Modeling Networks with Missing Data in statnet

ERGM Refresher

Carter T. Butts
Department of Sociology and
Institute for Mathematical Behavioral Sciences
University of California, Irvine

2011 Political Networks Conference
June 16, 2011
Ann Arbor, MI
http://polnet2011.statnet.org/

Exponential Families for Random Graphs (w/Covariates)

- **For random graph G w/countable support $\mathcal{G}$ and covariate set X, pmf can be written in ERG form:**

$$\Pr\left(G=g \middle| \mathbf{t},\theta,\mathcal{G},X\right)=\frac{\exp\left(\theta^T \mathbf{t}(g,X)\right)}{\sum_{g'\in\mathcal{G}} \exp\left(\theta^T \mathbf{t}(g',X)\right)} I_{\mathcal{G}}(g)$$

  - $\theta^T\mathbf{t}$: linear predictor
    - $\mathbf{t}:\mathcal{G}\to\mathbb{R}^m$: vector of sufficient statistics
    - $\theta\in\mathbb{R}^m$: vector of parameters
    - $\sum_{g'\in\mathcal{G}}\exp(\theta^T\mathbf{t}(g',X))$: normalizing factor

- **Intuition: ERG placed more/less weight on structures with certain features, as determined by $\mathbf{t}$, θ**
  - Framework is complete for pmfs on $\mathcal{G}$, few constraints on $\mathbf{t}$

Exponential Families for Random Graphs (w/Covariates)

- **For random graph G w/countable support $\mathcal{G}$ and covariate set X, pmf can be written in ERG form:**

$$\Pr\left(G=g \middle| \mathbf{t}, \theta, \mathcal{G}, X\right) = \frac{\exp\left(\theta^T \mathbf{t}(g, X)\right)}{\sum_{g' \in \mathcal{G}} \exp\left(\theta^T \mathbf{t}(g', X)\right)} I_{\mathcal{G}}(g)$$

  - $\theta^{T}\mathbf{t}$: linear predictor
    - $\mathbf{t}: \mathcal{G} \rightarrow \mathbb{R}^m$: vector of sufficient statistics
    - $\theta \in \mathbb{R}^m$: vector of parameters
    - $\sum_{g' \in \mathcal{G}} \exp(\theta^{T}\mathbf{t}(g', X))$: normalizing factor

Source of great difficulty!

- **Intuition: ERG placed more/less weight on structures with certain features, as determined by $\mathbf{t}$, θ**
  - Framework is complete for pmfs on $\mathcal{G}$, few constraints on $\mathbf{t}$

Equivalent Expression: Modeling the Adjacency Matrix

- **For adjacency matrix $\mathbf{Y}$ w/countable support $\mathcal{Y}$ and covariate set X, <u>any</u> pmf can be written in ERG form:**

$$\Pr\left(\mathbf{Y}=\mathbf{y}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)=\frac{\exp\left(\theta^T\mathbf{t}(\mathbf{y},X)\right)}{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^T\mathbf{t}(\mathbf{y}',X)\right)}I_{\mathcal{Y}}(\mathbf{y})$$

  - $\theta^T\mathbf{t}$: linear predictor
    - $\mathbf{t}:\mathcal{Y}\rightarrow\mathbb{R}^m$: vector of sufficient statistics
    - $\theta\in\mathbb{R}^m$: vector of parameters
    - $\sum_{\mathbf{y}'\in\mathcal{Y}}\exp(\theta^T\mathbf{t}(\mathbf{y}',X))$: normalizing factor
- **Additional notation:**
  - $\mathbf{y}^+_{\mathbf{ij}},\mathbf{y}^-_{\mathbf{ij}}$: y w/$ij$th cell set to 1 or 0 (respectively)
  - $\mathbf{y}^c_{\mathbf{ij}}$: all elements of y other than the ijth

ERGs and Conditional Odds of an Edge

- **Can easily specify the conditional odds of an edge:**

$$\frac{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{+}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{-}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)} = \frac{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{+},X)\right)}{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}',X)\right)}\frac{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}',X)\right)}{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{-},X)\right)}$$

ERGs and Conditional Odds of an Edge

- **Can easily specify the conditional odds of an edge:**

$$\frac{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{+}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{-}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)} = \frac{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{+},X)\right)}{\cancel{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}',X)\right)}} \frac{\cancel{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}',X)\right)}}{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{-},X)\right)}$$

$$= \frac{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{+},X)\right)}{\exp\left(\theta^{T}\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{-},X)\right)} = \exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{+},X)-\mathbf{t}(\mathbf{y}_{\mathbf{ij}}^{-},X)\right)\right)$$

ERGs and Conditional Odds of an Edge

- **Can easily specify the conditional odds of an edge:**

$$\frac{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{+}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}_{\mathbf{ij}}^{-}\middle|\mathbf{t},\theta,\mathcal{Y},X\right)} = \frac{\exp\left(\theta^T\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{+},X\right)\right)}{\cancel{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^T\mathbf{t}\left(\mathbf{y}',X\right)\right)}}\frac{\cancel{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left(\theta^T\mathbf{t}\left(\mathbf{y}',X\right)\right)}}{\exp\left(\theta^T\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{-},X\right)\right)}$$

$$= \frac{\exp\left(\theta^T\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{+},X\right)\right)}{\exp\left(\theta^T\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{-},X\right)\right)} = \exp\left(\theta^T\left(\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{+},X\right)-\mathbf{t}\left(\mathbf{y}_{\mathbf{ij}}^{-},X\right)\right)\right)$$

  - Log-odds depend on the changescore, $\Delta_{ij}=\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}},X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}},X)$
  - Useful implication: each unit change in $\mathbf{t}_{\mathbf{k}}$ for (i,j) edge present (versus absent) increases the conditional log-odds of (i,j) by θ_k
    - ***Important***: this is only conditionally true! The marginal log-odds of an (i,j) edge can depend on a complex way on other aspects of the graph

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) = \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right)$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) = \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right)$$

$$= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)} \left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}\right]^{-1}$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$
\begin{aligned}
\Pr\left(\mathbf{Y}_{ij}=1 \middle| \mathbf{Y}^{c}_{ij}=\mathbf{y}^{c}_{ij}, \mathbf{t}, \theta, \mathcal{Y}, X\right) &= \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{ij} \middle| \mathbf{Y}^{c}_{ij}=\mathbf{y}^{c}_{ij}, \mathbf{t}, \theta, \mathcal{Y}, X\right) \\
&= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{ij} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{ij} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)} \left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{ij} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{ij} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}\right]^{-1} \\
&= \frac{\exp\left(\theta^{T}\left(\mathbf{t}\left(\mathbf{y}^{+}_{ij}, X\right)-\mathbf{t}\left(\mathbf{y}^{-}_{ij}, X\right)\right)\right)}{1+\exp\left(\theta^{T}\left(\mathbf{t}\left(\mathbf{y}^{+}_{ij}, X\right)-\mathbf{t}\left(\mathbf{y}^{-}_{ij}, X\right)\right)\right)}
\end{aligned}
$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$
\begin{aligned}
\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) &= \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) \\
&= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)} \left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}\right]^{-1} \\
&= \frac{\exp\left(\theta^{T}\left(\mathbf{t}\left(\mathbf{y}^{+}_{\mathbf{ij}}, X\right)-\mathbf{t}\left(\mathbf{y}^{-}_{\mathbf{ij}}, X\right)\right)\right)}{1+\exp\left(\theta^{T}\left(\mathbf{t}\left(\mathbf{y}^{+}_{\mathbf{ij}}, X\right)-\mathbf{t}\left(\mathbf{y}^{-}_{\mathbf{ij}}, X\right)\right)\right)} \\
&= \left[1+\exp\left(\theta^{T}\left(\mathbf{t}\left(\mathbf{y}^{-}_{\mathbf{ij}}, X\right)-\mathbf{t}\left(\mathbf{y}^{+}_{\mathbf{ij}}, X\right)\right)\right)\right]^{-1}
\end{aligned}
$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$
\begin{aligned}
\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) &= \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) \\
&= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)} \left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}\right]^{-1} \\
&= \frac{\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)}{1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)} \\
&= \left[1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)\right)\right)\right]^{-1} \\
&= \operatorname{logit}^{-1}\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)
\end{aligned}
$$

Conditional Edge Probability

- **From the above, we can also determine conditional edge probability using fact that prob=odds/(1+odds):**

$$
\begin{aligned}
\Pr\left(\mathbf{Y}_{\mathbf{ij}}=1 \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) &= \Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{Y}^{\mathbf{c}}_{\mathbf{ij}}=\mathbf{y}^{\mathbf{c}}_{\mathbf{ij}}, \mathbf{t}, \theta, \mathcal{Y}, X\right) \\
&= \frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)} \left[1+\frac{\Pr\left(\mathbf{Y}=\mathbf{y}^{+}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}{\Pr\left(\mathbf{Y}=\mathbf{y}^{-}_{\mathbf{ij}} \middle| \mathbf{t}, \theta, \mathcal{Y}, X\right)}\right]^{-1} \\
&= \frac{\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)}{1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)} \\
&= \left[1+\exp\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)\right)\right)\right]^{-1} \\
&= \operatorname{logit}^{-1}\left(\theta^{T}\left(\mathbf{t}(\mathbf{y}^{+}_{\mathbf{ij}}, X)-\mathbf{t}(\mathbf{y}^{-}_{\mathbf{ij}}, X)\right)\right)
\end{aligned}
$$

Conditional Edge Probability, Cont.

- **So, the conditional probability of an (i,j) edge is simply the inverse logit of $\theta^T\Delta_{ij}$**
  - Obvious idea: to find θ, why not set this up as a logistic network regression problem (regressing $\mathbf{y}$ on Δ)?
    - This is an "autologistic regression," and the resulting estimator is known as a *pseudolikelihood* estimator (Besag, 1975)
  - Problem: the probability here is only conditional – can use for any one ij, but joint likelihood of $\mathbf{y}$ is not generally the product of $\Pr(\mathbf{Y}_{ij}=\mathbf{y}_{ij}|\mathbf{Y}^c_{ij}=\mathbf{y}^c_{ij})$
    - Another view: $\mathbf{y}$ appears on both sides – can't regress w/out accounting for the "feedback" (i.e., dependence) among edges
    - Does work iff edges are independent – i.e., the logistic case
- **Still, useful aid in interpretation**
  - Can consider probability of $\mathbf{y}_{ij}$ under various scenarios to understand *local* model behavior

Fitting ERGs to Data

- **After positing a model, generally want to estimate parameters from data**
  - Benefit of the framework: standardized inferential framework
- **Most common current method: maximum likelihood estimation**
  - Find θ^* that maximizes $\Pr(\mathbf{Y}=\mathbf{y}_{obs}|\mathbf{t},\theta^*,\mathcal{Y})$
    - Exists for regular model so long as observed data is "non-extreme" (in a specific sense to be discussed later); always unique
    - Also has first order interpretation: $\mathbf{E}_{\theta*}\mathbf{t}(\mathbf{Y})=\mathbf{t}(\mathbf{y}_{obs})$
  - Approximate standard errors based on Hessian of maximized likelihood
    - Not guaranteed, but simulation studies suggest to be reasonable
- **Important point: estimation is conditional on the support**
  - Often taken for granted, but (as we shall see!) can be consequential

ERG Fitting Using `ergm`

- **Dedicated statnet package for fitting, simulating models in ERG form**
- **Basic call structure: `ergm(y~term1(arg)+term2(arg))`**
  - `y`: network object
  - `term1, term2`, etc.: terms to use (see `?"ergm-terms"`)
    - `arg`: where relevant, arguments to the term functions
- **Output: ergm object**
  - Summary, print, and other methods can be used to examine it
  - `simulate` command can also be used to take draws from the fitted model

ERGM Inference from Partially Observed Data

- **What if $\mathbf{y}_{obs}$ is not an entire adjacency matrix, but only a subset thereof?**
- **If pattern of missingness is ignorable, natural to use a "latent missing data approach" (Handcock and Gile, 2007)**
  - Treat unknown portion of $\mathbf{y}_{obs}$ as latent variables; by assumption, must be drawn from the estimated model
  - Conditional probability of the unknown data is similar to original ERGM, with restricted normalizing factor (sum is over all observable "completions" of $\mathbf{y}_{obs}$):

$$\Pr(\mathsf{Y}^{\bar{o}} = \mathsf{y}^{\bar{o}} | \mathsf{Y}^{o} = \mathsf{y}^{o}, X, \mathcal{Y}) = \frac{\exp(\theta^T \mathsf{t}(\mathsf{y}^{\bar{o}} \cup \mathsf{y}^{o}, X))}{\sum_{y'^{\bar{o}}: y'^{\bar{o}} \cup y^{o} \in \mathcal{Y}} \exp(\theta^T \mathsf{t}(\mathsf{y}'^{\bar{o}} \cup \mathsf{y}^{o}, X))}$$

Partially Observed Data, Cont.

- **Observed data likelihood is then proportional to sum of likelihoods for "compatible" realizations of $\mathbf{y}_{obs}|\mathbf{y}^o$:**

$$\Pr(\mathsf{Y}^{\,o} = \mathsf{y}^o | X, \mathcal{Y}) \propto \frac{\sum_{y'^{\overline{o}}: y'^{\overline{o}} \cup y^o \in \mathcal{Y}} \exp(\theta^T \mathsf{t}(\mathsf{y}'^{\overline{o}} \cup \mathsf{y}^o, X))}{\sum_{\mathsf{y}' \in \mathcal{Y}} \exp(\theta^T \mathsf{t}(\mathsf{y}', X))}$$

  - Automatically computed by default in `ergm`
  - Coefficients, other properties have standard interpretation
  - Will lose power, of course! (Perhaps substantially)
- **Reminder: not a fix for biased samples!**
  - If missingness depends on parameters other than through X, $\mathbf{y}^o$, then results may be very misleading
    - E.g., sampling based on unobserved properties of missing nodes