

Análise bayesiana de pesquisas RDS com incerteza no desfecho

Lucas Moschen

Escola de Matemática Aplicada
Fundação Getulio Vargas

23 de junho de 2021

- ① Introdução
- ② Justificativa
- ③ Objetivos
- ④ Metodologia
- ⑤ Resultados preliminares
- ⑥ Cronograma

- ① Introdução
- ② Justificativa
- ③ Objetivos
- ④ Metodologia
- ⑤ Resultados preliminares
- ⑥ Cronograma

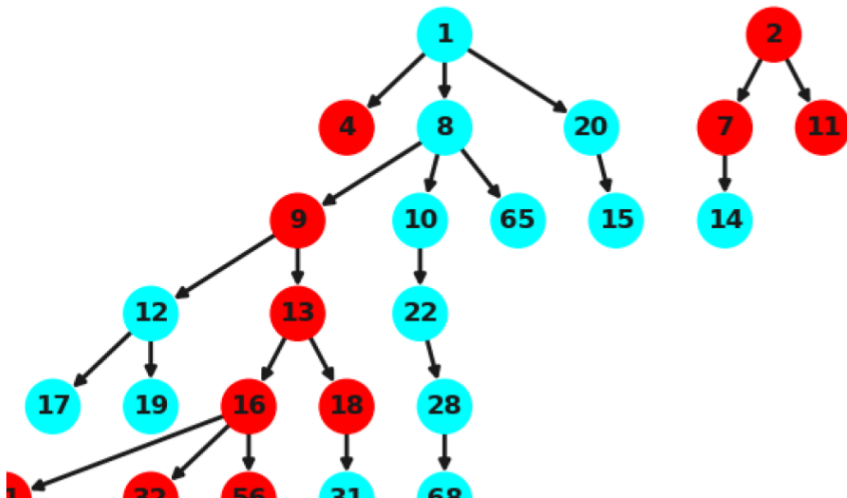
Populações de difícil acesso ou “escondidas”

- ▶ Não existe um esquema de amostragem clássico: tamanho e fronteiras da população são desconhecidos;
- ▶ Preocupações com a privacidade e medo de exposição: comportamento estigmatizado ou ilegal;
- ▶ Exemplos: Usuários de drogas pesadas, profissionais do sexo, pessoas em situação de rua, entre outros;
- ▶ Abordagens existentes de amostragem têm vários espaços para desenvolvimento.

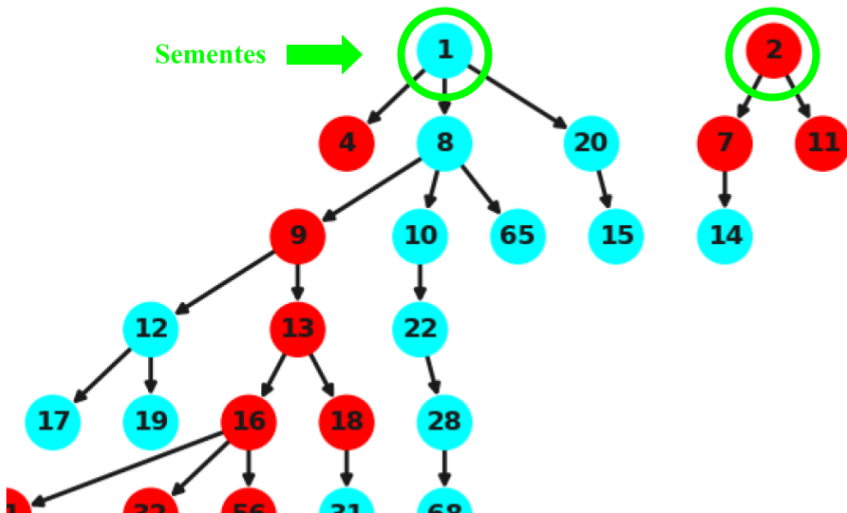
Respondent-driven sampling

- ▶ Proposta em [[Heckathorn, 1997](#)] como uma abordagem de estimar proporções em uma população alvo;
- ▶ Teoria baseada em cadeias de Markov;
- ▶ [[Crawford, 2016](#)] modela como uma rede com interações e nós faltantes e define uma distribuição de probabilidade sobre o subgrafo observado;
- ▶ Amostragem sem reposição.

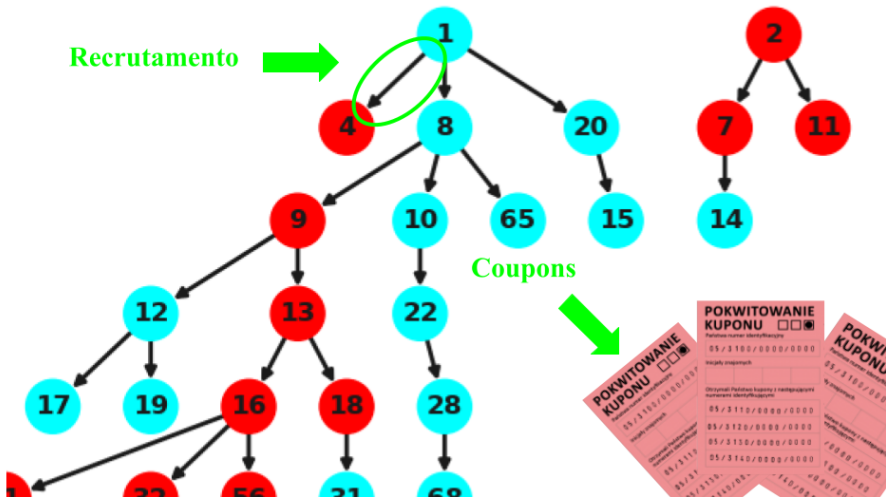
Estudo RDS: refugiados e ativistas na Síria



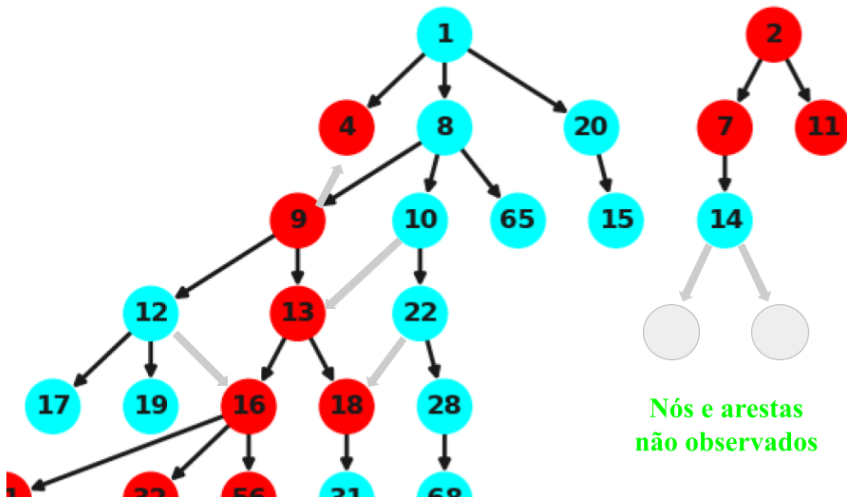
Estudo RDS: refugiados e ativistas na Síria



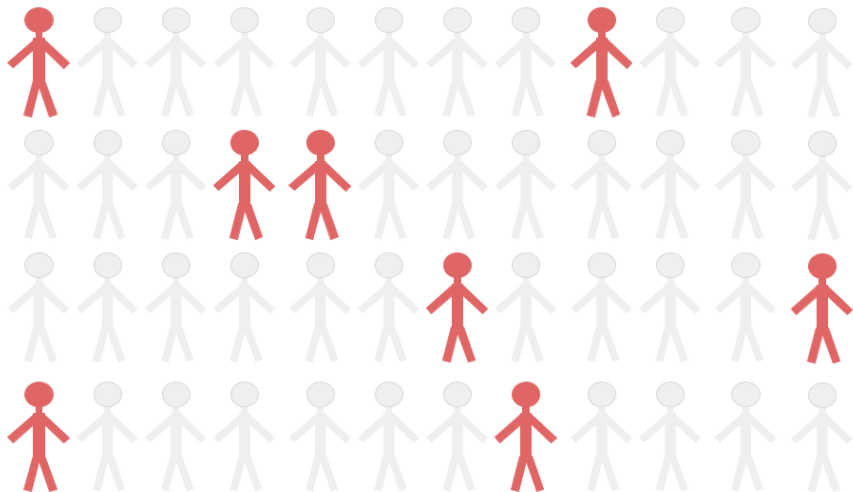
Estudo RDS: refugiados e ativistas na Síria



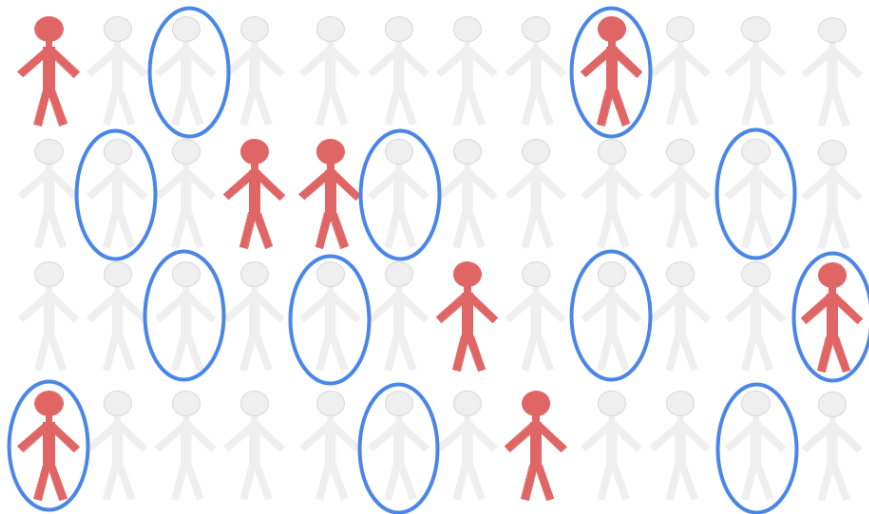
Estudo RDS: refugiados e ativistas na Síria



Estimação de prevalência com testes imperfeitos



Estimação de prevalência com testes imperfeitos



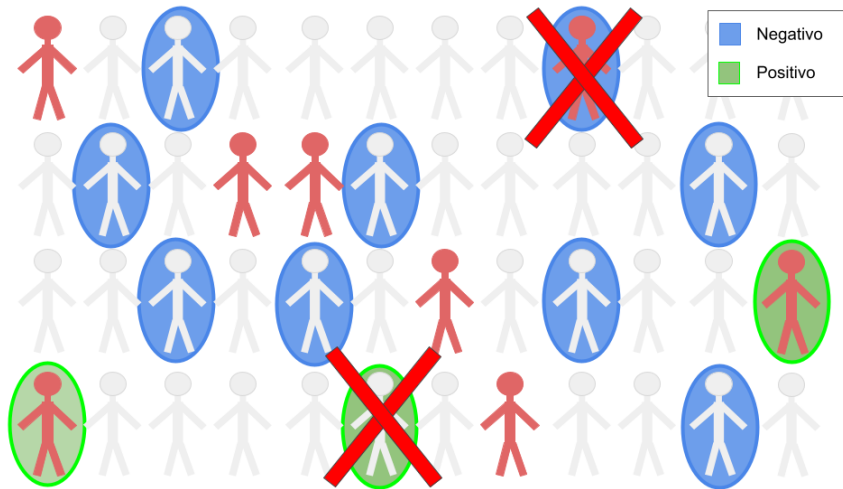
Estimação de prevalência com testes imperfeitos



Especificidade e Sensibilidade

Especificidade: Probabilidade de resultado negativo nos não doentes.

Sensibilidade: Probabilidade de resultado positivo nos doentes.



- ▶ Interpretação baseada no grau de crença em uma afirmação por um indivíduo;
- ▶ A fórmula de Bayes relaciona a probabilidade de um parâmetro após observar novos dados com a evidência e a informação prévia sobre ele;
- ▶ Permite a quantificação da incerteza de uma forma direta, dado que o processo não precisa ser aleatório.

- 1 Introdução
- 2 Justificativa
- 3 Objetivos
- 4 Metodologia
- 5 Resultados preliminares
- 6 Cronograma

- ▶ Populações escondidas são sub-representadas em pesquisas nacionais e têm maior risco de abuso de drogas ou contrair infecções sexualmente transmissíveis;
- ▶ O tópico tem vários gaps na estatística e abordagens de regressão para estimar prevalência considerando a estrutura de rede podem ser construídas [[Bastos et al., 2012](#)].

- 1 Introdução
- 2 Justificativa
- 3 Objetivos**
- 4 Metodologia
- 5 Resultados preliminares
- 6 Cronograma

Este trabalho tem o objetivo de estudar o problema da estimação de prevalência em uma estrutura de rede RDS considerando a Especificidade e a Sensitividade do diagnóstico.

Também se pretende aplicar esse framework de forma eficiente, comparando algoritmos Monte Carlo e Aproximações de Laplace.

- 1 Revisão bibliográfica, descrição matemática do problema e propagação da incerteza com métodos Bayesianos;
- 2 Distribuição conjunta a priori da especificidade e sensibilidade;
- 3 Implementação eficiente utilizando pacotes estatísticos, como *rstanarm* e *INLA*;
- 4 Análise de estudos epidemiológicos RDS.

- 1 Introdução
- 2 Justificativa
- 3 Objetivos
- 4 Metodologia
- 5 Resultados preliminares
- 6 Cronograma

Pesquisa bibliográfica

A fundamentação teórica se dará por meio de artigos nos tópicos indicados na introdução: RDS, estimativa de prevalência por meio de regressão, e estatística bayesiana.

Recursos técnicos

Toda a programação necessária será feita nas linguagens de programação *Python* e *R*.

Estudo formal

Disciplinas do Doutorado em Modelagem Matemática da EMap: Estatística Bayesiana e Ciências de Redes.

- ① Introdução
- ② Justificativa
- ③ Objetivos
- ④ Metodologia
- ⑤ Resultados preliminares
- ⑥ Cronograma

Observe also that for a network the independence of the sample is also not valid. Other estimators are better. Adicionar o modelo 1 e comentar sobre a beta bivariada.

Se π é a probabilidade do teste ser positivo,




$$\pi = \theta\gamma_s + (1 - \theta)(1 - \gamma_e),$$

e podemos estudar θ considerando γ_s, γ_e .

- ① Introdução
- ② Justificativa
- ③ Objetivos
- ④ Metodologia
- ⑤ Resultados preliminares
- ⑥ Cronograma

Cronograma

References I

-  Bastos, L. S., Pinho, A. A., Codeço, C., and Bastos, F. I. (2012). Binary regression analysis with network structure of respondent-driven sampling data.
-  Crawford, F. W. (2016). The graphical structure of respondent-driven sampling. *Sociological Methodology*, 46(1):187–211.
-  Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, 44(2):174–199.