

Assignment II for

EP16: Missing Values in Clinical Research

Multiple Imputation

13 – 17 May, 2019

Data

The **MIdata31** data comprise 1217 observations of 338 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
<code>id</code>	subject identifier
<code>gage</code>	gestational age at measurement (<code>gage</code> = 0 refers to a measurement before pregnancy)
<code>weight</code>	maternal weight
<code>stress</code>	self reported stress score (0 – 5)
<code>gestbir</code>	gestational age at birth
<code>sex</code>	child sex
<code>income</code>	household income
<code>parity</code>	number of pregnancies of more than 20 weeks the mother had (nulliparity: this was the first pregnancy, ≥ 1 child: mother had previous pregnancies)
<code>bd_mom</code>	birth date of the mother
<code>BMI</code>	self reported maternal BMI before pregnancy
<code>preterm</code>	was the baby born before 37 weeks of gestation (preterm) or later?
<code>visit_center</code>	was intake performed at the study center? (0: no, 1: yes)
<code>trimester</code>	trimester of measurement
<code>date_incl</code>	date of inclusion in the study
<code>smoke</code>	smoking behaviour of the mother during pregnancy
<code>kcal</code>	average daily kcal intake (calculated from food frequency questionnaire)
<code>educ</code>	educational level of the mother

Analysis model of interest

The analysis model of interest is a linear mixed model for weight with random intercept and slope for `gage`. Covariates are `smoke`, `kcal`, `stress`, `preterm`, `parity`, `educ` and `income`.