

*NeuroData lab*

*Johns Hopkins University - Biomedical Engineering*

 [\*j1c@jhu.edu\*](mailto:j1c@jhu.edu)



*@j1c (Github)*



# Outline

- Background
  - **What is heritability?**
  - Graphs, networks, connectomes
  - Where do connectomes come from?
- Problem
  - What are we trying to estimate?
  - Causal models
  - Dcorr
  - Distance functions
- Results
  - Dataset
  - Results 1
  - Results 2
  - Results 3

# What is heritability?

- Variations in phenotype caused by variations in genotype.
- Example:

# Brains as connectomes

(aka networks or graphs)

- Vertex = a region of interest
- Edges = connectivity measure between a pair of vertices
- Diffusion MRI = # of estimated neuronal fibers

# Connectome Generation

 center

Source:

<https://www.biorxiv.org/content/biorxiv/early/2021/11/03/2021.11.01.466686.full.pdf>

# Overall DAG

Insert dag

# Subject specific DAGs

Insert 4 Dags

# Statistical problem

- Want an independence test!
- $H_0 : F(\text{Genome}, \text{Connectome}) = F(\text{Genome})F(\text{Connectome})$   
 $H_A : F(\text{Genome}, \text{Connectome}) \neq F(\text{Genome})F(\text{Connectome})$
- Test statistic: Distance correlation



# Statistical problem

- Want an independence test!
- $H_0 : F(\text{Genome}, \text{Connectome} | \text{Covariates}) = F(\text{Genome} | \text{Covariates})F(\text{Connectome} | \text{Covariates})$   
 $H_A : F(\text{Genome}, \text{Connectome} | \text{Covariates}) \neq F(\text{Genome} | \text{Covariates})F(\text{Connectome} | \text{Covariates})$
- Test statistic: Conditional distance correlation

# Human Connectome Project

-

# Assumptions

- Only 2 children per family
- Only identical twins, fraternal twins, non-twin siblings

# Monozygotic vs Dizygotic

- Assumptions:
  - Controls environment variable

Insert figure

# All three groups

- Assumptions:
  - Add in environmental and genetic variance

Insert figure

# Neuroanatomy Mediator

Test the existence of arrow

# Conditional Test as causal effect estimator

- Using conditional distance correlation

**The End**



Additional slides

# **Information on Distance Correlation**

# Shortcomings

- Network models
- Problems with connectome estimation.
- dominant genetic effects and epistasis.
- No interaction between environment and genetics

# Environmental effects

- Shared
  - Common experiences of siblings living in the same household.
    - household income, the family's living situation, the dynamics between the parents, food consumed
- Non-shared
  - Everything else
  - Epigenetics
  - Luck
  - schools, peers

