

Assignment II for

EP16: Missing Values in Clinical Research

Multiple Imputation

13 – 17 May, 2019

Data

The **MI_{dat12}** data comprise 588 observations of 588 mothers measured repeatedly before and during pregnancy. All women were scheduled to have their weight measured once each trimester and were asked for their pre-pregnancy weight and BMI.

The dataset contains the following variables:

variable	explanation
ID	subject identifier
gender	gender of the child
sun_birth	average sun light duration in the month before birth in hours/day
bdate	child'd birth date
lean	lean mass in kg
length	child's length at time of DXA scan in meters
ancestry	child's ancestry
sun	average sun light duration in minutes/day in the month before blood sampling
birthwgt	birthweight in kg
BMC	bone mineral content of the child in grams, determined by DXA scan
weight	child's total weight in kg at DXA scan
gravidity	number of times the mother has been pregnant (primigravida: this was the first pregnancy, multigravida: the mother had previous pregnancies)
leanfrac	proportion of child's lean mass (lean mass/total mass; lean mass = total mass - fat mass)
vitD	mother's serum vitamin D concentrations in 10 nmol/L
season	season of blood sampling
sports	does the child do sports regularly?

Analysis model of interest

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

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