Assignment for
EP16: Missing Values in Clinical Research

Multiple Imputation

14 – 18 May, 2018

Data

The **MIdat11** data comprise data of 816 children and their mothers on vitamin D exposure of the mother during pregnancy and child bone health, measured by DXA scan, at 6 years of age. Maternal serum samples were taken in the third trimester of pregnancy.

The dataset contains the following variables:

variable	explanation
`ID`	subject identifier
`sex`	child's sex
`vitD`	mother's serum vitamin D concentrations in 10 nmol/L
`weight`	child's total weight in kg at DXA scan
`BMC`	bone mineral content of the child in grams, determined by DXA scan
`season`	season of blood sampling
`bdate`	child'd birth date
`singleton`	is the child a singleton birth?
`length`	child's length in cm at time of DXA scan
`sun`	average sun light duration in minutes/day in the month before blood sampling
`sun_birth`	average sun light duration in the month before birth in hours/day
`sports`	does the child play sports regularily?
`leanfrac`	proportion of lean mass (lean mass/total mass; lean mass = total mass - fat mass)
`birthwgt`	birthweight (standard deviation score)
`ethn`	child's ethnicity
`lean`	lean mass in kg

Analysis model of interest

The analysis model of interest is a linear regression with outcome `BMC` and covariates `vitD`, `ethn`, `sex`, `leanfrac`, `sports`, `sun`, `season`, `length` and `weight`.

We assume that `vitD` has a non-linear (quadratic) effect.