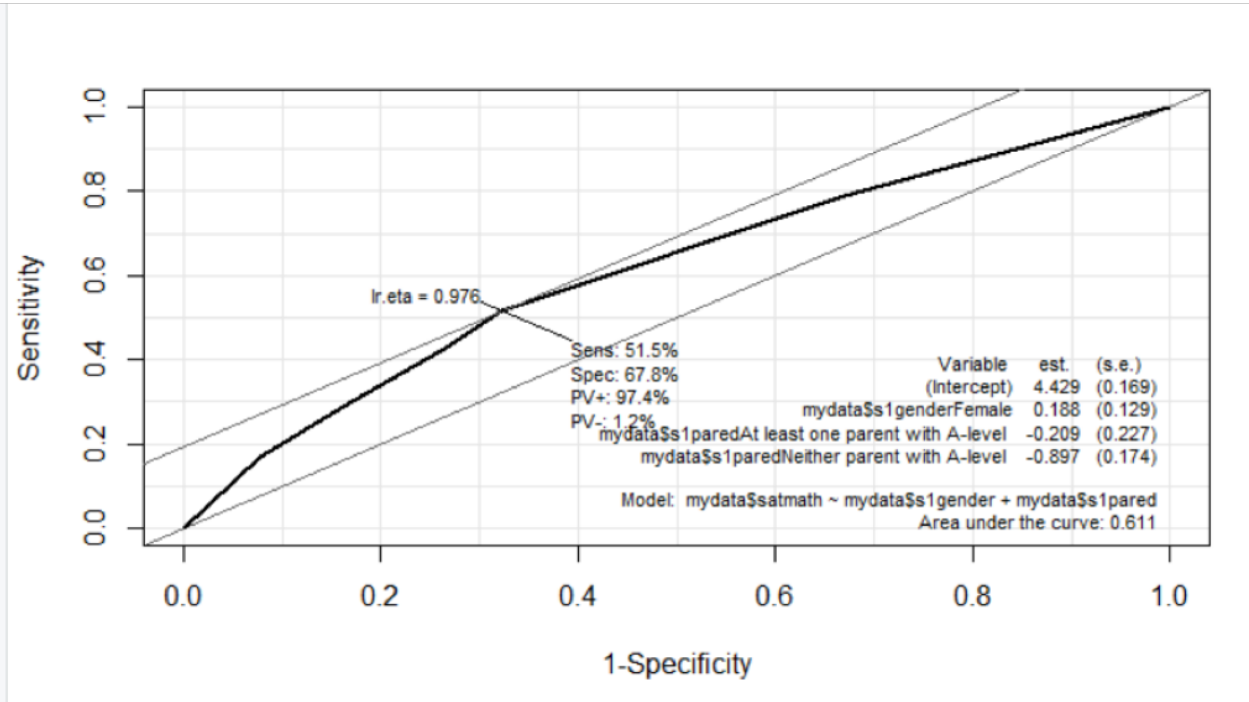
Interpreting the output

- The positive predictive value is the percentage of the cases the model classifies as having the characteristic that is actually observed in this group.
  - Calculate by dividing the number of predicted cases yes/ the total number predicted (97.4%)
- The negative predictive value is the percentage of the cases the model classifies as not having the characteristic that is actually observed in this group. (1.2%)

ROC Curve

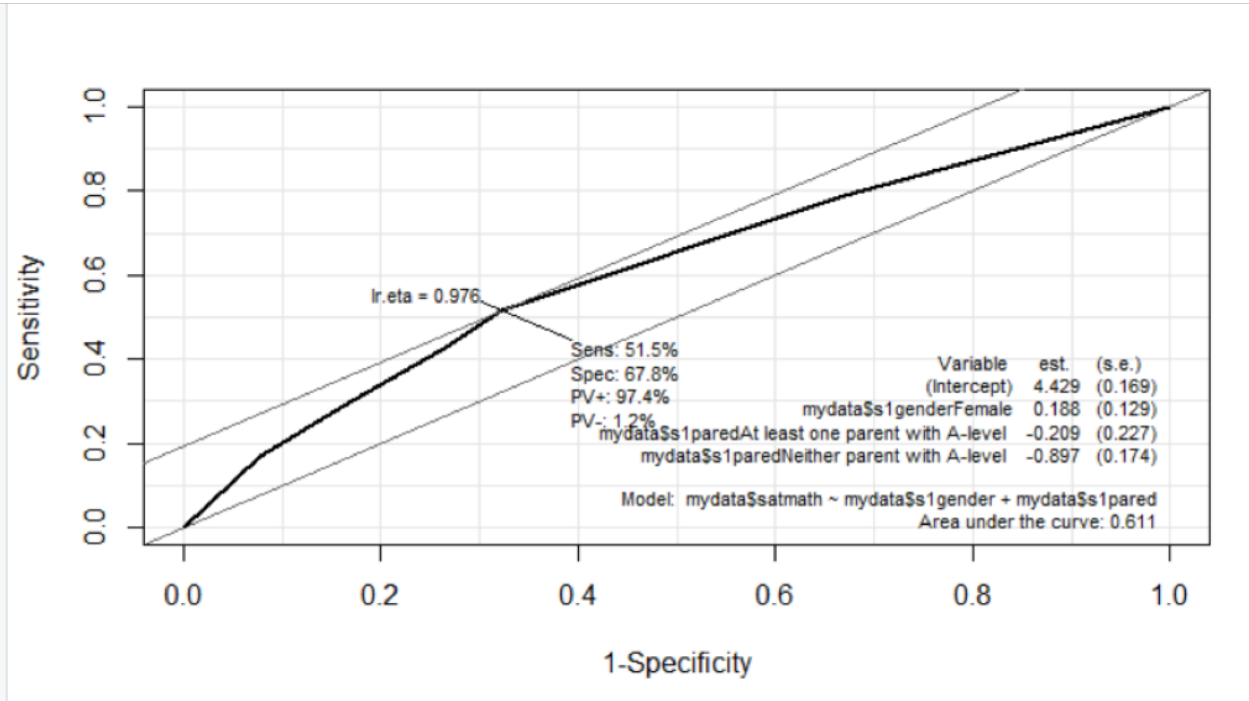
- Receiver Operating Characteristic.
  - Originates from sonar in the 1940s.
  - Were used to measure how well a sonar signal (e.g., from an enemy submarine) could be detected from noise (a school of fish).
- Used to see how any predictive model can distinguish between the true positives and negatives by plotting sensitivity, the probability of predicting a real positive will be a positive, against 1-specificity, the probability of predicting a real negative will be a positive.
- The ROC curve plots out the sensitivity and specificity for every possible decision rule cutoff between 0 and 1 for a model.

ROC



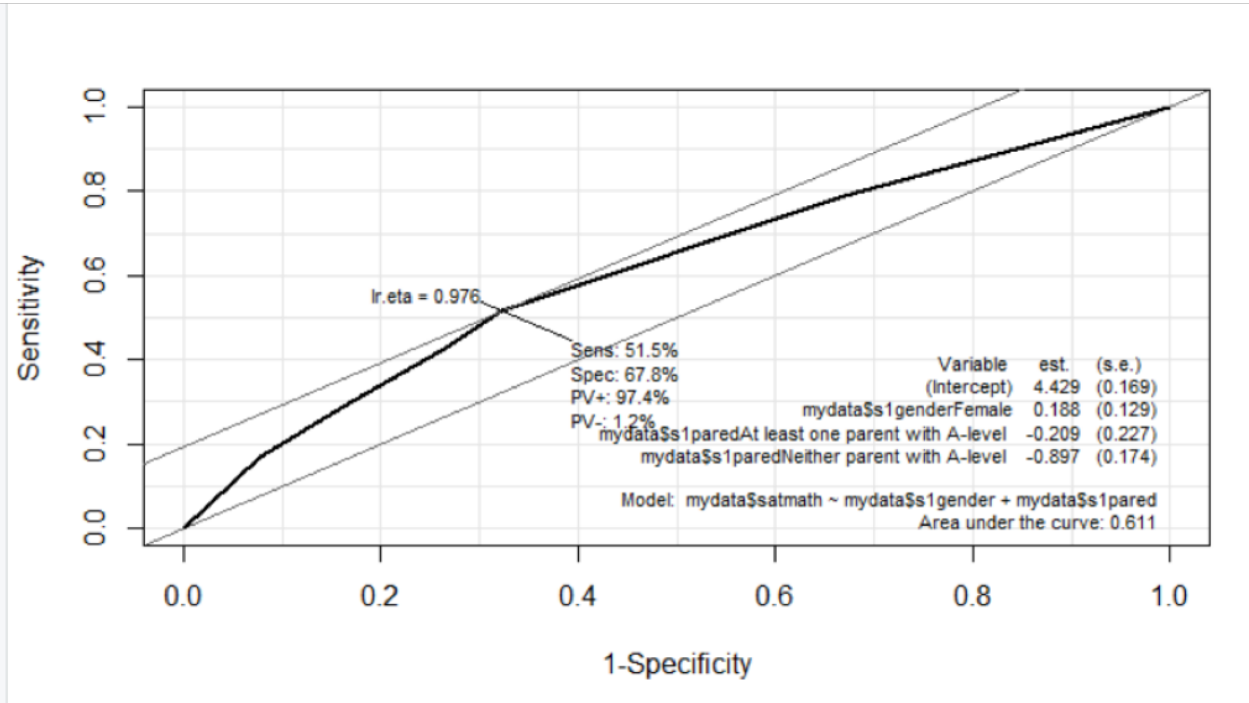
A model that predicts at chance will have an ROC curve that looks like the lower diagonal line.

ROC



The further the curve is from the lower diagonal line, the better the model is at discriminating between positives and negatives in general.

ROC



To use the ROC curve to **quantify the performance of a classifier** use the **Area Under the Curve**. AUC is literally just the percentage of this graph that is under this curve. This AUC is 0.61 (.5 is considered weak, 0.8 strong).

Interpreting the output

- Reporting odds ratios for categorical predictors we must remember that we are comparing the odds for two categories.
- For more than two we are comparing each category to the reference category (value 0) - which we specify in the regression.
- For odds ratios less than 1 we can invert these and report the inversion to help interpretation.

Multinomial logistic regression

- Logistic regression to predict membership of more than two categories.
- It (basically) works in the same way as binary logistic regression.
- The analysis breaks the outcome variable down into a series of comparisons between two categories.
  - E.g., if you have three outcome categories (A, B and C), then the analysis will consist of two comparisons that you choose:
  - Compare everything against your first category (e.g. A vs. B and A vs. C),
  - Or your last category (e.g. A vs. C and B vs. C),
  - Or a custom category (e.g. B vs. A and B vs. C).
- The important parts of the analysis and output are much the same as we have just seen for binary logistic regression.

Assumptions Logistic Regression

- Dependent variable must be binary for binomial, ordinal/multinomial for multinomial.
- Observations to be independent of each other. In other words, the observations should not come from repeated measurements or matched data.
- Little or no multicollinearity among the independent variables. This means that the independent variables should not be too highly correlated with each other.
    - Assess as you would for linear using vif from the car package.
- Assumes linearity of independent variables and log odds. This does not require the dependent and independent variables to be related linearly, it requires that the independent variables are linearly related to the log odds.
    - Box-Tidwell test (in the car package) can be used.
- Typically requires a large sample size. A general guideline is that you need at minimum of 10 cases with the least frequent outcome for each independent variable in your model. For example, if you have 5 independent variables and the expected probability of your least frequent outcome is .10, then you would need a minimum sample size of 500 (10*5 / .10).

15 mins break

Lab Exercise

- Use the youthcohort dataset.
- Check how many people sat math and how many didn't.
- Create a logistic regression model to predict satmath based on gender.
- Produce summary data of the model with stargazer and summery.
- Perform a likelihood ratio test.
- Test for pseudo R^2 and chi square.
- Exponentiate coefficients.
- Create confusion matrix.
- Plot a ROC curve and look for specificity and sensitivity.