# Bootstrap methods

1. The statistic $F$ of a regression table is used as a test statistic in the so-called *ANOVA* test for regression, whose null hypothesis establishes that none of the covariates contains relevant (linear) information about the response variable, while the alternative is that some of them do. The formula of the statistics is:

$$\frac{SSR/k}{SSE/(n-k-1)} = \frac{\sum_{i=1}^{n}(\widehat{y}_i - \overline{y})^2/k}{\sum_{i=1}^{n}\widehat{\varepsilon}_i^2/(n-k-1)}$$

   **a)** What is the $F$-statistic of the *least squares* model that fits the prestige in terms of `income` and `education` in the Duncan dataset?

   `search.r-project.org/CRAN/refmans/aspect/html/duncan.html`

   Write a formula to obtain the $F$-statistic from the fitted values and the residuals.

   **b)** Simulate $B = 1000$ bootstrap replicates of the $F$ statistic (use the bootstrap pairs approach). Represent them in a histogram and estimate the standard error of the $F$-statistic.

   **c)** Simulate $B = 1000$ bootstrap replicates of the $F$-statistic (use the bootstrapping residuals approach). Represent them in a histogram and estimate the standard error of the $F$-statistic.

   **d)** What is the value of the $F$-statistic evaluated in the robust regression model? You can use several options. The simplest one is with the `rlm()` function.

   **e)** Repeat part (b) with the robust regression model.

   **f)** Repeat part (c) with the robust regression model.

2. Simulate 1200 observations of an *ARMA*(1,1) time series model with parameter values not very close to 1 in absolute value. Estimate the model parameters using a maximum likelihood method.

   Calculate the standard errors of the parameters using a bootstrap procedure based on residuals and another bootstrap procedure based on blocks. Using various block sizes, compare the results.

   Use several alternatives or programs to do this, *by hand* and with the `boot` library.