Regression Commands

Chris Mainey

14/10/2019

Tutorial Cheatsheet

The tutorial on regression model uses many different command, but most of them are simple and based on default `R` objects and methods. Never-the-less, it takes time to learn, search and try these commands out, so this document is a summary of the commands show in the slides. The will use the `LOS_model` dataset from the `NHSRdatasets` package, but will not usually show the output of each command.

You can install/load the package and data with:

```r
install.packages("NHSRdatasets")

library(NHSRdatasets)
data(LOS_model)
```

Correlation

Although we are not using correlation in the practical, it is mentioned in the slides.

```r
# Correlation coefficient for Age and LOS
cor(LOS_model$Age, LOS_model$LOS)

# With a significane test:
cor.test(LOS_model$Age, LOS_model$LOS)

# Correaltion matrix plot of the data. frame (numeric columns 3 - 5)
library(corrplot)

M <-cor(LOS_model[3:5])
corrplot(M, type="upper", order="hclust")
```

Linear regression

Linear regression is the building block for what we are looking at, and we will treat regressions as 'object', i.e. assign each regression to a variable.

```r
#Linear model
lm(LOS ~ Age, data=LOS_model)

# Linear model assigned to variable, this includes two parameters,
# age and death, with death as a categorical
mod1 <- lm(LOS ~ Age + factor(Death), data=LOS_model)

# View model summary
summary(mod1)

# Exctract coefficients
coef(mod1)
```

```r
# calculate confidence interval
confint(model1)

# plot the residuals
plot(mod1)


# interaction model
mod2 <- lm(LOS ~ Age *factor(Death), data=LOS_model)

# Predict from your model
predict(mod2, newdata=my_newdata)
```

Generalized Linear Models (GLM)

GLMs share most of the architecture of linear models, with the additional 'family' argument

```r
# GLM to nodel death as binary (binomial).  We'll also mean-centre and scale then
mod3 <- glm(Death ~ scale(Age) + scale(LOS), data=LOS_model, family="binomial")

# Extract AIC
aic(mod3)

# predict on scasle of the data
predict(mod3, newdata=my_data, type = "response")
```