Carnegie Mellon University
School of Computer Science

A social scientist's guide to network statistics

http://mominmalik.com/network-stats-guide.pdf

Momin M. Malik
November 10, 2016 (updated December 23, 2016)

70/73-449: Social, Economic and Information Networks
Instructor: Dr. Katharine Anderson
Carnegie Mellon University, Fall 2016

Why statistics?

- If we have a philosophical belief that there is *variability* in the world:
  - Entities/processes have *underlying similarities* without being identical
  - We want to not be fooled by variability, thinking that we have found patterns when there are none
- Statistics is a way to systematically manage variability (using probability as a model)
- Measures and metrics in networks do not account for variability, and so we worry about them leading us astray
- Potentially: institutional/professional pressure, that if there isn't statistics, it isn't "science" and won't get published

The problem with network statistics:

Everything is terrible, and nothing works.

Overview

- Dependencies cause problems
- A reasonable default is to do a logistic regression and add massive caveats
- Use whatever model is accepted by your community
- Or…
  - Give up on empirical analysis and do simulation modeling
  - Give up on modeling and do qualitative analysis

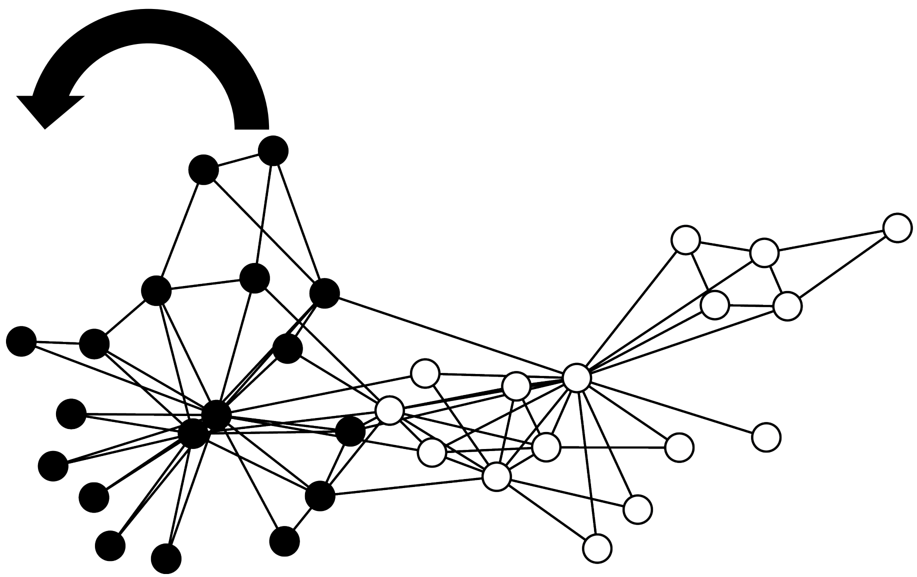
Outline

1. Why do “obvious” approaches not work?
    a. First pass: add network metrics as covariates
        i. Conceptual and technical problems
        ii. Dependencies
    b. Second pass: model edges
2. Overview of existing models
    a. Models that *control* for dependencies
        i. MRQAP
        ii. Network autocorrelation
        iii. “Bootstrapping”
    b. Models that *model* dependencies
        i. Stochastic Block Models
        ii. Propensity score matching (doesn’t work even for non-networks)
        iii. Latent Space Models (justifiable but useless)
        iv. p_1, p_2, and ERGMs (perfect but unjustifiable)
        v. SAOMs (highly stylized, layers upon layers of assumptions)
        vi. REMs (appropriate for lots of data from the Internet but poor predictions)
3. The difficulties of causality in networks

1. Why do “obvious” approaches not work?

“Obvious” first pass:

	Y	X_1	X_2	$\cdots$	X_k
v_1	y_1	x_{11}	x_{12}	$\cdots$	x_{1k}
v_2	y_2	x_{21}	x_{22}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
v_n	y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}



“Obvious” first pass:

	Y	X_1	X_2	$\cdots$	X_k	C_d
v_1	y_1	x_{11}	x_{12}	$\cdots$	x_{1k}	d_1
v_2	y_2	x_{21}	x_{22}	$\cdots$	x_{2k}	d_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
v_n	y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}	d_n

Add centralities (degree, eigenvector, etc.) or other network metrics as a covariate (explanatory variable, or maybe, make it the response variable)

Conceptual problems

- *Ceteris paribus* interpretation
- "Holding all else constant": how do we change the (undirected) degree of one node (or some centrality like eigenvector, betweenness, closeness) and hold those of all other nodes constant?
- Deeper question: **what are we trying model?**
- Centralities are a very crass way of capturing network structure
  - Are a *by-product* of network structure, not the cause of it
- Even if we have a directed graph, e.g. an advice network,
  - Modeling in-degree centrality would be getting at who is sought out
  - But not *by whom*
  - Out-degree centrality would be getting at who seeks out advice
  - But not *from whom*

Technical problems

- "Model misspecification"
  - The wrong functional form, and/or
  - The wrong variables
- Endogeneity
- **Omitted variable bias** (OVB)
- "Non-iid data" (independent and identically distributed, an important property in statistics)

- What is "dependent" or "independent"?

Dependence as factoring joint distributions

Full joint distribution (probability of observing the data as a whole):

$$p(Y, \mathbf{X}) = p(y_1, \ldots, y_n, x_{11}, x_{12}, \ldots, x_{1k}, x_{21}, \ldots, x_{n1}, \ldots, x_{nk})$$

$$p(Y, \mathbf{X}) = p(Y, X_1, \ldots, X_k) \overset{\text{ind}}{=} p(Y | X_1, \ldots, X_k) \prod_{j=1}^{k} p(X_j)$$

If you can't factor by columns: graphical models

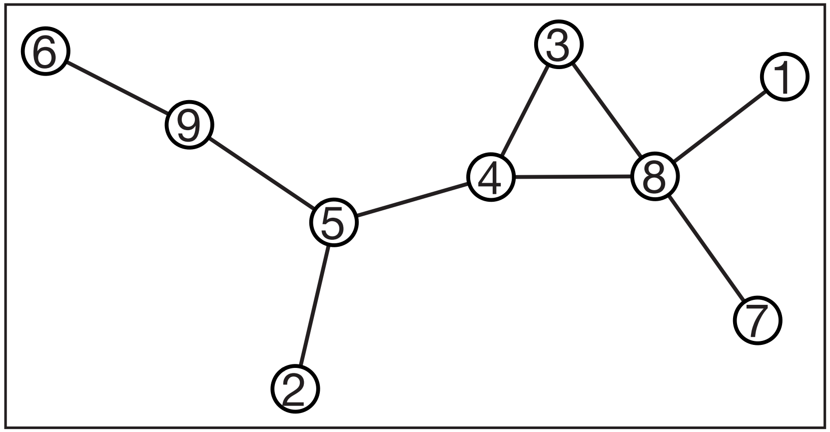
$$p(Y, \mathbf{X})$$
$$= p(y_1, \ldots, y_n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$$
$$\overset{\text{iid}}{=} \prod_{i=1}^{n} p(y_i, \mathbf{x}_i)$$

If you can't factor by rows: network models

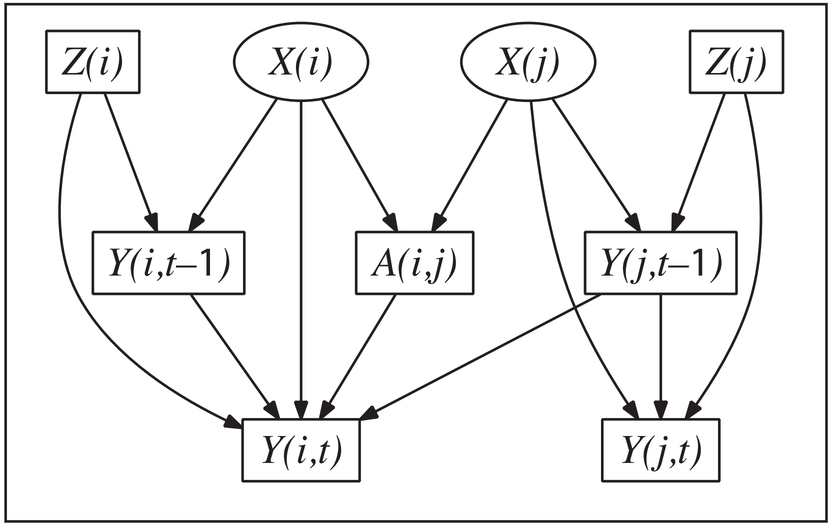
	Y	X_1	X_2	$\cdots$	X_k
v_1	y_1	x_{11}	x_{12}	$\cdots$	x_{1k}
v_2	y_2	x_{21}	x_{22}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
v_n	y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}

Graphs model both types of dependence!

Network model
(observations)



Graphical model
(variables)

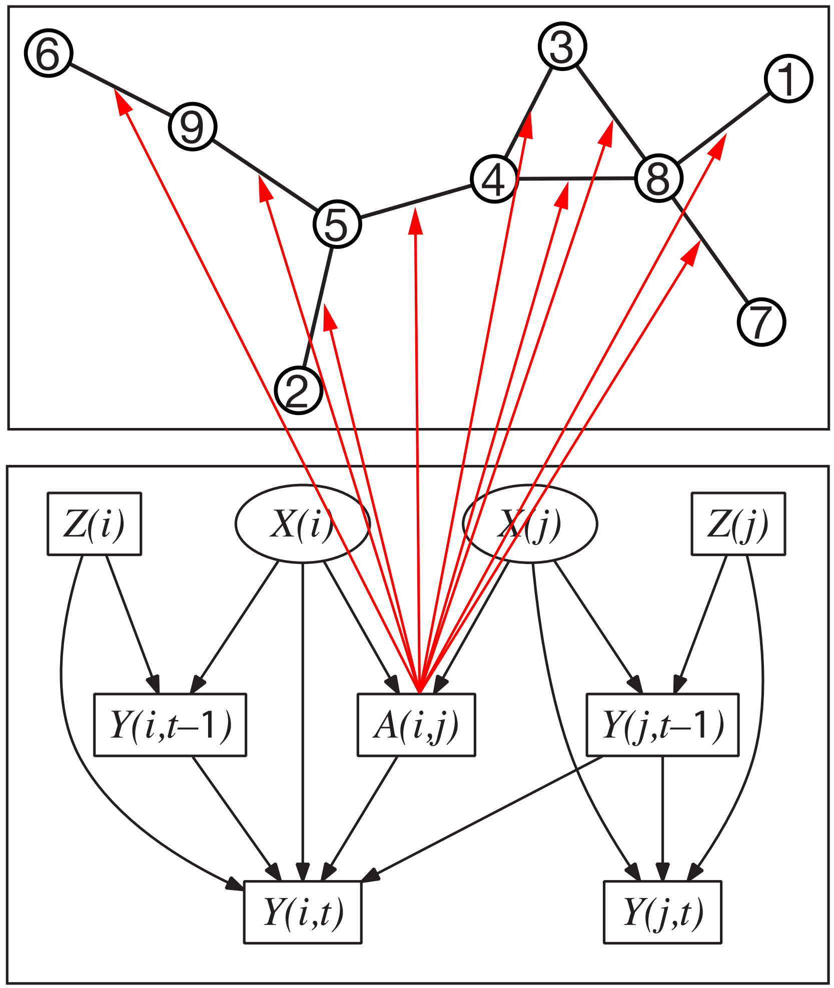


Graphical model from Shalizi & Thomas (2011)

Graphs model both types of dependence!

Network model
(observations)

Graphical model
(variables)

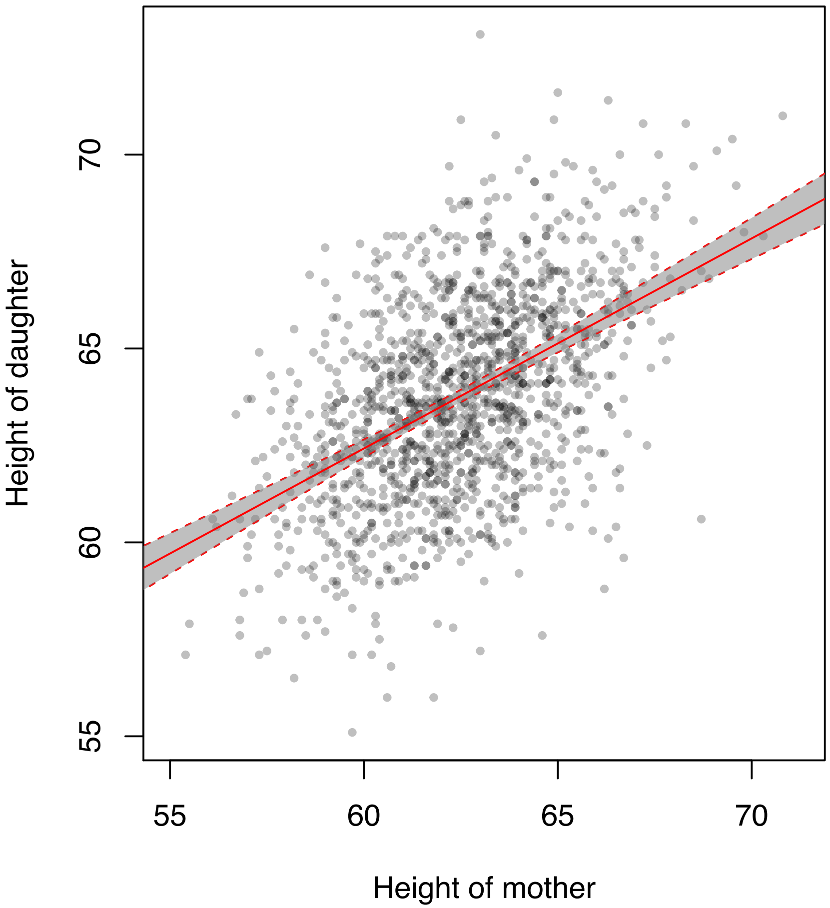


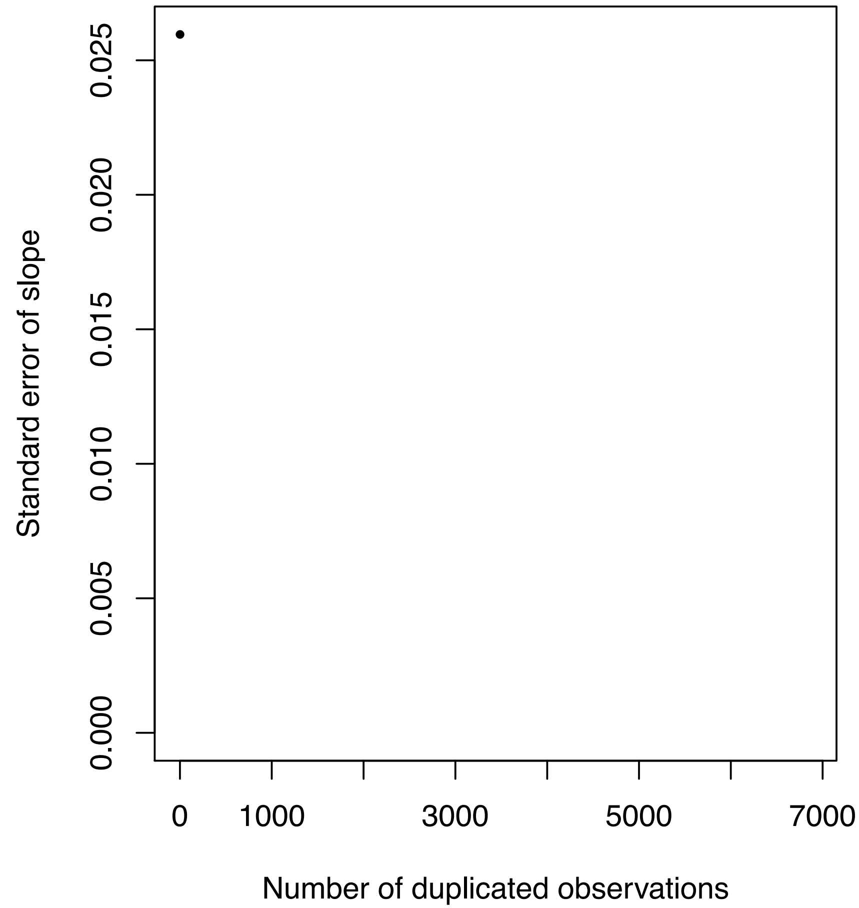
Graphical model from Shalizi & Thomas (2011)

Dependence: A simple illustration

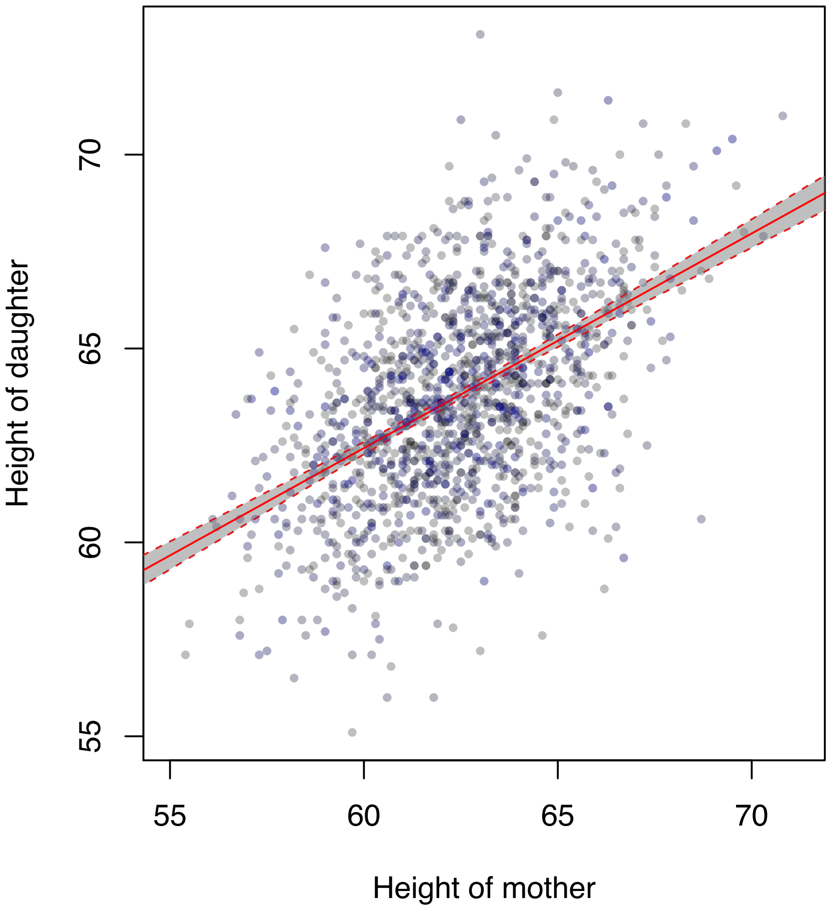
- From Wikipedia: "Asking two people in the same household whether they watch TV, for example, does not give you statistically independent answers. The sample size, *n*, for independent observations in this case is one, not two."
- **The simplest form of dependence: duplicate observations**
- Let's use Galton's height data
- Sample from the observations at random, and append a copy of that observation to the data set.

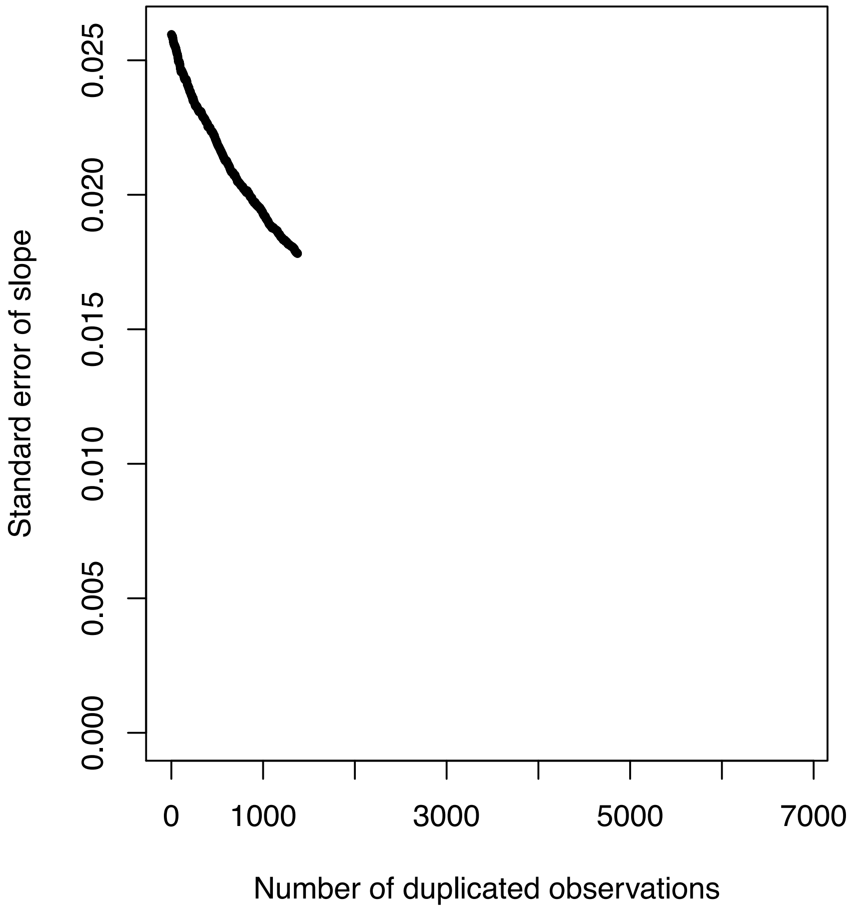
99% CI, duplicates in blue



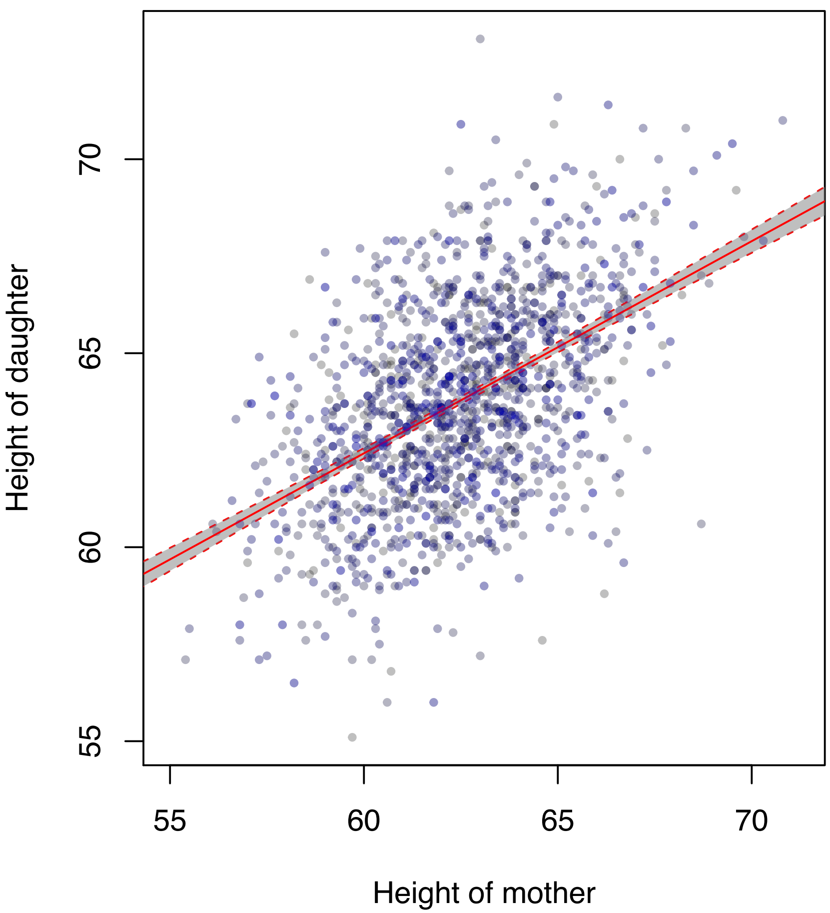


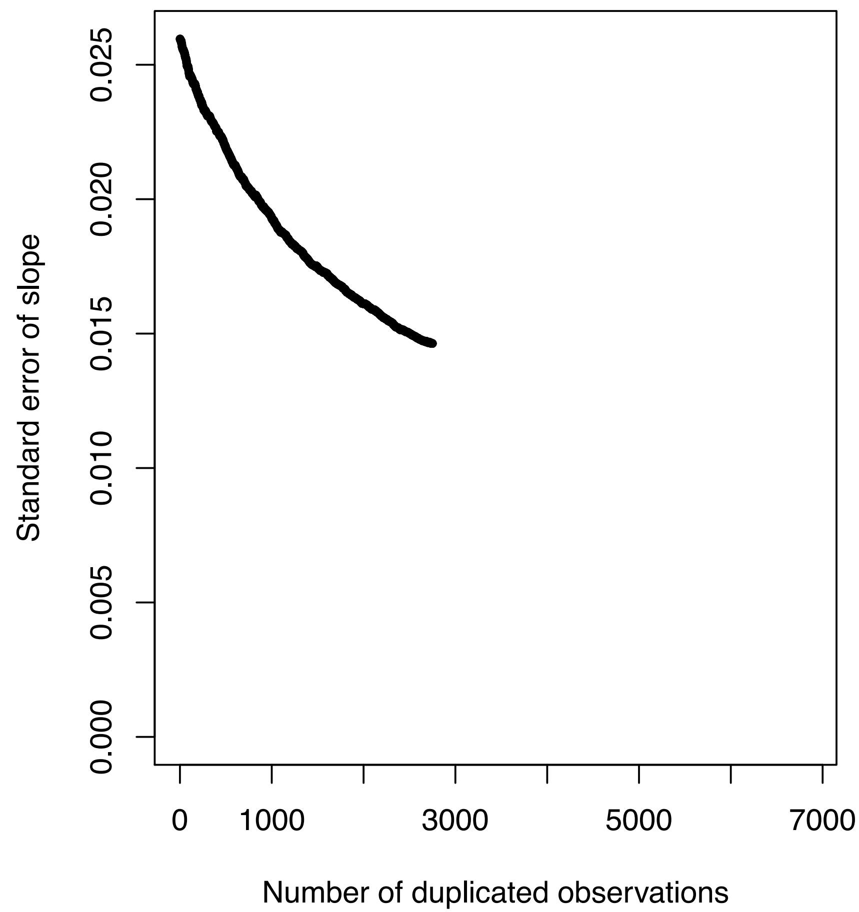
99% CI, duplicates in blue



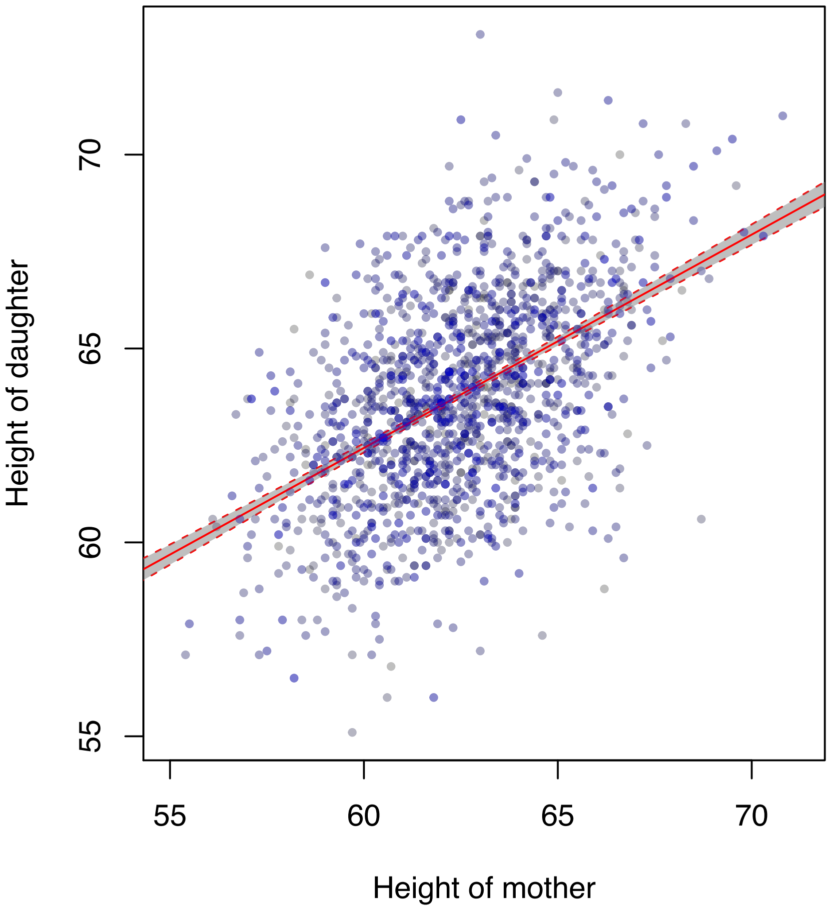


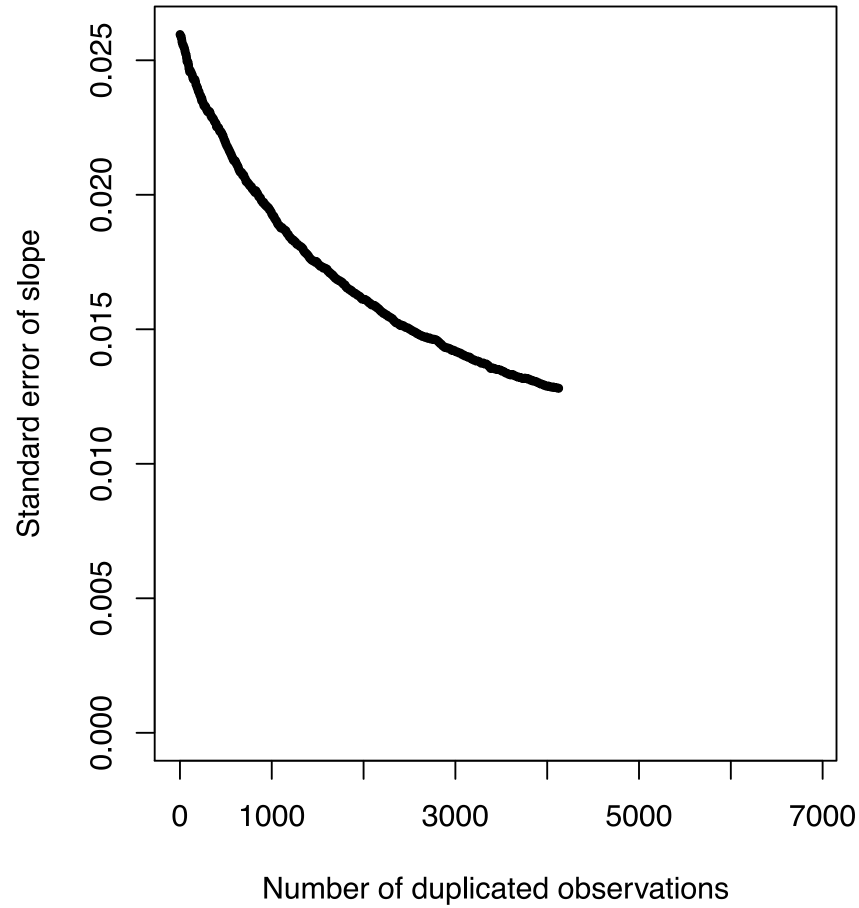
99% CI, duplicates in blue



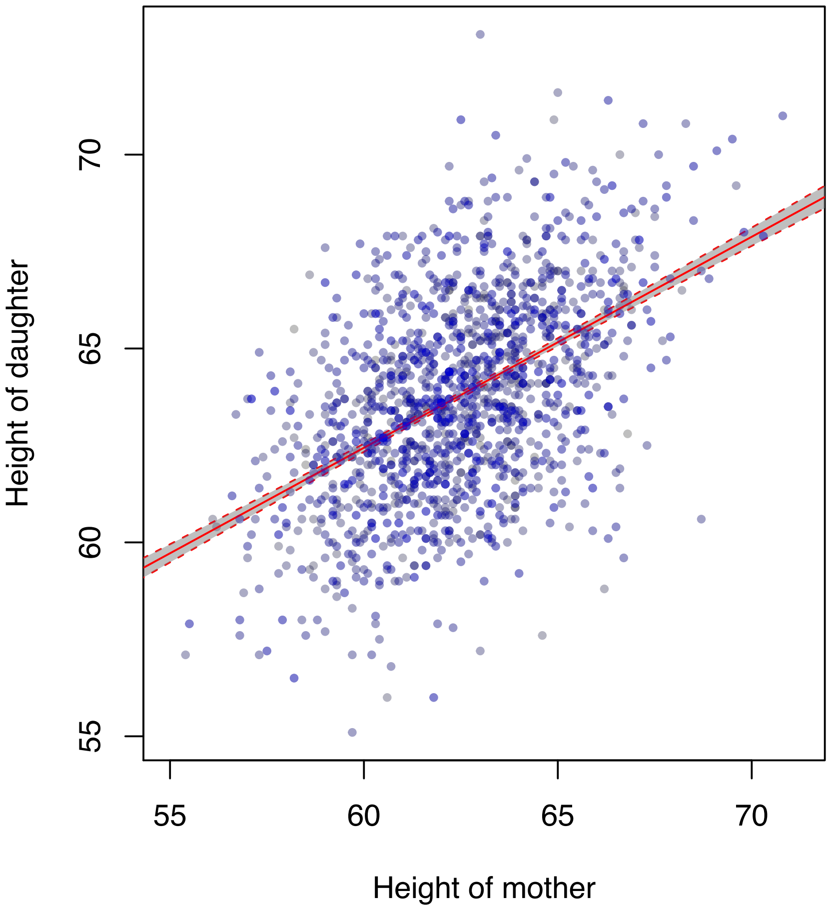


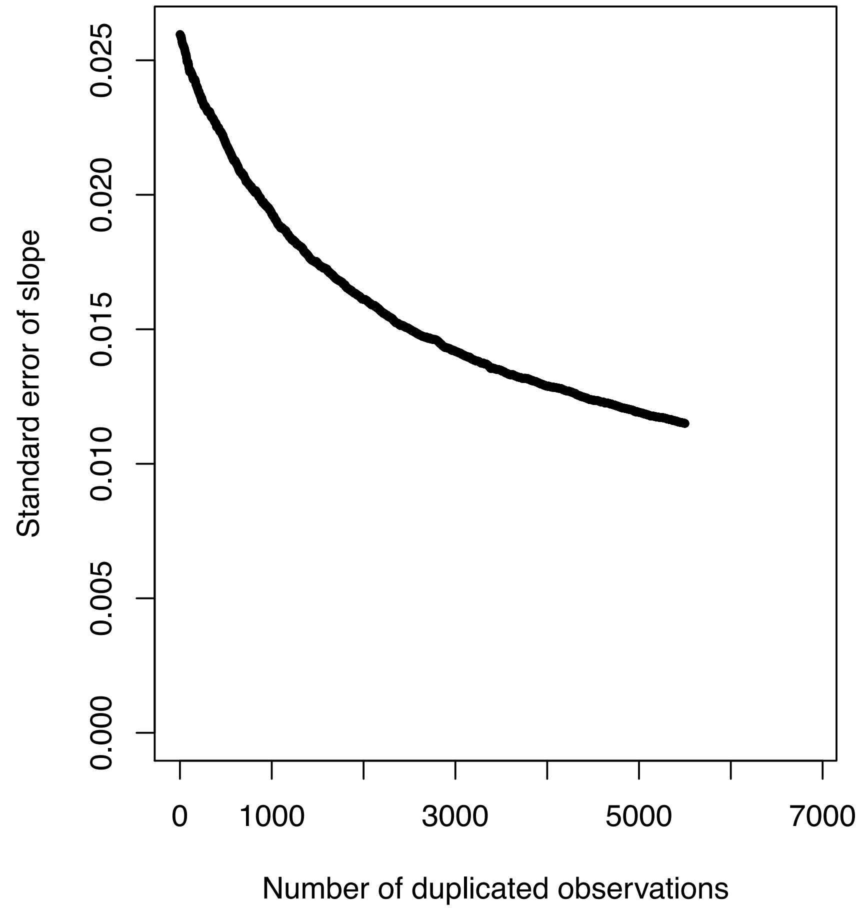
99% CI, duplicates in blue



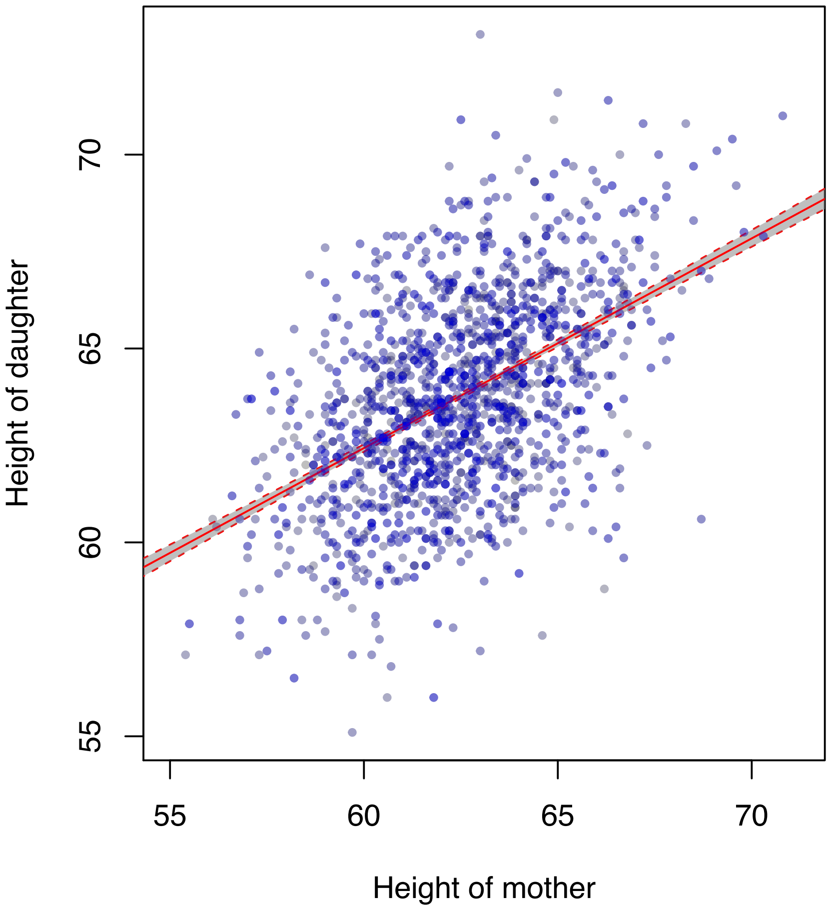


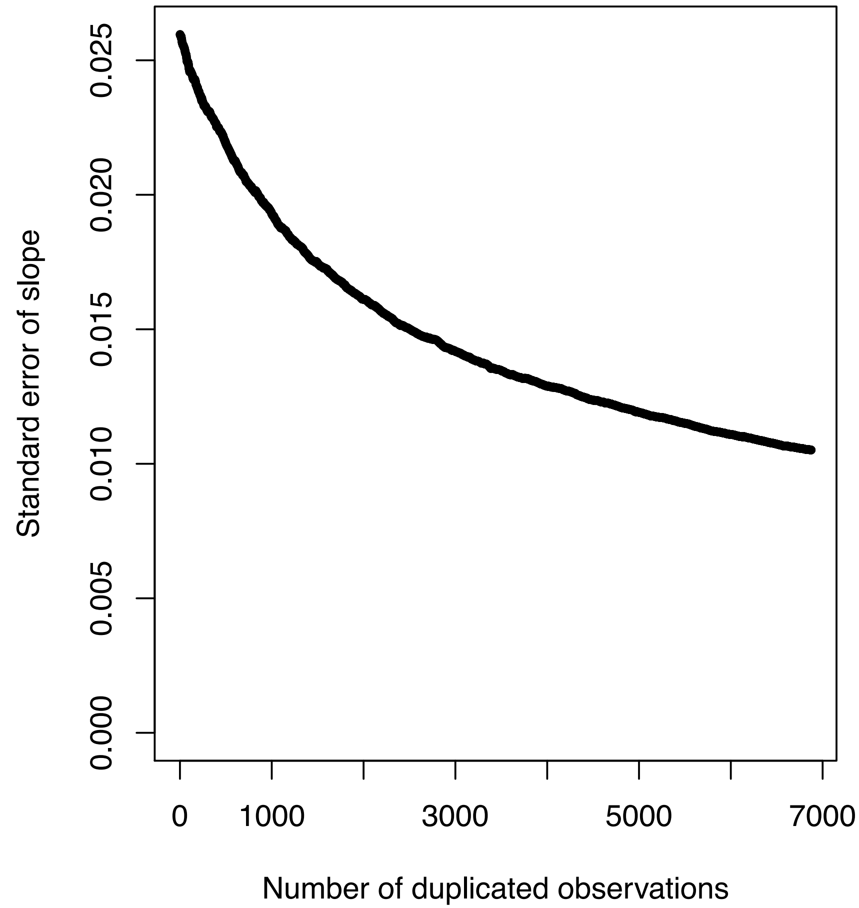
99% CI, duplicates in blue





99% CI, duplicates in blue



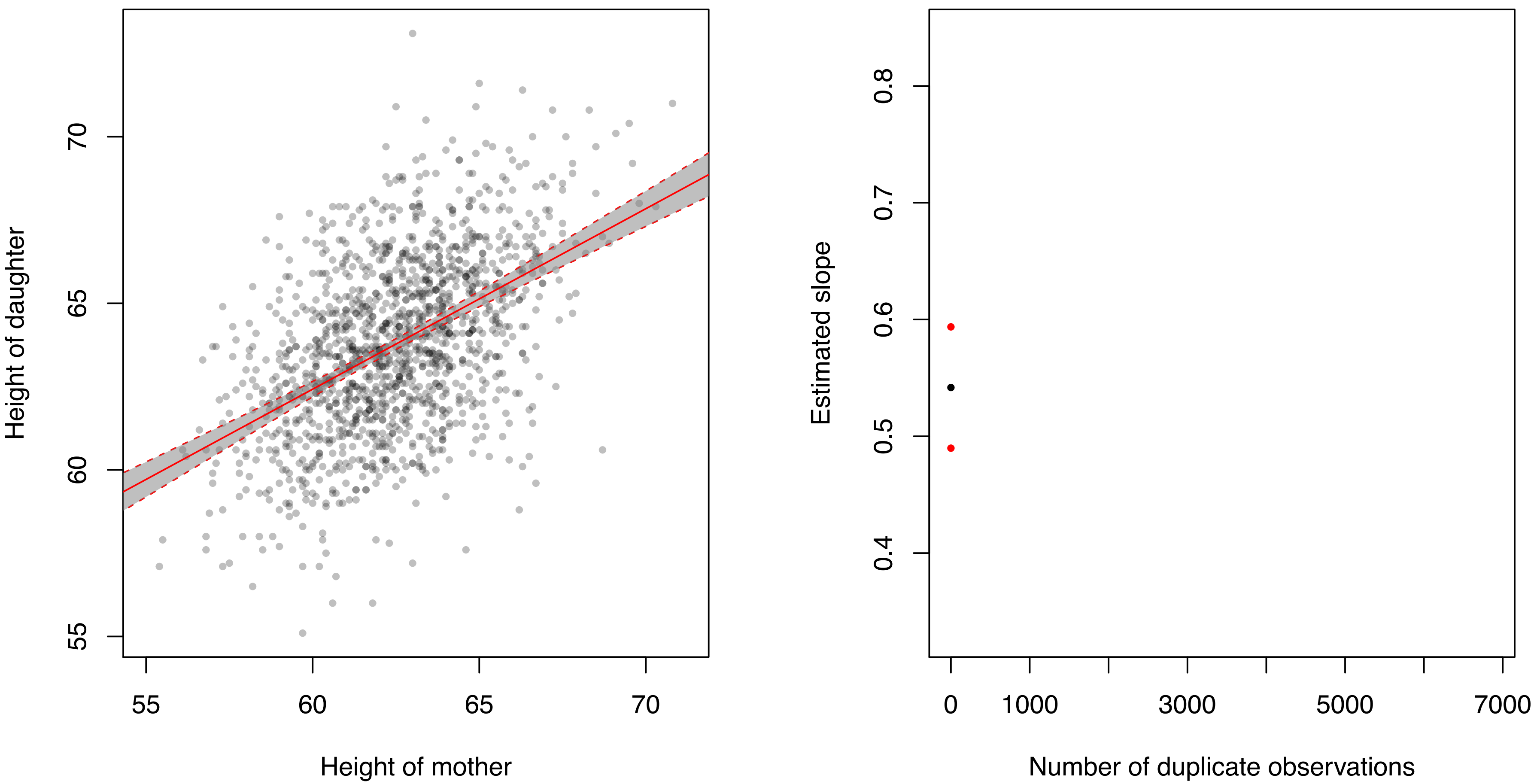


“Duplicated at random” is not so bad

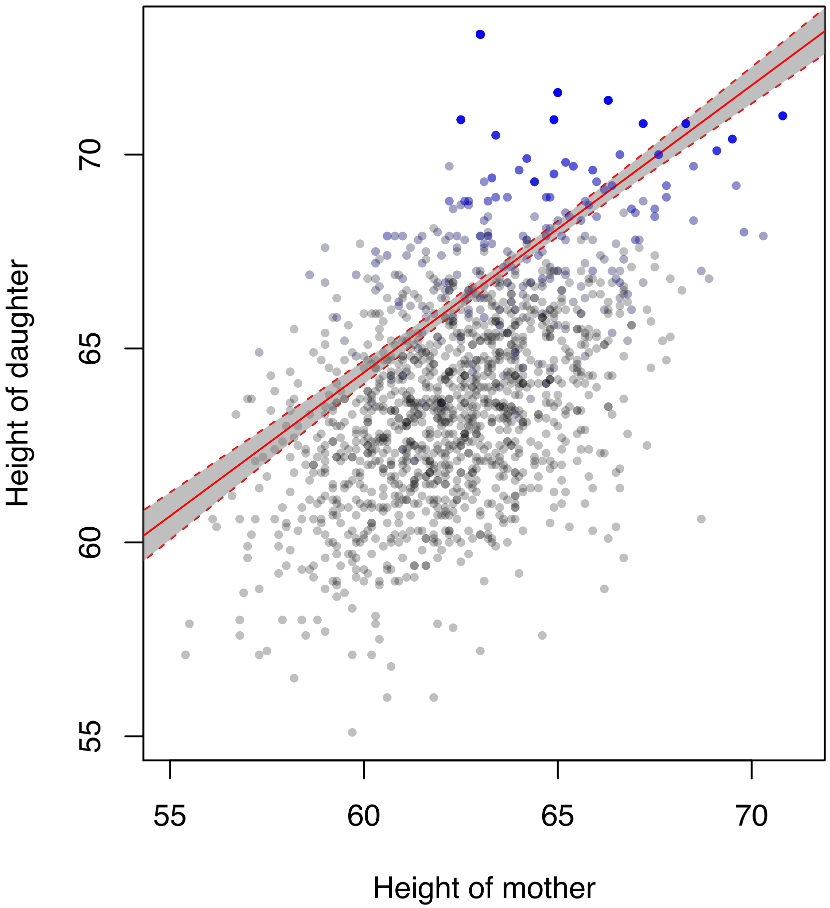
- Standard errors shrink (at a rate of $n^{-1/2}$), but no bias.

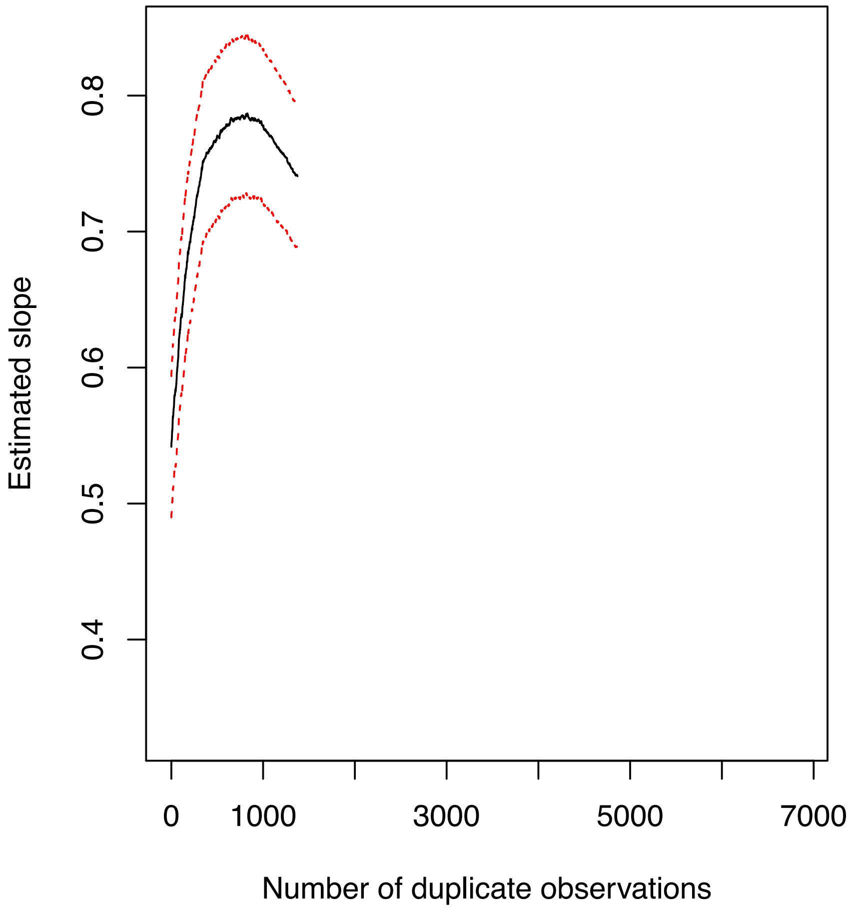
- If observations duplicated *not* at random, but instead proportionately to the dependent variable…

99% CI, duplicates in blue

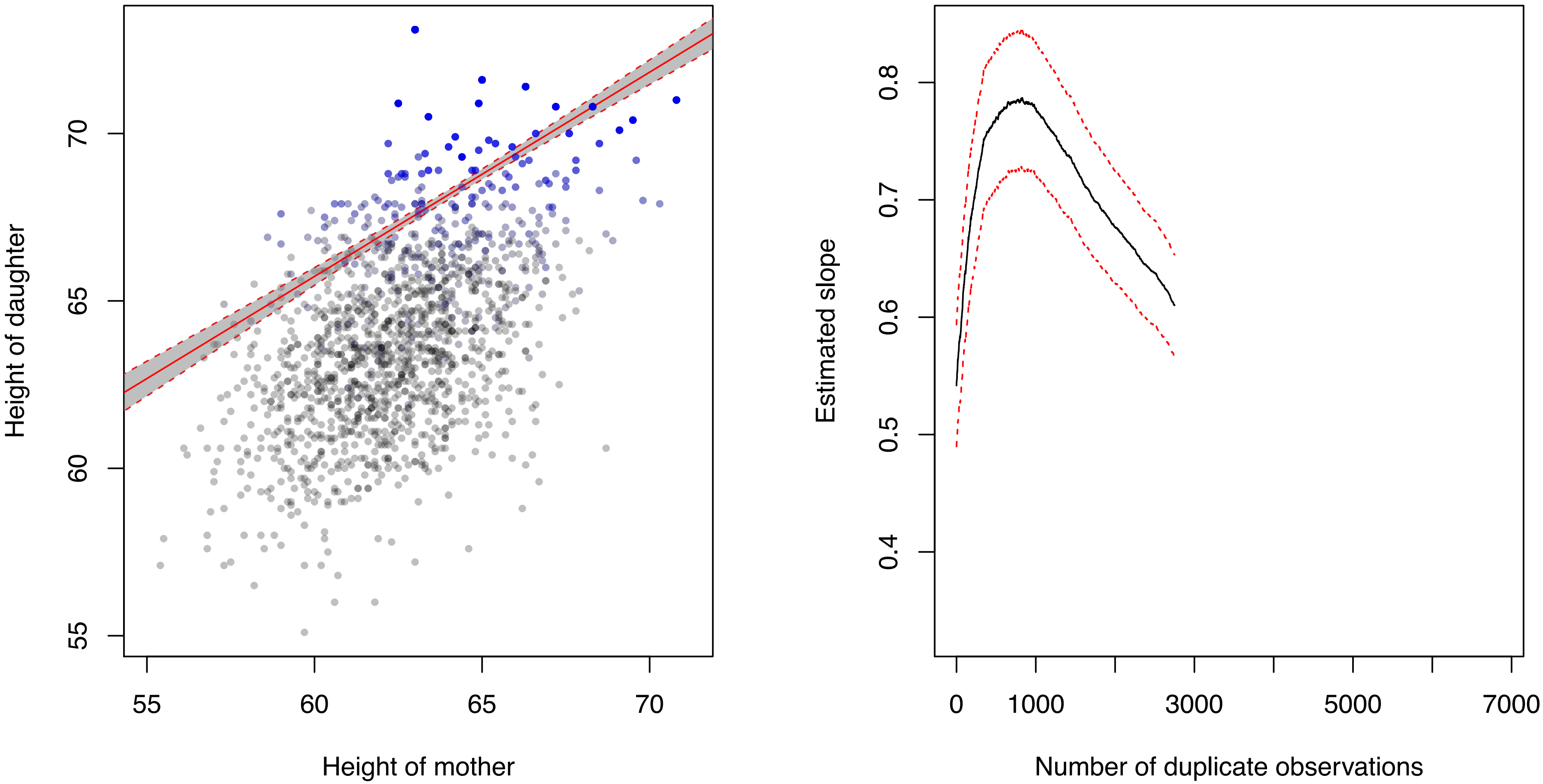


99% CI, duplicates in blue

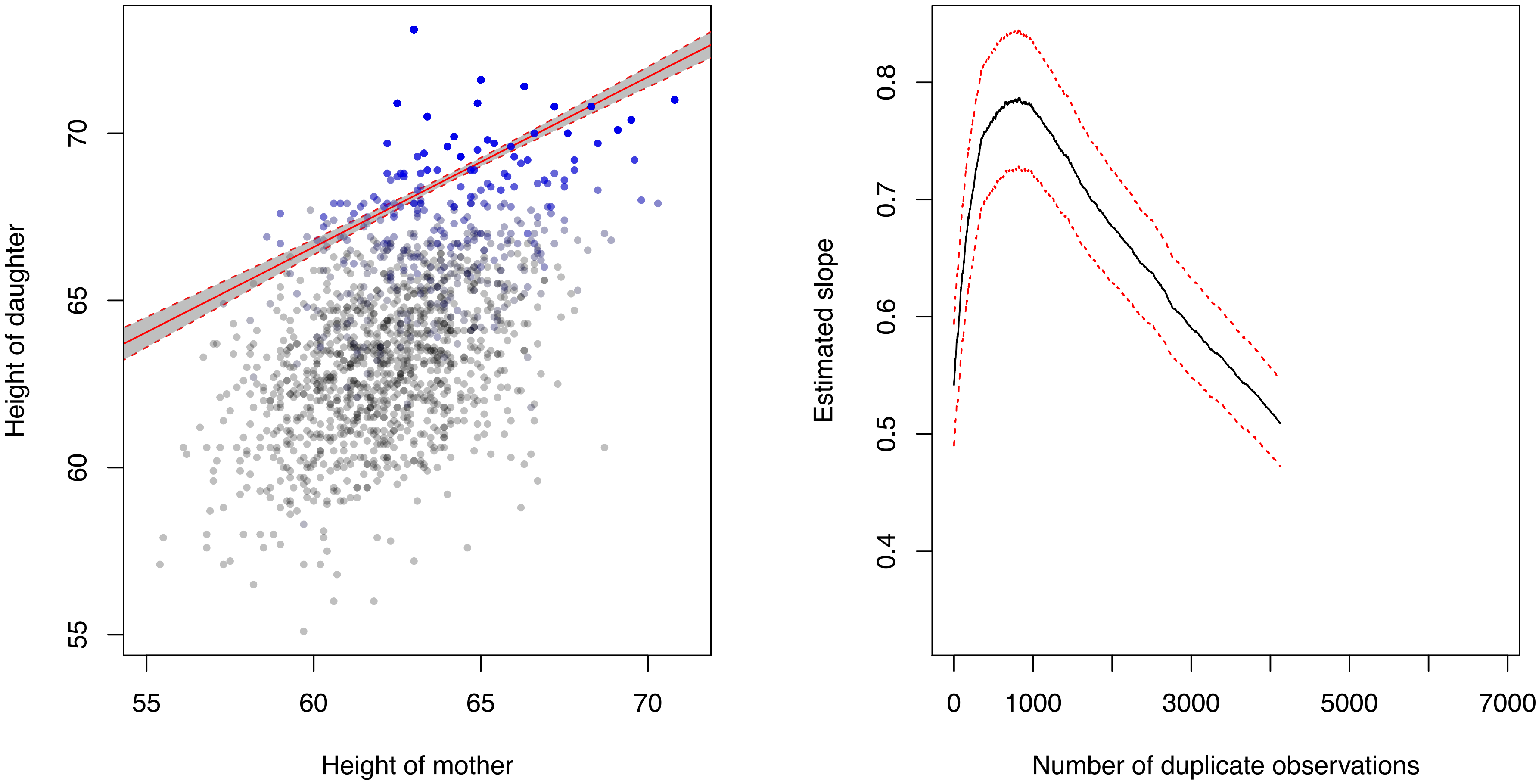




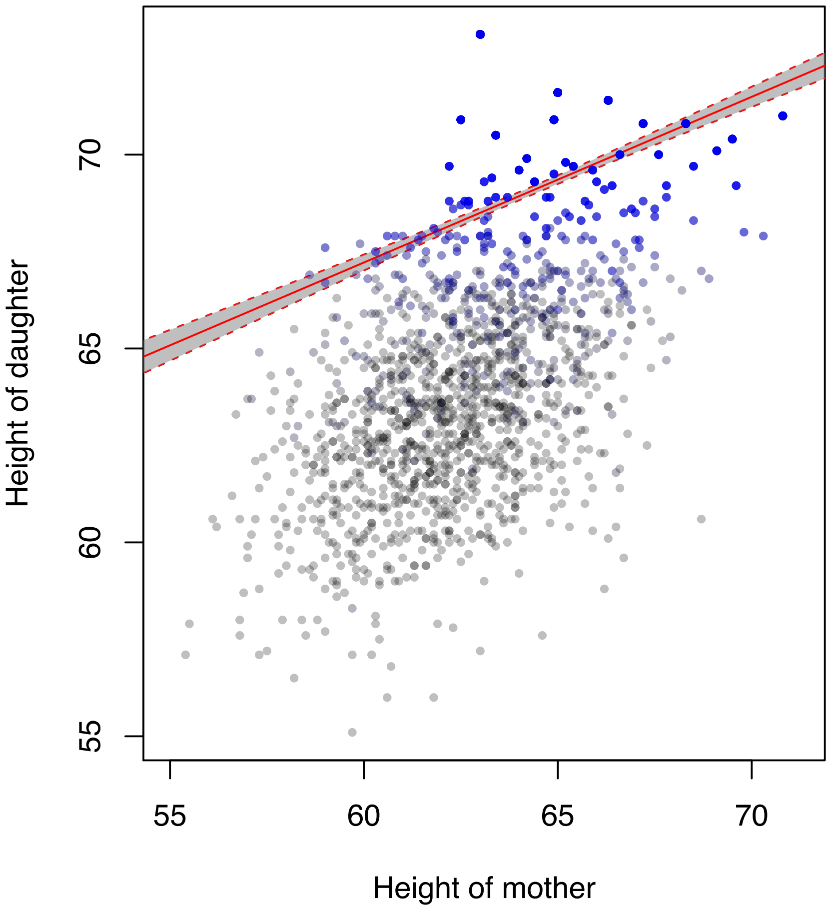
99% CI, duplicates in blue

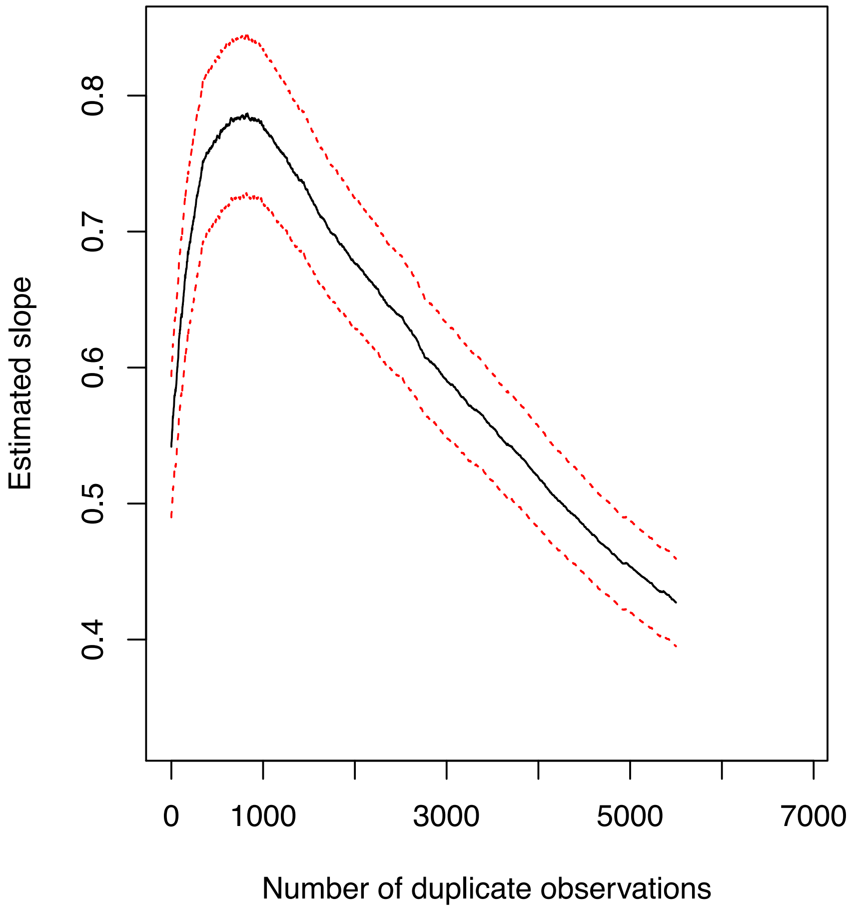


99% CI, duplicates in blue

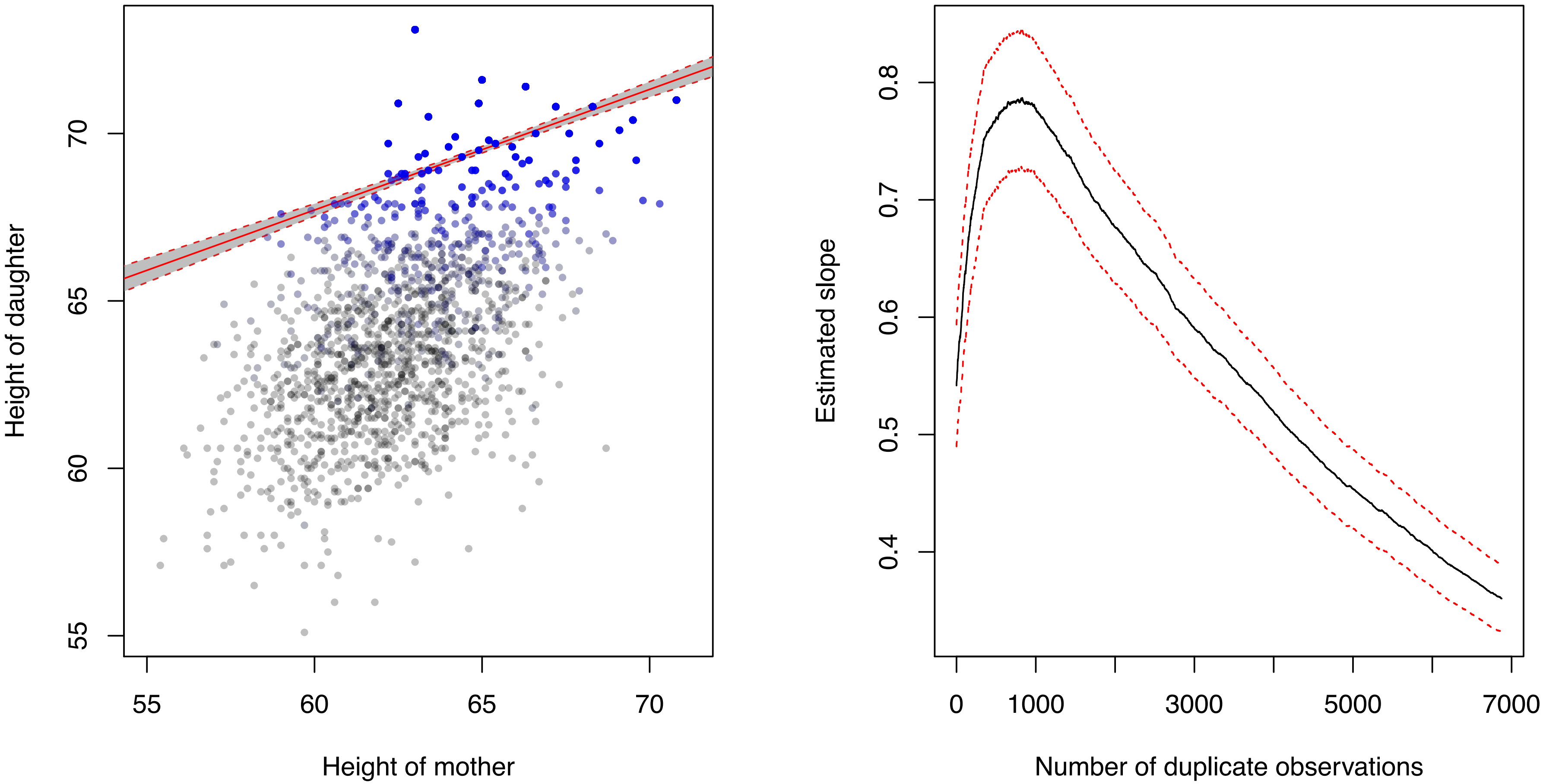


99% CI, duplicates in blue





99% CI, duplicates in blue



Endogeneity/Autocorrelation

- “Dependencies” can shrink standard errors *and* cause bias
- Can phrase as a problem of endogeneity, or of OVB
- If the dependence is regular enough, we can try to
- model it directly...
- Time series does this
- “Temporal autocorrelation”: an observation is dependent with “itself” at different times
- Network dependencies don’t have the same regularity

Second pass: Model the *edges*

	Y	X_1	X_2	$\cdots$	X_k
v_1	y_1	x_{11}	x_{12}	$\cdots$	x_{1k}
v_2	y_2	x_{21}	x_{22}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
v_n	y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}

(Need to turn node covariates into edge covariates, can do in different ways)

	from	*to*	Y	W_1	W_2	W_3	$\cdots$
e_1	v_1	v_2	y_{12}	$\mathbb{1}(x_{11} = x_{21})$	$x_{12} - x_{22}$	x_{13}	$\cdots$
e_2	v_2	y_3	y_{23}	$\mathbb{1}(x_{11} = x_{31})$	$x_{12} - x_{32}$	x_{13}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
e_{n+1}	v_2	v_1	y_{21}	$\mathbb{1}(x_{21} = x_{11})$	$x_{22} - x_{12}$	x_{23}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$e_{2\binom{n}{2}}$	v_{n-1}	v_n	$y_{(n-1)n}$	$\mathbb{1}(x_{(n-1)1} = x_{n1})$	$x_{(n-1)2} - x_{n2}$	$x_{(n-1)3}$	$\cdots$

Second pass: Model the *edges*

- Yes, you actually transform your data set, going from *n* rows to $2\times\binom{n}{2}$ rows (or $\binom{n}{2}$ for undirected graphs)
- The edges represent the dependencies between observations
- Problem: *the edges are dependent, too!*
- Transitivity, reciprocity, Dunbar’s number: these are dependencies *between edges*
- Not only are we not measuring important forces, but we assume them away (probably get OVB!)

Second pass: Model the *edges*

- Still: *a logistic regression on the edges is a reasonable first pass*

- Also, understand: most network models are models of the *edges*

2. Overview of existing models

Controlling for network structure 1: MRQAP

- The "Quadratic Assignment Procedure" (Krackhardt, 1987) is a *nonparametric permutation test,* same as the Mantel test in ecology
- The procedure: take the adjacency matrix *A* and another matrix *X* of attributes/similarities, turn both into vectors, find the correlation
- Permute the node labels of the graph, take the new adjacency matrix *A′*, again turn into a vector, and calculate correlation again
- If *X* were be correlated with *A* "by chance," we should see the actual correlation fall in the middle of a distribution of permutations

Controlling for network structure 1: MRQAP

- Can extend to “Multiple Regression QAP” (Dekker, Snijders, & Krackhardt, 2007), same as “Mantel regression”
- Problem: permutation tests are *tests*, not models
- When using them as models, you “get the standard errors from the null model”: your standard errors are a feature of the variability of permutations on the graph, not the variability of your data X
- Further problem: can only control for network structure, not model it

Controlling for network structure 2: Network autocorrelation

- A great frame for understanding dependencies (Dow et al., 1984)
- Analagous to temporal autocorrelation and time series models: fit a parameter for lag
- Problem: can only fit a single parameter for all network autocorrelation
- Problem: is the adjacency matrix the “right” weights matrix? Maybe not! (Leenders, 2002)
- Further problem: again, control for network dependencies at best, still don’t model them

Controlling for network structure 3: "Bootstrapping"

- Find your graph density. Initialize an empty graph; consider each pair independently, and connect with a probability equal to the density
- Repeat many times to get a distribution of whatever relationship you want to test. See where the empirical measure falls in the center of that distribution (not significant) or at the tails (significant)
- Equivalent to a "parametric bootstrap from a Bernoulli random graph model" (or, from an Erdos-Renyi random graph model)

Controlling for network structure 3: “Bootstrapping”

- A very good sanity check if you develop new metrics. But not very good as a null model, because it doesn’t capture how networks actually form (again, neglects dependencies between edges like reciprocity, transitivity, etc.)
- Can do similar bootstrapping with other models, like the “configuration model” (split the graph apart into nodes with edges equal to their degree, then join up edges randomly; preserves degree distribution)
- Such bootstrapping from various models is sometimes also called “conditional uniform random graphs”
- Problem: no good null model for social networks

Modeling the dependencies 1: Stochastic Block Models

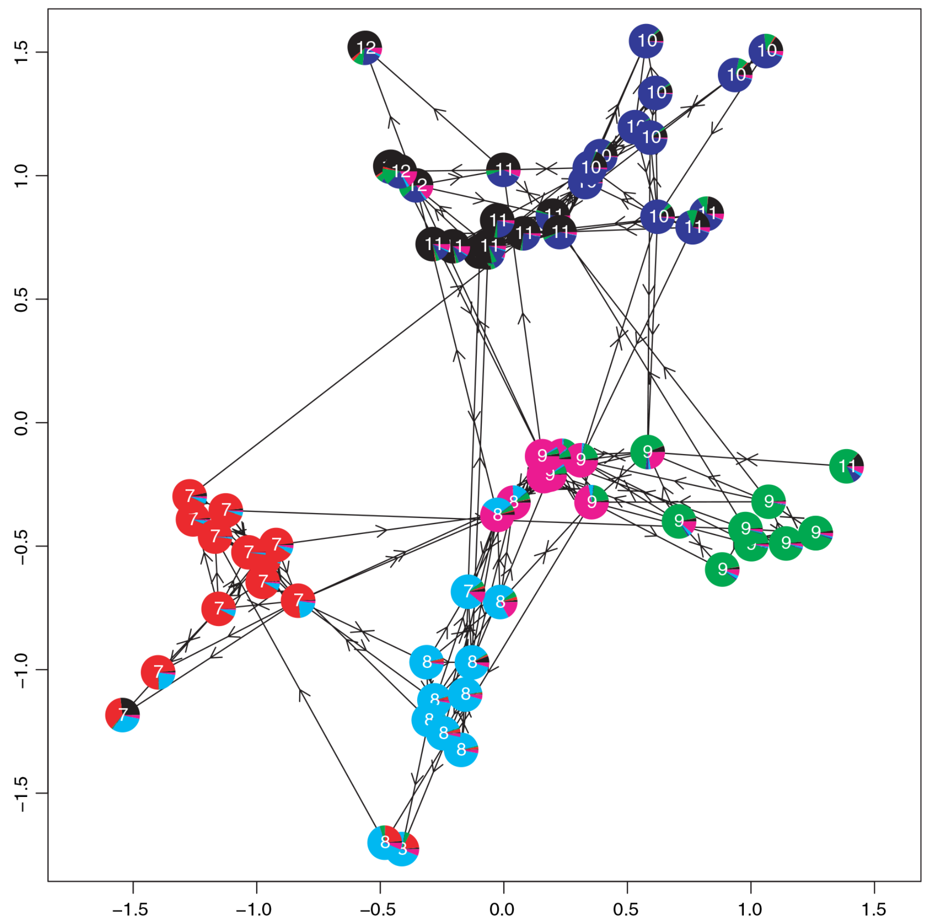
- A random graph model with community structure: separate parameters for within-group ties and out-of-group ties, otherwise everything is independent and Bernoulli
- A foundational model for statisticians, because it is analytically tractable
- But can only model community dependencies, so really would never use for empirical analysis

Modeling the dependencies 2: Propensity score matching

- Idea: pair up people in the network and observe their outcomes. Difference in outcomes can be attributed to the network (Aral, Muchnik, & Sundararajan, 2009)
- Problem: even perfect nonparametric matching cannot overcome OVB (Arceneaux, Gerber, & Green, 2010)

Modeling the dependencies 3: Latent Space Models

- Consider networks as existing in an extremely high-dimensional space, where the graph neighbors of a node are the ones it is geometrically closest to
- The dimensions of this space (somehow) “soak up” all dependencies
- Latent Space Models model such a space in lower dimension
- Pro: Unlike other models, this has good theoretical properties
- Con: Pretty much the only information is pictures like this: tells nothing about processes of interest, just gives a visual grouping.



Handcock et al. (2006)

Modeling the dependencies 4: p_1, p_2

- Logistic regression on edges can't model dependencies between edges, like reciprocity
- Solution: multinomial regression. Each pair is an observation, with values in $\{i \quad j, i \longrightarrow j, i \longleftarrow j, i \longleftrightarrow j\}$
- Fixed effects for sending, receiving, and reciprocity
- This is the "p_1 model", recently redescribed as the "β model" or "sender-receiver model"
- p_2 model: random effects version of p_1

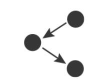
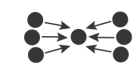
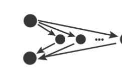
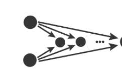
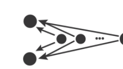
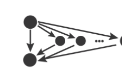
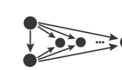
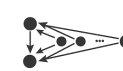
Exponential[-family] Random Graph Models (ERGMs)

- The crown jewel of 30+ years of research, came out of p_2 model
- (Main version treats graphs as the response: graphs as explanatory are called "autologistic actor attribute models" [ALAAMs], aren't really done)
- Logic: specify a set of *sufficient statistics*
- These can include terms for anything you can think of
- By construction, these are the sufficient statistics for a graph. Question is if there is any weighting of these statistics that can produced the observed graph

$$S_1(y) = \sum_{1 \le i < j \le n} y_{ij} \qquad \text{number of edges}$$

$$S_k(y) = \sum_{1 \le i \le n} \binom{y_{i+}}{k} \qquad \text{number of } k\text{-stars } (k \ge 2)$$

$$T(y) = \sum_{1 \le i < j < h \le n} y_{ij}\, y_{ih}\, y_{jh} \qquad \text{number of triangles}$$

Network statistics	Description	Structural signature
Univariate parameters		
Dyadic parameters		
Reciprocity	Occurrence of mutual ties	
Degree parameters		
Mixed 2-star	Correlation of indegrees and outdegrees	
Alternating-in-star (A-in-S)	Network centralisaiton around indegree	
Isolate	Occurrence of actors with zero indegree and zero outdegree	
Sink	Occurrence of actors with an outdegree of zero and indegree of at least one	
Triangle parameters		
Multiple connectivity (A2P-T)	Multiple paths of indirect connectivity	
Shared out-ties (A2P-U)	Activity based structural equivalence: multiple sets of out-ties to the same third others	
Shared in-ties (A2P-D)	Popularity based structure equivalence: multiple sets of in-ties from the same third others	
Transitive closure (AT-T)	Transitive closure of multiple 2-paths	
Activity closure (AT-U)	Closure of multiple in-2-stars	
Popularity closure (AT-D)	Closure of multiple out-2-stars	

Snijders et al. (2006); Simpson (2015)

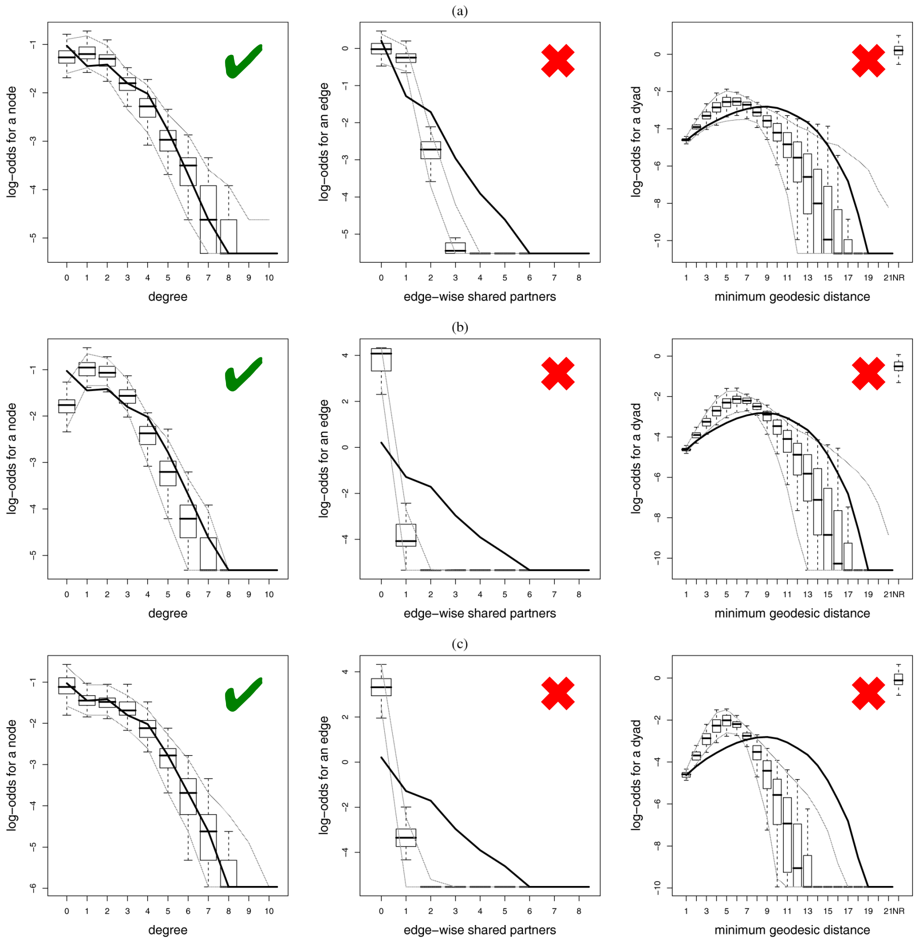
ERGMs: Procedure

- Take the observed graph, do counts of sufficient statistics, and initialize weights (through logistic regression)
- Holding the rest of the graph constant, consider a single edge.
- How would removing this edge (if present) or adding it (if absent) change the count of sufficient statistics? Would a higher/lower count make the graph more likely based on current weights?
- If yes, adjust weights so that the observed graph remains most likely.
- Do this for some time to explore the parameter space (an MCMC procedure)
- At the end: if the terms put in were indeed the “correct” ones, these would be their weights
- (Version for time series data: Separable Temporal ERGMs [STERGMs], needs observations at regular intervals.)

ERGMs: Goodness-of-fit testing

Excellent goodness-of-fit (GOF) testing framework.

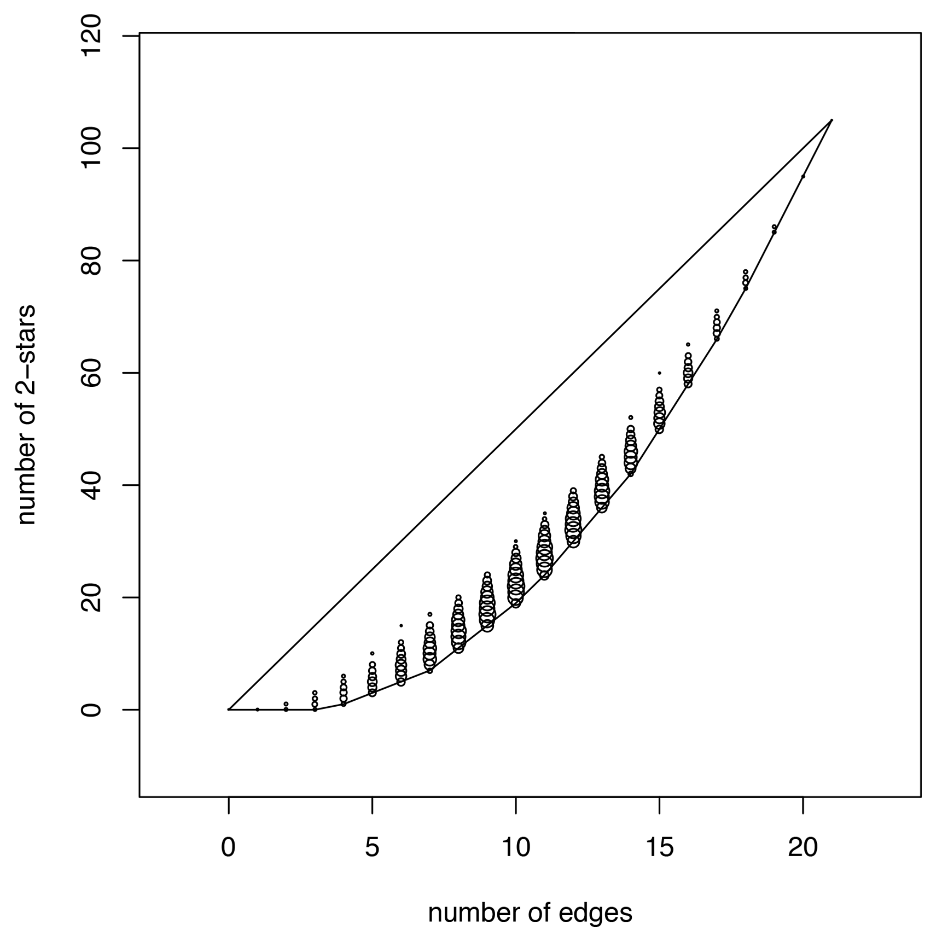
- See if the sufficient statistics that you put into the model can recover the distribution of statistics that were not among your sufficient statistics
- E.g., can density, reciprocity and transitivity alone as sufficient statistics recover the graph's degree distribution?
- Can test with anything (e.g., any subgraph/graph motif density), but should be theoretically important
- Gives a complete framework for finding a parsimonious explanation



Hunter et al. (2008)

ERGMs: The bad news

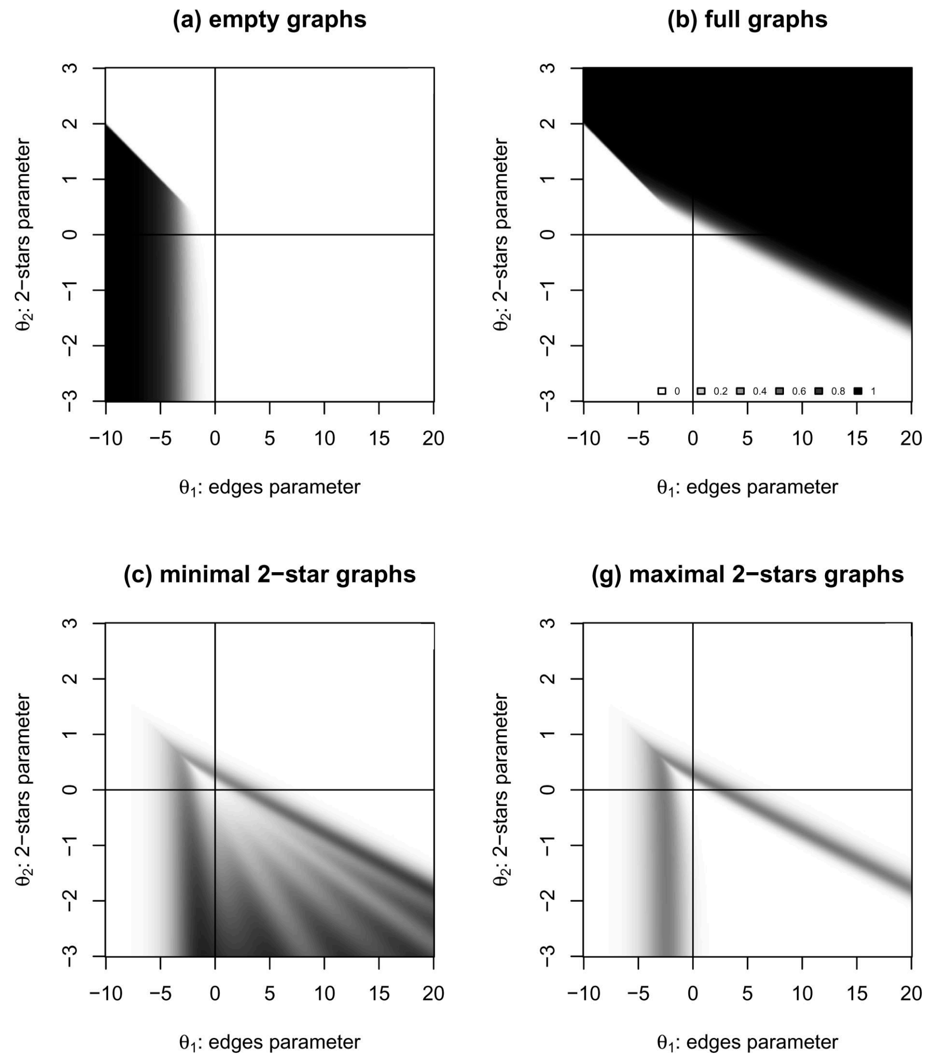
- LOTS of problems.
- The space of graphs doesn’t play nice with probabilities
- There are only a certain number of graphs of any given size, and only a certain number of graphs with a combination of sufficient statistics



Handcock (2003)

ERGMs: The bad news

- Sometimes, under large portions of the parameter space, the most likely graph is either the complete graph or the empty graph: such specifications are are *degenerate*
- Because the space of graphs is so large, don't know if a model is degenerate or if our MCMC procedure is bad
- Model degeneracy (arguably) has nothing to do with the social phenomena of interest
- Better specifications are (arguably) technical, not sociological, entities: e.g., "geometrically weighted edgewise shared partners"



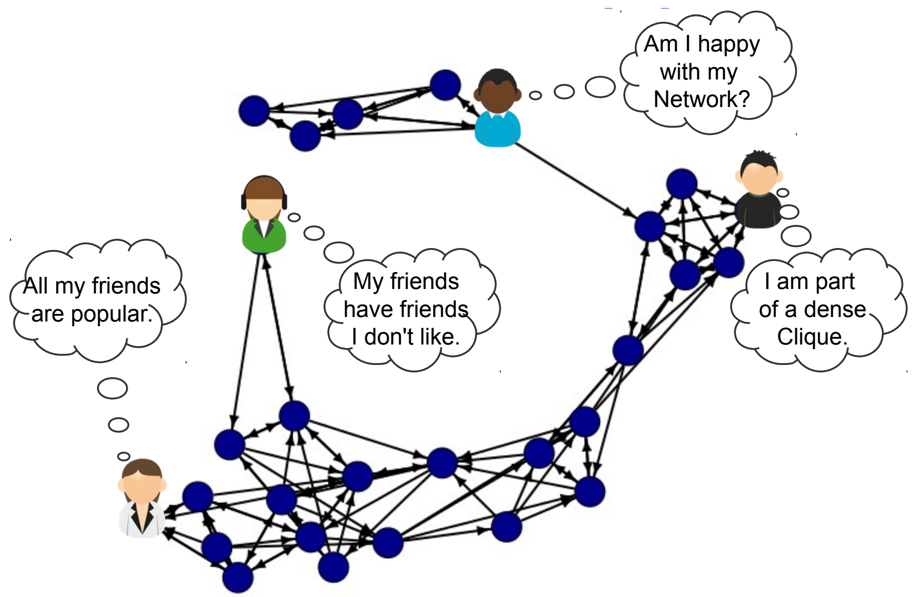
Handcock (2003)

ERGMs: The bad news

- Another: ERGMs are not “projective” (Shalizi & Rinaldo, 2013)
- Short explanation: if you are missing one node, it could have ties to every other single node, which would completely change the estimates of all the network effects
- In this sense ERGMs (and many other possible similar network models) are extremely fragile in substance (i.e., even if we can get them to work technically, they might be leading us astray)

Stochastic Actor-Oriented Models (SAOMs)/SIENA

- A different perspective: model actor decision-making ("utility")
- Currently, only SAOM is SIENA (Simulation Investigation for Empirical Network Analysis)
- Create utility functions with ERGM-like terms (SIENA manual gives 100+ built-in terms)
- Uses something like an agent-based model to fit the terms
- Elegant, only model to get at co-evolution of behavior and networks, but layers upon layers of assumptions
- And in practice, SIENA can be very temperamental, it's hard to models to successfully run



Steglich et al. (2010)

Relational Event Models (REMs)

- Relational Event Models (Butts, 2008b) model continuous-time network data (network ties with time steps, e.g. emails or calls, each of which is called an “event”)
- It is similar to (and builds on) ERGMs and SIENA in the terms it uses to express processes like transitivity, reciprocity, etc. Like SIENA, it models actor decision-making (the likelihood function tries to capture actor “utility”)
- Note that much of online data can be seen as “relational events,” so this is a very useful model. Applying SIENA or STERGMs requires binning time, which is artificial
- REMs normalize the probability of an observed event stream not by *all* possible event streams, but only by possible alternative actions (e.g., all other possible sender-receiver pairs) at a the time of each event given all previous events until then
- A good, reasonable model, but has extremely low predictive performance (in predicting who will send a tie to whom, and when/in what order it will happen)
  - We usually think of good predictive performance as (at least somewhat) necessary, but not at all sufficient, for getting close to “truth” (see also Shmueli 2010; Breiman 2001)
  - By that standard, this is not a good model. But maybe the core phenomenon is too variable for us to expect a model to predict it with any accuracy/recall
- Ultimately comes down to whether sequential alternative actions is the null set you want to do statistics over, and (as always) if you “believe” it or not

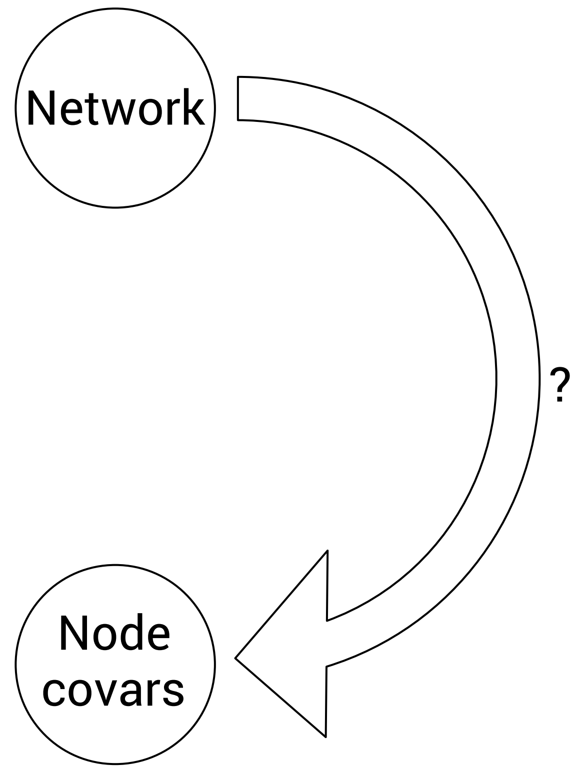
Scalability

- Yet another problem: none of the “good” models (LSMs, ERGMs, SIENA, REMs) scale past a few hundred nodes at best
  - They all require intensive computation (generally, MCMC procedures through a space of graphs or at least alternative edges)
  - In computer science terms, they scale as $O(n^2)$ or worse
- So, forget using any of these to model all of Facebook, or any other big dataset

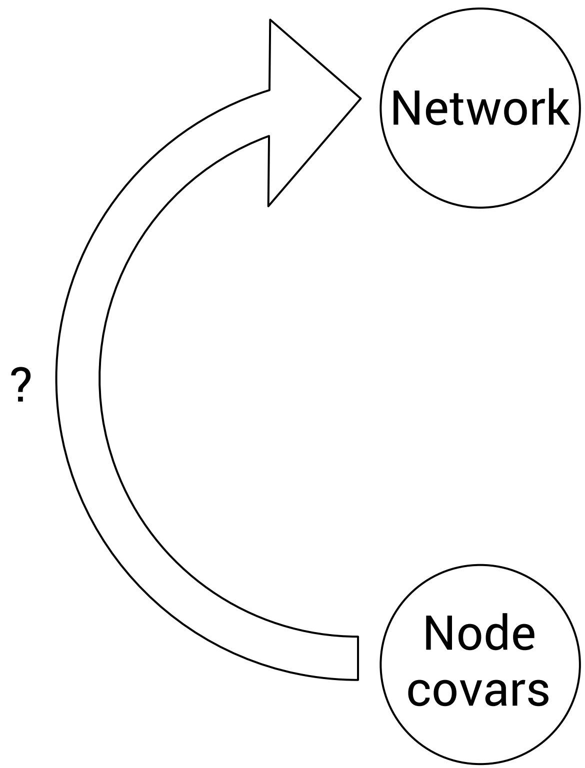
3. Causality

Network: explanatory or response?

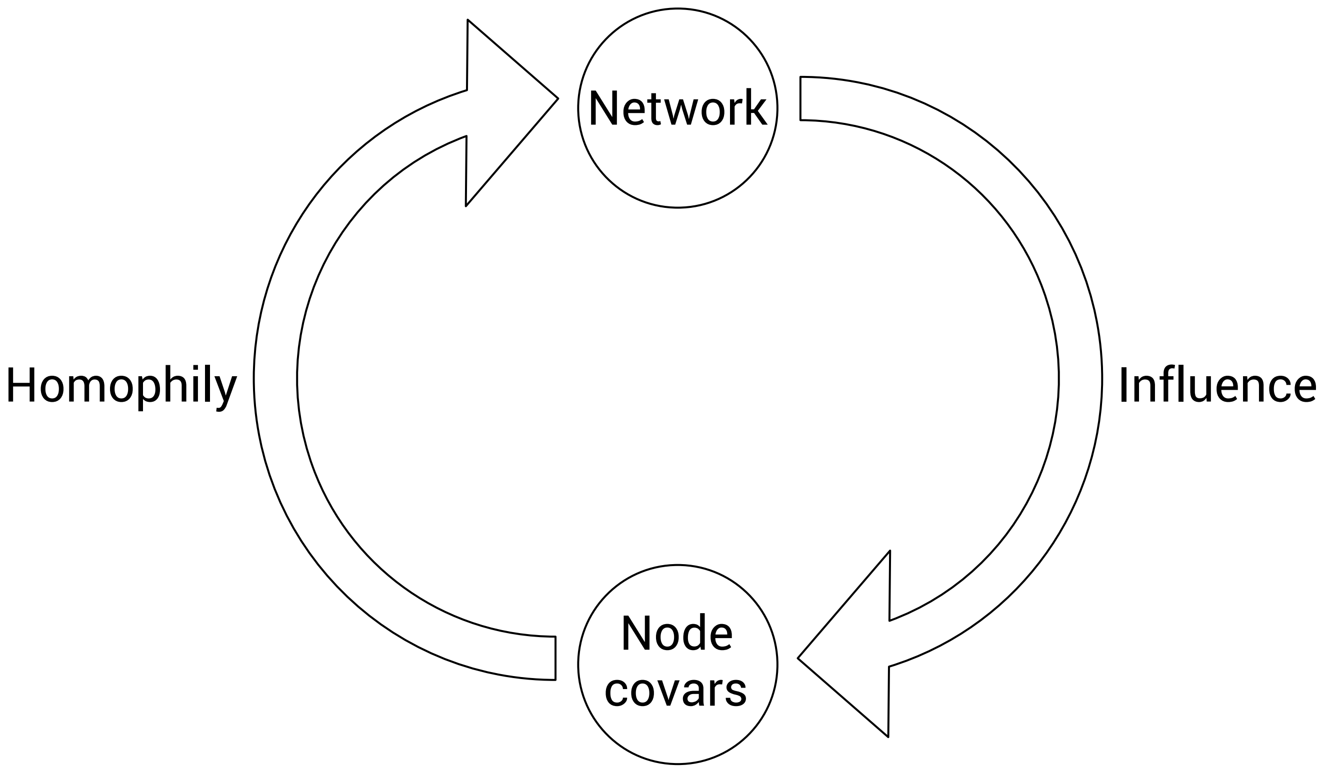
Network as cause? (as explanatory/IV?)



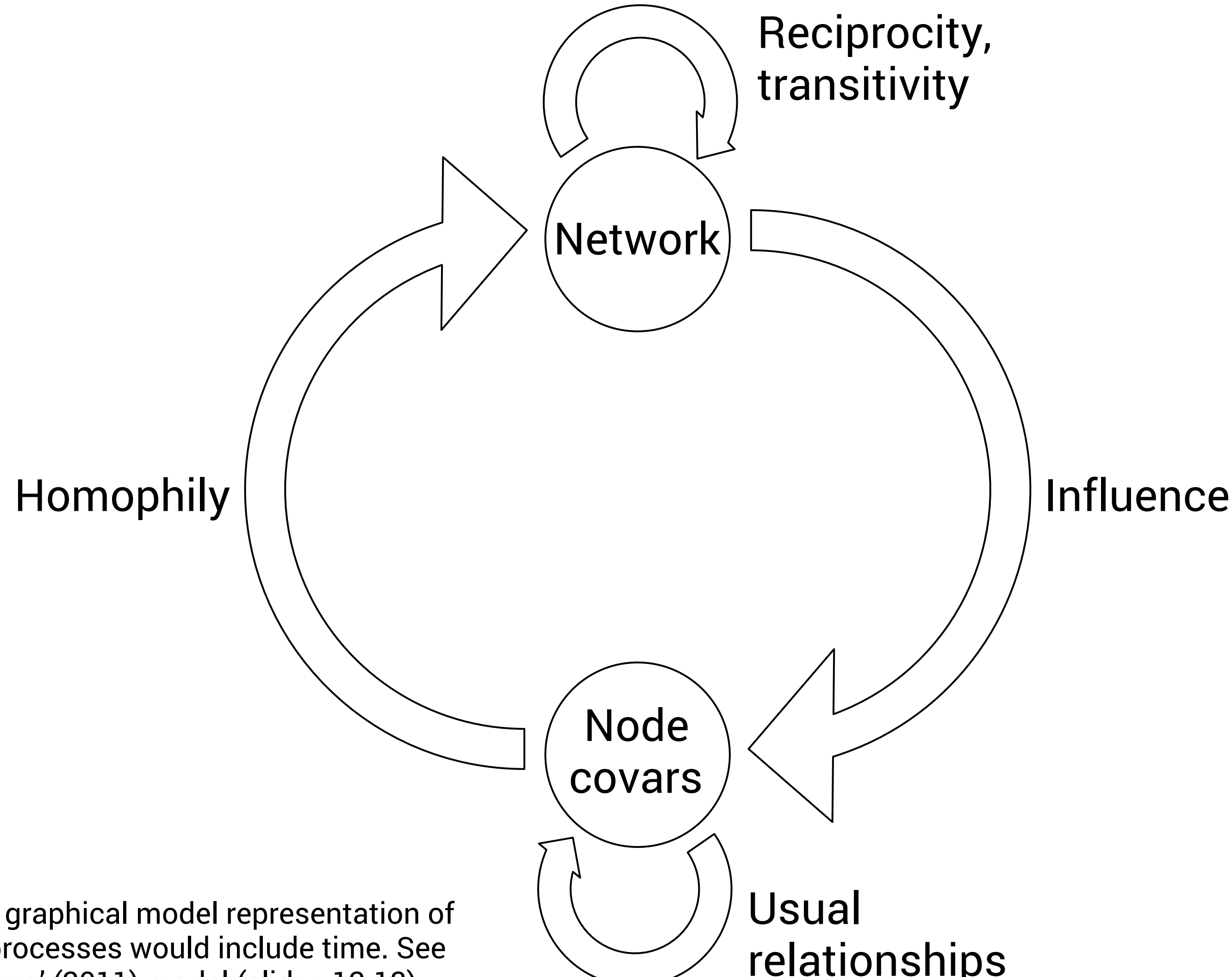
Network as effect? (as response/DV?)



The problem: both happen.



And they aren't the only things.



Note: a proper graphical model representation of these causal processes would include time. See Shalizi & Thomas' (2011) model (slides 12-13).

Sorting it all out is a challenge.

- There's too much going on: we might never be able to get enough data (or the right kind of data) to sort out different effects
- (SIENA does try to model this whole process, but it comes down to if you believe it or not)
- Experimentation isn't really an option either: since networks arise spontaneously, experiments that force people into network structures lack ecological validity

Final thoughts

Lessons

- Everything is terrible, and nothing works. Sorry.
- Don't bother with regression with centralities
- Do a logistic regression on the edges as a first pass (for yourself or an audience)
- Maybe mess with ERGMs, SAOMs, or REMs… if you believe them (and if nobody will yell at you for it)
- Fit a Latent Space Model if you need to satisfy statisticians
- Wait for better models?
- Maybe there's a really clever study design or IV out there you can find

Come across a fancy (new) network model and wondering if it’s the answer?

- (Don’t worry, it’s not.)
- My heuristic: “[how] does it model transitivity?”
- If it doesn’t, I’m not interested
  - I care about network processes, for which transitivity (which happens between node triplets) is exemplary
- E.g., “degree-corrected stochastic block model”? Nope. “Kronecker graphs”? Nope. The “influence model”? Nope.
- Caveat: if you are doing *prediction*, not explanation (Shmueli, 2010; Breiman, 2001), the data-generating process is irrelevant and you should use whatever can perform well

The eternal caveat

- “All models are wrong…”
- “…but some are useful.” –George Box
- Networks are hard to measure
- All network data is highly uncertain
  - Perfect and complete trace data (e.g., online social media) doesn’t give us what’s important
  - Getting at what’s important (e.g., through surveys and interviews) gives us imperfect and incomplete data
- Networks are an abstraction. They may not be the “right” abstraction.

References

Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *PNAS*, *106*(51), 21544–21549.

Arceneaux, K., Gerber, A. S., & Green, D. P. (2010). A cautionary note on the use of matching to estimate causal effects: An empirical example comparing matching estimates to an experimental benchmark. *Sociological Methods & Research*, *39*(2), 256–282.

Besag, J. (2001). Markov Chain Monte Carlo for statistical inference. Working Paper No. 9, Center for Statistics and the Social Sciences, University of Washington.

Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, *16*(3), 199–231.

Butts, C. T. (2008a). Social network analysis: A methodological introduction. *Asian Journal of Social Psychology*, *11*, 13–41.

Butts, C. T. (2008b). A relational event model for social action. *Sociological Methodology, 38*, 155–200.

Cranmer, S. J., Leifeld, P., McClurg, S. D., & Rolfe, M. (2016). Navigating the range of statistical tools for inferential network analysis*. American Journal of Political Science*. DOI: 10.1111/ajps.12263

Dekker, D., Krackhardt, D., & Sniggers, T. A. B. (2007). Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. *Psychometrika*, *72*(4), 563–581.

Dow, M. M., Burton, M. L., White, D. R., and Reitz, K. P. (1984). Galton's problem as network autocorrelation. *American Ethnologist*, *11*(4), 754–770.

Goldenberg, A., Zheng, A. X., Fienberg, S. E., and Airoldi, E. M. (2010). A survey of statistical network models. *Foundations and Trends in Machine Learning*, *2*(2), 129–233.

Handcock, M. S. (2003). Degeneracy in statistical models of social networks. Working Paper No. 39, Center for Statistics and the Social Sciences, University of Washington.

Handcock, M. S., Raftery, A. E., & Tantrum, J. M. (2007). Model-based clustering for social networks. *Journal of the Royal Statistical Society A*, *170*(Part 2), 301–354.

Hanneman, R. A. and Riddle, M. (2005). Chapter 18: Some statistical tools. In *Introduction to social network methods*. University of California, Riverside.

Hunter, D. R., Goodreau, S. M., & Handcock, M. S. (2008) Goodness of fit of social network models. *Journal of the American Statistical Association*, *103*(481), 248–258.

Kass, R. E. (2011). Statistical inference: The big picture. *Statistical Science*, *26*(1), 1–9.

Kolaczyk, E. D. and Csárdi, G. (2014). *Statistical analysis of network data with R*, volume 65 of Use R! Springer-Verlag.

Krackhardt, D. (1987). QAP partialling as a test of spuriousness. *Social Networks*, *9*, 171–186.

Leenders, R. T. A. J. (2002). Modeling social influence through network autocorrelation: Constructing the weight matrix. *Social Networks*, *24*, 21–47.

O'Malley, A. J., & Marsden, P. V. (2008). The analysis of social networks. *Health Services and Outcomes Research Methodology*, *8*, 222–269.

Pattison, P. and Robins, G. (2008). Chapter 18: Probabilistic network analysis. In T. Rudas (Ed.), *Handbook of probability: Theory and applications* (pp. 291–313). SAGE.

Robins, G. (2015). Drawing conclusions: Inference, generalization, causality and other weighty matters. In *Doing social network research: Network-based research design for social scientists* (pp. 211–230). SAGE.

Robins, G., Lewis, J. M., & Wang, P. (2012). Statistical network analysis for analyzing policy networks. *The Policy Studies Journal*, *40*(3), 375–401.

Shalizi, C. R., & Rinaldo, A. (2013). Consistency under sampling of exponential random graph models. *The Annals of Statistics*, *41*(2), 508–535.

Shalizi, C. R., & Thomas, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, *40*(2), 211–239.

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289–310.

Steglich, C., Snijders, T. A. B., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology*, *40*(1), 329–393.

Snijders, T. A. B. (2011). Statistical models for social networks. *Annual Review of Sociology*, *37*(1):131–153.

Snijders, T. A. B., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New specifications for Exponential Random Graph Models. *Sociological Methodology*, *36*, 99–153.

Acknowledgements (chronological order)

Tom Snijders
Cosma Shalizi
Andrew Thomas
Brian Junker
Cohen Simpson
David Krackhardt
Hemank Lamba
Alex Loewi
Philipp Burckhardt
Dave Choi
Valerie Ventura
The stat-network group at CMU

All errors, overgeneralizations, and unfair characterizations of models are my own.