

Microbiome Simulation Analysis

Olajumoke Evangelina Owokotomo

6/13/2019

General information

- This is a simulation study to study the relationship between the probability of inclusion from the BVS approach, the R_h^2 from the information theory approach and the adjusted association from the joint moelling approach.

Simulation procedure (One dataset analysis)

For simulating the dataset the following steps and values were used;

- The joint error term for X and Y was assumed to have mean = (0,0) and covariance matrix = ε
$$\begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{bmatrix}$$
- The $\sigma_x^2 = \sigma_y^2 = 1$.
- $\rho_{xy} = \sigma_{xy}^2$ because $\sigma_x^2 = \sigma_y^2 = 1$.
- $\rho_{xy} = c(0.25, 0.50, 0.75, 0.90)$.
- $\varepsilon_1 = \begin{bmatrix} 1 & 0.25 \\ 0.25 & 1 \end{bmatrix}$, $\varepsilon_2 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$, $\varepsilon_3 = \begin{bmatrix} 1 & 0.75 \\ 0.75 & 1 \end{bmatrix}$, $\varepsilon_4 = \begin{bmatrix} 1 & 0.90 \\ 0.90 & 1 \end{bmatrix}$.
- The number of subjects in each treatment group is 50, therefore $Z_i = 100$.
-

$$\begin{aligned} X_i &= \alpha_0 + \alpha_1 Z_i + \epsilon_i^x \\ Y_i &= \beta_0 + \beta_1 Z_i + \epsilon_i^y \end{aligned}$$

- ϵ_i^x and ϵ_i^y were obtained using `rmvnorm(100, c(0,0), ε)` for each correlation.
- Using the following true value: $\alpha_0 = 1$, $\alpha_1 = 5$, $\beta_0 = 5$, $\beta_1 = 2$.
- The combination of the true parameter values and ϵ_i^x and ϵ_i^y was then used to simulate 1000 dataset for each correlation; we therefore have 1000 dataset for each $\text{cor}=0.25$, for $\text{cor}=0.5$, for $\text{cor}=0.75$ and for $\text{cor}=0.90$.
- Models for the Joint model approach

$$\begin{aligned} X_i &= \mu_i^x + \alpha Z_i + \epsilon_i^x \\ Y_i &= \mu_i^y + \beta Z_i + \epsilon_i^y \\ R^2 &= \text{corr}(\epsilon_i^x, \epsilon_i^y) \end{aligned}$$

- Models for the Information theory approach

$$\begin{aligned} Y_i^1 &= \mu_i^{y1} + \alpha Z_i + \epsilon_i^{y1} \text{ --- --- --- Model1} \\ Y_i^2 &= \mu_i^{y2} + \beta Z_i + \gamma X_i + \epsilon_i^{y2} \text{ --- --- --- Model2} \\ G^2 &= -2(\log 1 - \log 2) \\ R_h^2 &= 1 - e^{-\frac{G^2}{n}} \end{aligned}$$

- Models for the BVS approach

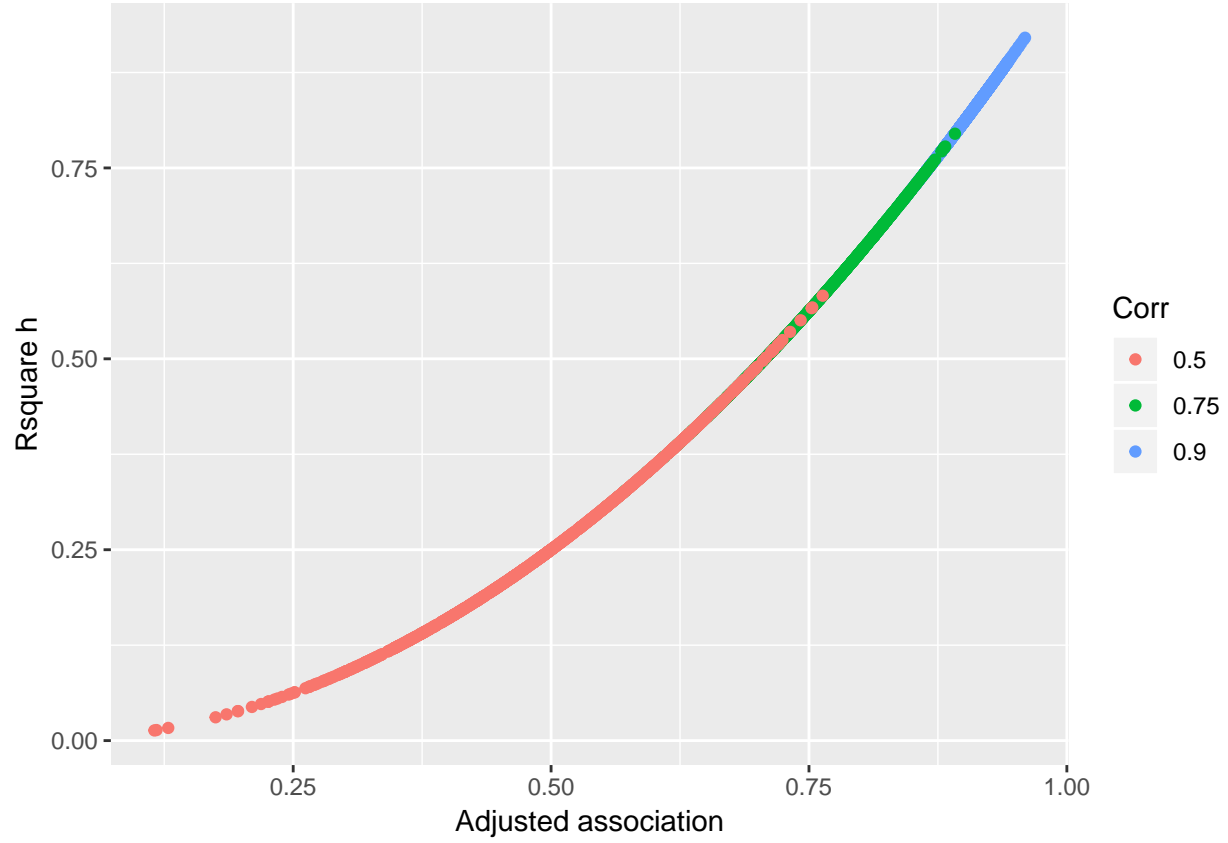
$$X_i = \mu_i^x + \theta_1 \phi_1 Z_i + \epsilon_i^x \quad \text{where} \quad \alpha = \theta_1 \phi_1$$

$$Y_i = \mu_i^y + \theta_2 \phi_2 Z_i + \theta_3 \phi_3 X_i + \epsilon_i^y \quad \text{where} \quad \beta = \theta_2 \phi_2 \quad \text{and} \quad \gamma = \theta_3 \phi_3$$

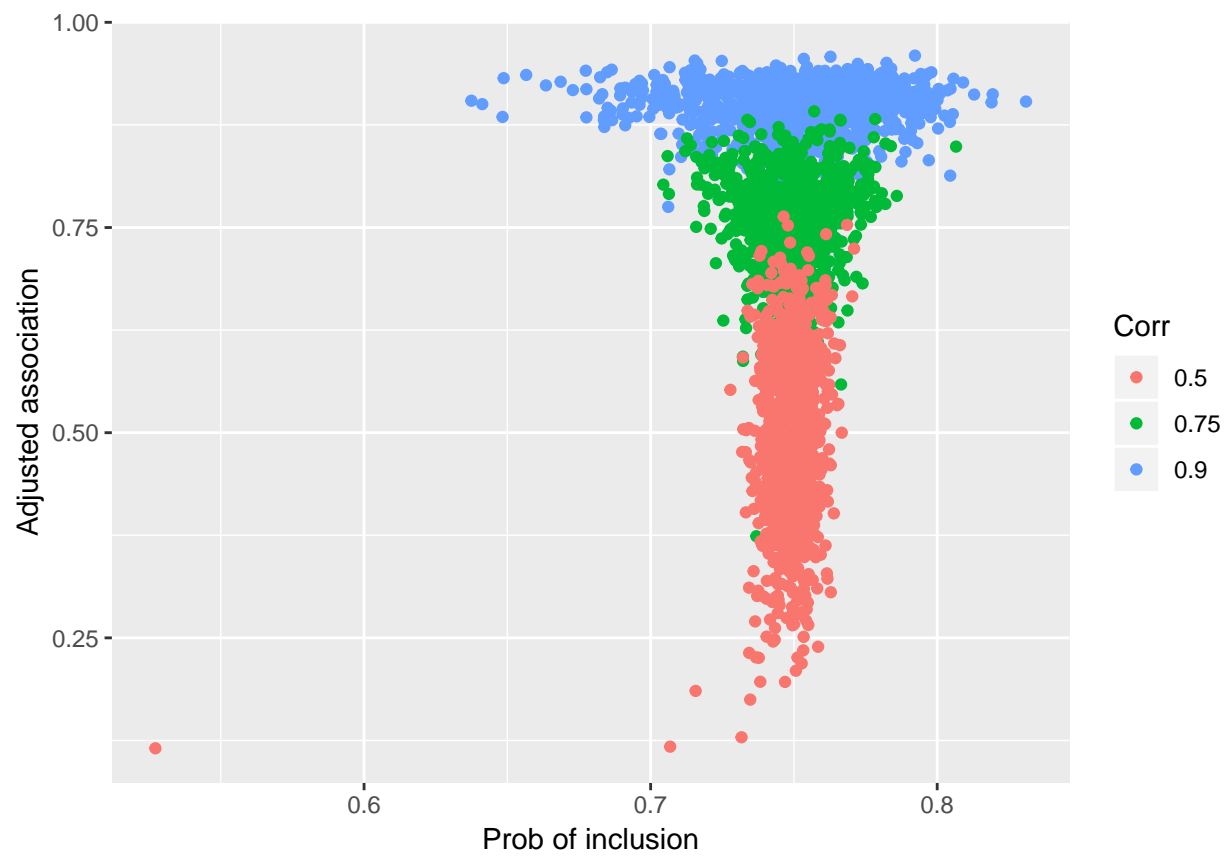
$$R_{BVS}^2 = P(\phi_3 = 1)$$

All timepoints

ADJUSTED ASSOCIATION AND INFORMATION APPROACH



PROBABILITY OF INCLUSION AND ADJUSTED ASSOCIATION



PROBABILITY OF INCLUSION AND INFORMATION APPROACH

