Assessing Effects of Exposures to DDE and PCBs on Premature Delivery via Ordinal Logistic Regression

Raphael Morsomme Rihui Ou Alessandro Zito

Case Study 1- Stat 723

January 20, 2020

Overview

Introduction

Data

The data reported the gestational age (in weeks) of the mother, the DDE and PCBs concentration, socio-economic info and scores (race, occupation, education and income), and amount of triglycerides and cholesterol. Total sample size (after preprocessing) = 2336

We contruct the following variables:

- Total level of lipids[1]

$$lipid = 2.27 * cholesterol_i + triglycerides_i + 0.623$$

- Gestational age group

$$gestgroup_i = \begin{cases} 0 & if \text{ Dangerous preterm} \\ 1 & if \text{ Preterm} \\ 2 & if \text{ At term} \end{cases}$$

- Average (standardized) PCB

$$PCB_i = \frac{1}{11} \sum_{j=1}^{1} 1 \frac{PCB_{ij} - mean_i(PCB_{ij})}{sd_i(PCB_{ij})}$$

Model (I) - Ordinal Logistic Regression

We run the following regression model

Model (II) - Bayesian Ordinal Logistic Regression

Results

Conclusions

Preprocessing:

- Drop obs. with gestational age > 45 (the world record)
- Standardize and average the differnt PCBs (to avoid their correlation)
- Mean impute of occupation, education and income scores $\Longrightarrow$ Total obs. $= 2336$