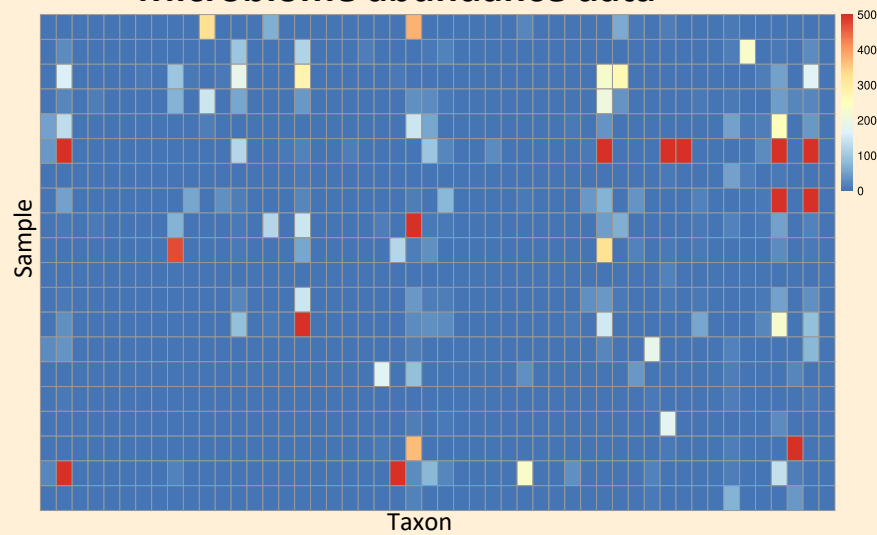


## Microbiome abundance data



compositionality  
sparsity  
over-dispersion

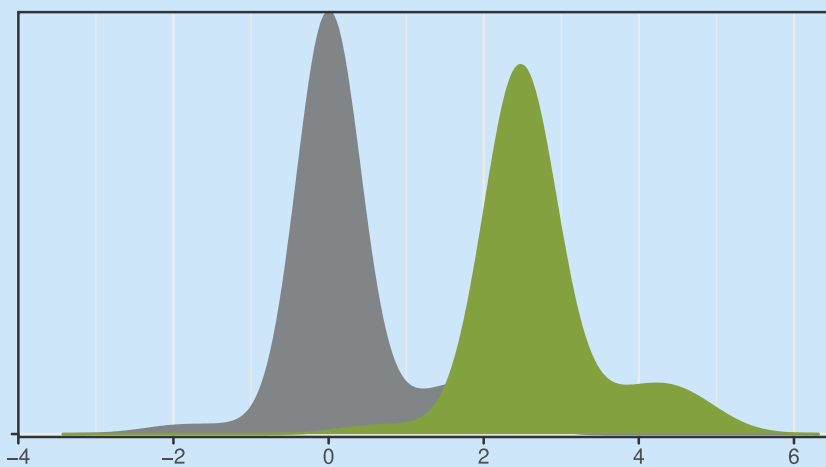
## ZINB model

$$Z_{ik} \sim \text{Bern}(\eta_{ik}),$$

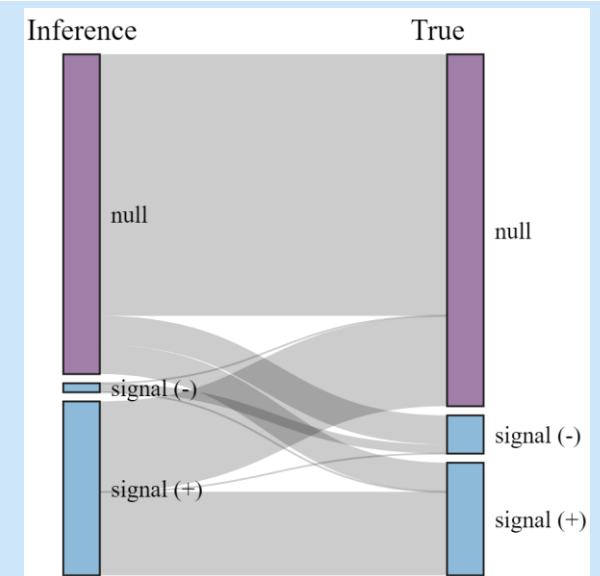
$$Y_{ik} | Z_{ik} \sim \begin{cases} 0, & Z_{ik} = 1, \\ \text{NB}(\mu_{ik}, \phi_k), & Z_{ik} = 0. \end{cases}$$

## EM algorithm

True signals Biased estimates

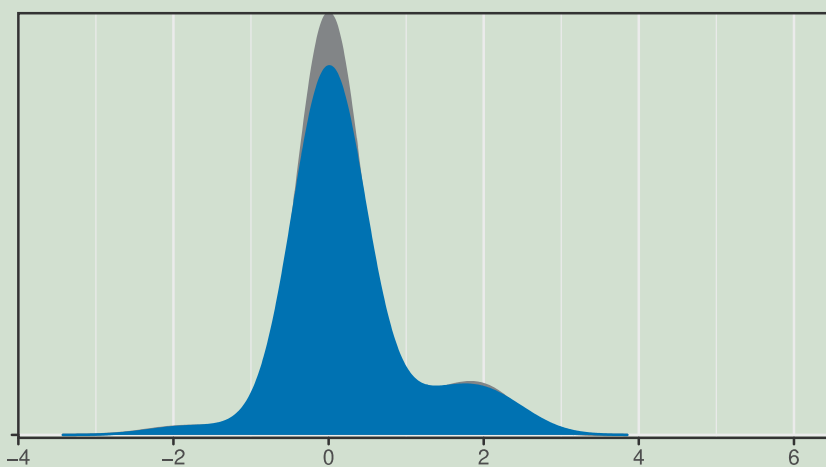


invalid inferences



## Bias correction

True signals Bias-corrected estimates



valid inferences

