

# 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 =  $\varepsilon$   
$$\begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{bmatrix}$$
- 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  $Z_i = 100$ .
- 

$$X_i = \alpha_0 + \alpha_1 Z_i + \epsilon_i^x$$

$$Y_i = \beta_0 + \beta_1 Z_i + \epsilon_i^y$$

.

- $\epsilon_i^x$  and  $\epsilon_i^y$  were obtained using `rmvnorm(100, c(0,0),  $\varepsilon$ )` 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  $\epsilon_i^x$  and  $\epsilon_i^y$  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  $\text{cor}=0.90$ .