Law of Large Graphs

Michael Ketcha, Runze Tang, Daniel Sussman, Vince Lyzinski, Carey E. Priebe,
Presented by: Joshua T. Vogelstein

{BME, ICM, CIS, IDIES}@ Johns Hopkins University

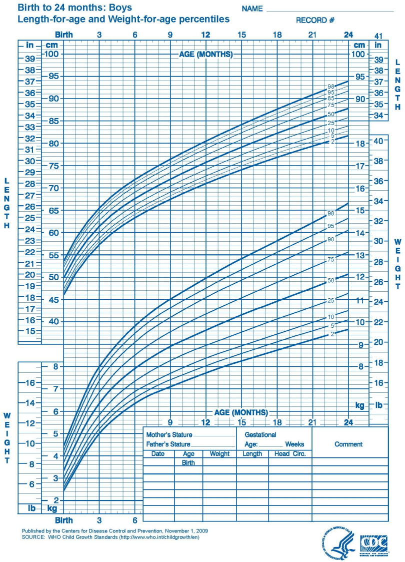
jovo@cis.jhu.edu

March 27, 2015

Overview

Table of Contents

Background



Birth to 24 months: Boys
Length-for-age and Weight-for-age percentiles
NAME
RECORD #
AGE (MONTHS)
LENGTH
WEIGHT
Mother's Stature
Father's Stature
Gestational Age: Weeks
Comment
Date
Age
Weight
Length
Head Circ.
Birth
Published by the Centers for Disease Control and Prevention, November 1, 2009
SOURCE: WHO Child Growth Standards (http://www.who.int/childgrowth/en)

News from the Human Connectome Project (HCP)

February 24, 2015

Now available for download are an interim report and associated slides summarizing findings from the HCP Lifespan Pilot project being conducted by the WU-Minn HCP consortium.

The ongoing HCP Lifespan Pilot is collecting multimodal imaging data acquired across the lifespan, in 6 age groups (4-6, 8-9, 14-15, 25-35, 45-55, 65-75) and using scanners that differ in field strength (3T, 7T) and maximum gradient strength (70-100 mT/m). The scanning protocols are similar to those for the WU-Minn Young Adult HCP, except shorter in duration.

The report should be of interest for those groups

Figure 7: Parcel-Wise Connectivity Matrices within Age

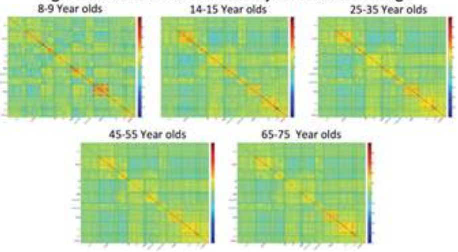


Table of Contents

Law of Large Numbers

A sample average converges almost surely to the expected value.

That is:

$$Pr\left(\lim_{n\to\infty} \bar{X}_n = \mu\right) = 1$$

Is there a Law of Large Graphs?

Independent Edge Model

Let us consider an edgewise probability matrix

$$P \in [0,1]^{n \times n}_{sym}$$

Then, conditioned on P, we define a symmetric adjacency matrix A

$$A_{ij} \stackrel{iid}{\sim} Bern(P_{ij}), i > j$$

Law of Large Graphs

Let $A^{(1)}, A^{(2)}, ...A^{(M)}$ be iid adjacency matrices each conditioned on a given P. Let

$$\bar{A}_M^{n\times n} = \frac{1}{M}\sum_{m=1}^{M} A^{(m)}$$

Then,

$$Pr\left(\lim_{M\to\infty} \bar{A}_M = P\right) = 1$$

Table of Contents

Can We Get Concentration Inequalities?

Concentration Inequalities

Hoeffding's Inequality

Let $Y_1, Y_2, ..., Y_n$ be iid observations such that $\mathbb{E}[Y_i] = \mu$ and $\forall i \in [n], a \leq Y_i \leq b$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

With probability $1 - \delta$

$$|\bar{Y} - \mu| \leq \sqrt{\frac{(b-a)^2}{2n} log\left(\frac{2}{\delta}\right)}$$

Concentration Inequalities

Hoeffding's Inequality for Graphs

$$\bar{A}^{n\times n} = \frac{1}{M}\sum_{m=1}^{M} A^{(m)}$$

Then with probability at least 1-η

$$\|\bar{A} - P\|_2 \leq 2\sqrt{\frac{\Delta}{M}\log\left(\frac{n}{\eta}\right)}$$

where $\Delta = \max_i P_{i\cdot}1$. [1]

[1] As an extension of Oliveira, R. I., Concentration of the adjacency matrix and of the laplacian in random graphs with independent edges. Arxiv preprint at http://arxiv.org/abs/0911.0600, 2010. 2910

Can We Improve the Finite Sample Bound?

Random Dot Product Graph

Suppose there are low-rank ($d < n$) latent positions $X^{n \times d} = [x_1, x_2, ..., x_n]^T$ such that:

$$A_{ij} \overset{iid}{\sim} Bern(\langle x_i, x_j \rangle)$$

This implies,

$$P = XX^T$$

Concentration Inequalities

Low-Rank Estimator: $\hat{P}$

By assuming that there exists a low rank ($d < n$) approximation for P, we have the eigen decomposition of $\bar{A}$:

$$\bar{A} = VSV^T$$

Let $\hat{S}^{d \times d}$ contain the d largest eigenvalues with eigenvectors $\hat{V}$

$$\hat{P} = \hat{X}\hat{X}^T, \text{ where } \hat{X} = \hat{V}\hat{S}^{1/2}$$

Variance for Naive Estimator $\bar{A}$

Since each $A^{(m)}$ is taken from a Bernoulli, we see that the element-wise variance is:

$$\mathbb{E}[(\bar{A} - P)^2_{ij}] = \frac{P_{ij}(1 - P_{ij})}{M}$$

Concentration Inequalities

Variance for $\hat{P}$

We have the element-wise variance of $\hat{P}$ to be:

$$\mathbb{E}[(\hat{P} - P)_{ij}^2] \approx \frac{2P_{ij}(1 - P_{ij})}{nM} \text{ for large n}$$

Concentration Inequalities

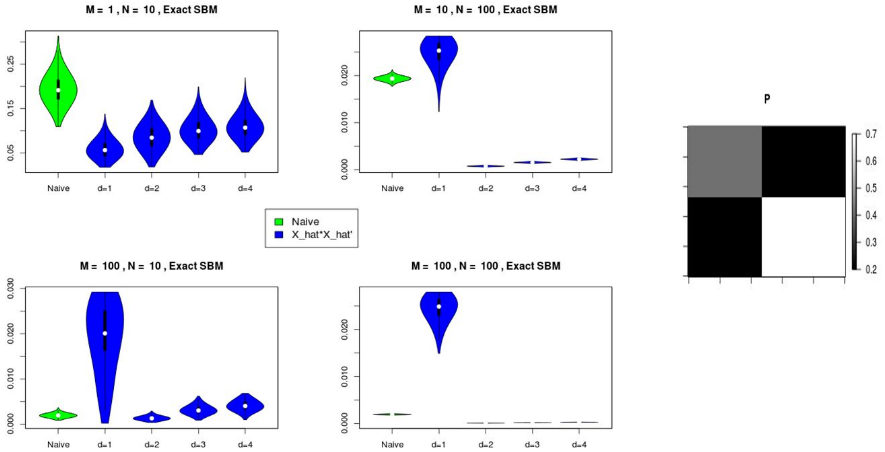
Main Result

The Relative Efficiency if the low-rank vs. naïve estimator is

$$\frac{\mathbb{E}[(\hat{P} - P)_{ij}^2]}{\mathbb{E}[(\bar{A} - P)_{ij}^2]} \approx 2/n$$

Numerical Simulations

Performance Under Mean Squared Error
Using a rank-2 P matrix



Numerical Simulations

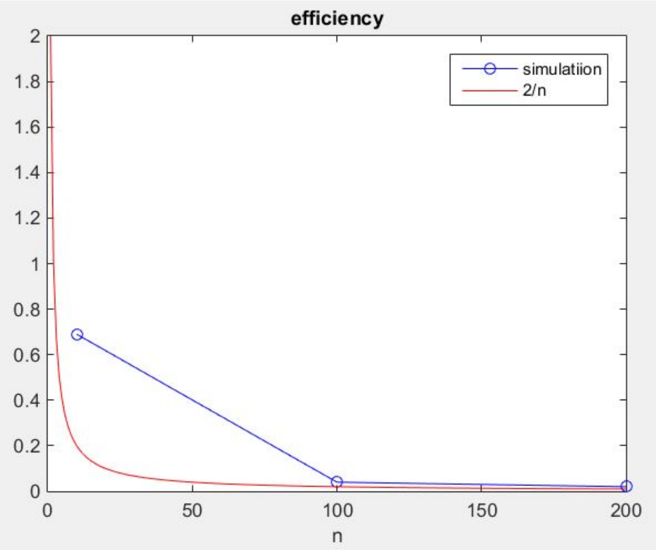


Table of Contents

Can We Get a Robust Estimator?

Hodges-Lehmann Estimator

- A robust and consistent measurement for the population mean.
- Let $Y_1, Y_2, ..., Y_n$ be iid observations
- Calculate the mean of $n(n+1)/2$ pairs:

$$B_{ij} = (Y_i + Y_j)/2$$

- Take the median

$$Y_{HL}(Y_1, Y_2, ..., Y_n) = \text{median}(B_{ij})$$

Robust Estimators

Poisson Weighted RDPG

Given the latent positions X

$$P = XX^T$$

$$A_{ij}^{(m)} \stackrel{iid}{\sim} Poisson(P_{ij})$$

$$\stackrel{iid}{\sim} Poisson(\langle x_i, x_j \rangle)$$

Contaminated Model

The latent position x_i associated with the contaminated vertex i follows a uniform distribution of the scaled feasible region $c \cdot \mathcal{S}$

$$x_i' \stackrel{iid}{\sim} (1 - \epsilon) \cdot x_i + \epsilon \cdot \mathcal{U}(c \cdot \mathcal{S})$$

$$P = X' X'^T$$

Robust Estimators

Hodges-Lehmann Estimator For Graphs: $\tilde{A}$

Calculate the average of $M(M+1)/2$ pairs of graphs:

$$B^{(ij)} = (A^{(i)} + A^{(j)})/2$$

Take the element-wise median

$$\hat{A}_{HL}(A^{(1)}, ..., A^{(M)}) = \text{median}_{1 \leq i \leq j \leq M} B^{(ij)}$$

Robust Estimators

Define $\hat{P}$ and P_{HL} similarly as above:

$$\bar{A} = VSV^T \text{ and } \hat{A}_{HL} = ULU^T$$

$$\hat{P} = \hat{X}\hat{X}^T, \text{ where } \hat{X}^{n \times d} = \hat{V}\hat{S}^{1/2}$$

$$\hat{P}_{HL} = \hat{X}_{HL}\hat{X}_{HL}^T, \text{ where } \hat{X}_{HL}^{n \times d} = \hat{U}\hat{L}^{1/2}$$

Analytical Result: $E[\bar{A}]$

We have

$$\mu_1(\epsilon, c) = E[X_i] = (1-\epsilon)\cdot\sum_{k=1}^{K}\pi_k\nu_k + \epsilon\cdot\frac{c}{2}\cdot\mathbf{1}_{\boldsymbol{d}},$$

$$E\left[\bar{A}_{ij}\right] = \begin{cases} \epsilon\frac{c^2}{12}d + \alpha(\epsilon, c) & \text{if } i = j \\ \mu_1(\epsilon, c)^T\mu_1(\epsilon, c) & \text{otherwise} \end{cases}$$

Robust Estimators

Numerical Results

	$\bar{A}$	A_{HL}	$H_A : \hat{A}_{HL} < \bar{A}$
Mean of $\|\hat{X}\hat{X}^T - P^*\|$	40.07	20.28	
p-Value of sign test			$< 10^{-16}$

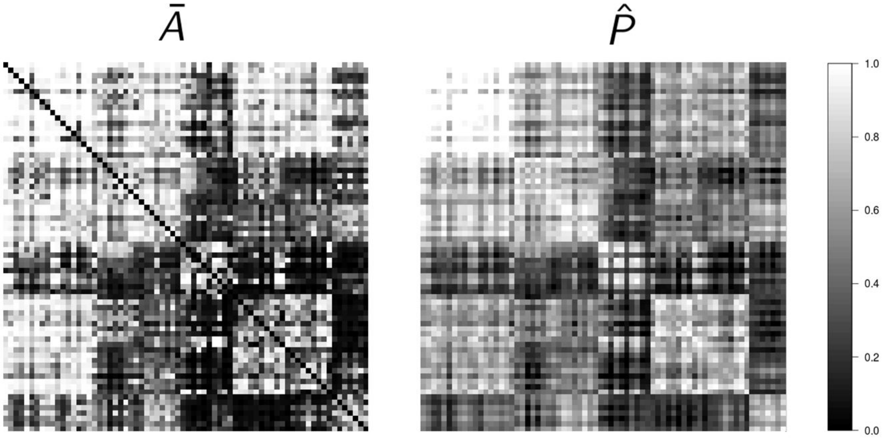
Table: $\epsilon = 0.2$, $\epsilon_B = 0.05$, $c = 2$, $m = 10$, 100 replicates.

	$\bar{A}$	A_{HL}	$H_A : A_{HL} < \bar{A}$
Mean of $\|\hat{X}\hat{X}^T - P^*\|$	19.15	13.93	
p-Value of sign test			$< 10^{-16}$

Table: $\epsilon = 0.1$, $\epsilon_B = 0.05$, $c = 2$, $m = 10$, 100 replicates.

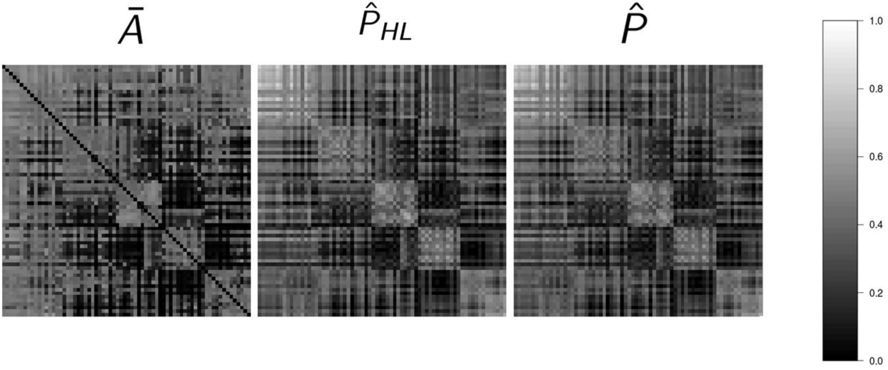
Real Data Application

Unweighted Graph: $\bar{A}$ vs. $\hat{P}$



Real Data Application

Weighted Graph: $\bar{A}$ vs. $\hat{P}_{HL}$ vs. $\hat{P}$



Future Work

- Choosing d
- Robust theory for Hodges-Lehman
- Is the estimator efficient?
- Unknown vertex correspondences