Module 11

Regression

- **lm(formula, data, subset, weights, na.action, method = 'qr', model = TRUE, x = FALSE, y = FALSE, qr = FALSE, ..)**

Fits a linear model to the given data and is used for linear regression. Returns the coefficients of the fit. The arguments are:

- **formula –** an object of class 'formula', which is a symbolic description of the model to be fitted (essentially, the model description in mathematical terms)
- **data –** an optional dataframe or list. If not specified, the arguments specified in ***formula*** are taken as variables by default
- **subset –** an optional vector specifying the subset of data values to be used in the fitting
- **weights –** an optional vector of weights to be used in the fitting process. Defaults to NULL, but if specified, uses a weighted least squares process to fit the model
- **na.action –** a function that indicates what should happen to NA values in the fitting process. The ***action*** values are:
    - **na.fail –** the regression fails
    - **na.omit –** excludes NA values
    - **na.exclude –** similar to na.omit, but behaves differently only when used with other functions computing residuals and predictions. It corrects for the vector lengths when these operations are conducted
    - **NULL**
- **method –** the fitting method ***'qr'*** is the default and is widely applicable
- **model, x, y, qr –** If TRUE, the function returns these components of the fit
- **linearHypothesis(model,…)**

Generic function for testing a linear hypothesis for a variety of linear models. (NOTE: For mixed effects models, the default test is the Chi-Square test for testing fixed effects).

- **ivreg(formula, instruments, data, subset, na.action, weights, offset, model = TRUE, y = TRUE, x = FALSE)**

Fit instrumental-variable regression by a two-stage least squares method. This is equivalent to direct instrumental-variables estimation when the number of instruments is equal to the number of predictors. The arguments are:

- **formula, instruments** formula specification(s) of the regression relationship and the instruments. Either **instruments** are excluded and **formula** has three parts as in y ~ x1 + x2 | z1 + z2 + z3 (which recommended) or **formula** is y ~ x1 + x2 and **instruments** is a one-sided formula ~ z1 + z2 + z3

- **data** an optional data frame containing the variables in the model. By default the variables are taken from the environment of the formula.
- **subset** an optional vector specifying a subset of observations to be used in fitting the model.
- **na.action** a function that indicates what should happen when the data contain NAs. The default is set by the **na.action** option.
- **weights** an optional vector of weights to be used in the fitting process.
- **offset** an optional offset that can be used to specify an *a priori* known component to be included during fitting.
- **model, x, y** logicals. If TRUE the corresponding components of the fit (the model frame, the model matrices , the response) are returned.