Mechanistic Modeling of Social Conditions in Disease-Prediction Simulations via Copula-Informed Probabilistic Graphical Models: HIV Case Study

Amir Khosheghbal[1] **Peter J. Haas**[1] **Chaitra Gopalappa**[1]
[1]University of Massachusetts Amherst
akhosheghbal@umass.edu, phaas@cs.umass.edu, chaitrag@umass.edu

Abstract

This work has already been published; here, we provide a brief overview [1]. Epidemic models typically simulate the spread of diseases as functions of behaviors, e.g., sexual and care behaviors for sexually transmitted diseases. However, multilevel factors, including poverty, housing or food insecurity, mental health, substance use disorder, etc., which are called social determinants of health (SDH), are drivers of those behaviors. There is increasing awareness of the need to incorporate SDH into epidemic simulation models to evaluate structural interventions alongside behavioral interventions. However, the multivariate joint associations between SDH and behaviors needed for modeling are not available. Data for SDH are mostly available as county-level marginal distributions, and associations between SDH and behaviors are mostly bivariate. To address this problem, we combine copula theory and probabilistic graphical models to estimate multivariate joint distributions. We estimate bivariate associations between SDH using a novel copula approach that transitions from continuous to discrete copulas.We then use these bivariate associations—together with bivariate associations between SDH and behaviors from the literature—as links in an undirected graphical model to calculate the multivariate joint distributions. As a case study, we used the joint distributions to model HIV-risk related behaviors as function of SDH in a national-level HIV/AIDS (PATH 4.0) model and studied the impact of hypothetical 100% efficacious SDH interventions on HIV prevention. We found that this intervention could lead to a cumulative 10-year reduction of 29% in HIV incidence.

1 Background

Although there has been significant progress in treatments for human immunodeficiency virus (HIV) such as antiretroviral therapy (ART) and pre-exposure prophylaxis (PrEP) that can prevent HIV transmission or acquisition [2, 3], HIV continues to impose substantial disease and economic burden, with an estimated 1.2 million people living with HIV (PWH) in 2019 and approximately 35,000 new infections in 2019 [4]. The average discounted lifetime HIV-related medical cost per person is estimated at $420,285 [5]. There is growing evidence highlights social and economic conditions are key drivers of behaviors that increase HIV risk, e.g., lower adherence to care, higher number of partners, and higher condomless sex [6, 7, 8, 9]. Surveillance data indicate that among persons with diagnosed HIV (PWDH), about 44% had a physical, mental, or emotional disability, 14% were unemployed, 43% lived at or below the federal poverty line, and 10% were homeless [10, 11, 12]. Structural interventions, including programs like comprehensive sex education, universal condom availability, expanded syringe access, healthcare coverage, subsidized housing, and mental healthcare access [13, 14, 15, 16, 17, 18], are key evidence-based recommendations for HIV prevention

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: MLxOR: Mathematical Foundations and Operational Integration of Machine Learning for Uncertainty-Aware Decision-Making.

[19]. While the costs of these interventions can be estimated, their population-level impact on HIV burden is difficult to assess through controlled trials alone. Mechanistic models, which simulate HIV projections based on care and sexual behaviors, can be used to evaluate intervention combinations and inform resource allocation, but typically do not incorporate SDH [20, 21, 22, 23, 24]. The key challenge is the unavailability of coherent statistical data for social conditions and behaviors, specifically the joint probability distributions that capture the complex associations between various social determinants and their impact on health behaviors. Data on social conditions are often reported as marginal distributions (e.g., proportions in poverty or homelessness), but their joint distributions (e.g., proportion of the population both in poverty and homeless) are largely unavailable. Similarly, available data between behaviors and social conditions are mostly limited to bivariate associations. We address this data gap by proposing a novel methodology that combines copula theory with undirected graphical models to first estimate the joint probability distributions and then developing a framework for incorporating social conditions into dynamic mechanistic simulation models. The numerical analysis includes care behavior using HIV Viral Load Suppression (VLS) as a proxy to evaluate the impact on HIV incidence from interventions related to care behaviors.

2 Method

2.1 Problem description

Mechanistic models in epidemiology often simulate disease outcomes as functions of individual behaviors. However, research has shown that social factors are key drivers of high-risk behaviors. Taking HIV as an example, structural intervention programs such as housing for HIV infected persons facing housing instability promote HIV "treatment adherence" behavior. Treatment leads to Viral Load Suppression (VLS)—i.e., suppresses the level of virus in the blood— preventing the transmission of HIV.

Our objective is to explicitly represent such dependencies in HIV simulations by modeling behaviors (e.g., VLS) as functions of SDH variables, for decisions analysis related to the need for and impact of structural interventions. This requires a joint distribution of multiple social and behavioral factors, which is typically unavailable. Data on the prevalence of social conditions are reported mostly as marginal distributions, e.g., proportions in poverty, or homeless rates across different jurisdictions coming from disparate datasets.

We assume that all variables are binary, where 0 represents a good status (e.g., housed) and 1 a disadvantaged status (e.g., homeless). When pairwise associations, e.g., relative risks or odds ratios are available or can be estimated by bivariate samples, we can reconstruct the multivariate joint distribution using undirected graphical modeling, also known as Markov Random Field (MRF) [25], or directly applying iterative proportional fitting (IPF) [26, 27]. The problem is to thus estimate the joint associations between all variables and then use them to enhance simulation modeling. Below we propose a method to estimate pairwise joint distributions from dependent marginal samples that can then be used to construct the multivariate joint distributions via MRF or IPF.

2.2 Proposed continuous to discrete copula to transform dependent continuous marginals to discrete joint distributions

This section addresses scenarios where odds ratios are unavailable, but dependent marginal samples, such as unemployment and poverty rates across counties, are observed. The goal is to estimate the pairwise joint distributions $\Pr(X, Y)$ for binary random variables $X, Y \in \{0, 1\}$, where $X \sim \text{Bernoulli}(\pi_X)$ and $Y \sim \text{Bernoulli}(\pi_Y)$.

The dependency structure of bivariate random variables can be represented by a family of odds ratios [28]. For two binary random variables (X, Y) with joint probability mass function (pmf) $p = \left(\begin{smallmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{smallmatrix}\right)$, the odds ratio is defined by

$$w = \frac{p_{00}p_{11}}{p_{01}p_{10}}. \tag{1}$$

Extending this to a bivariate discrete vector (X, Y) where $X \in S_X = \{0, 1, \ldots, R\}$ and $Y \in S_Y = \{0, 1, \ldots, S\}$, with $R, S \in \mathbb{N}$ and $1 \leq R, S < \infty$, the dependency takes the form of

$$w_{xy} = \frac{p_{00}p_{xy}}{p_{0y}p_{x0}}, \quad \forall (x,y) \in S_X \setminus \{0\} \times S_Y \setminus \{0\}. \tag{2}$$

Suppose we have n independent copies of a bivariate Bernoulli random variable (X, Y), forming a bivariate binomial vector $(x, y) = (\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i)$ with pmf:

$$\Pr(X = x, Y = y) = \sum_{k=\max(x+y-n,0)}^{\min(x,y)} \binom{n}{k, x-k, y-k, n-x-y+k} p_{00}^{n-x-y+k} p_{10}^{x-k} p_{01}^{y-k} p_{11}^{k}.$$

Then, as per [28], the odds ratios for the binomial distribution can be estimated as:

$$w_{xy} = \sum_{k=\max(x+y-n,0)}^{\min(x,y)} \frac{\binom{n}{k,x-k,y-k,n-x-y+k}}{\binom{n}{x}\binom{n}{y}} w^k, \tag{3}$$

where w is given by Equation (1). Given the marginals π_X, π_Y, and the odds ratio w, the bivariate Bernoulli distribution can be recovered as:

$$p_{11} = \frac{1}{2(w-1)}\{1+(w-1)(\pi_X+\pi_Y) - \sqrt{[1+(w-1)(\pi_X+\pi_Y)]^2 - 4w(w-1)\pi_X\pi_Y}\} \tag{4}$$

and subsequently, $p_{00} = 1 - \pi_X - \pi_Y + p_{11}$; $p_{10} = \pi_X - p_{11}$; $p_{01} = \pi_Y - p_{11}$ via equations introduced in [28]. Suppose that we are missing the odds ratio w for some pair (X, Y), but do have aggregate data $\{(x_j, y_j)\}_{j\in[1..J]}$, where x_j and y_j are marginal proportions of people with these respective social burdens in jurisdiction j. Then we can estimate the odds ratio as follows.

If we assume that each jurisdiction has the same population size n, then $\{l_j = n \cdot x_j\}_{j\in[1..J]}$ and $\{m_j = n \cdot y_j\}_{j\in[1..J]}$ can be considered as samples from $\mathrm{Bin}(n, \pi_X)$ and $\mathrm{Bin}(n, \pi_Y)$, where π_X and π_Y are the Bernoulli parameters at the national level. By the central limit theorem, for sufficiently large n, these counts approximately follow a Gaussian distribution:

$$l_j \sim \mathcal{N}(n\pi_X, \sqrt{n\pi_X(1-\pi_X)}), \quad m_j \sim \mathcal{N}(n\pi_Y, \sqrt{n\pi_Y(1-\pi_Y)}).$$

The dependency between Bernoulli X and Y carries over to l_j and m_j, so the joint distribution ($p_{l_j m_j}$) can be estimated by a continuous copula as described by Sklar's theorem [29], which expresses the joint CDF and joint pdf as

$$F_{LM}(l_j, m_j) = C\big(F_L(l_j), F_M(m_j)\big) \quad \text{and} \quad f_{LM}(l_j, m_j) = c\big(F_L(l_j), F_M(m_j)\big) f_L(l_j) f_M(m_j)$$

for a copula C with density c. Specifically, we apply a Gaussian copula density function fitted to jurisdiction data to estimate the $p_{l_j m_j}$:

$$c(u, v) = \frac{1}{\sqrt{1-r^2}} \exp\left(-\frac{(a^2+b^2)r^2 - 2abr}{2(1-r^2)}\right),$$

where $a = \sqrt{2}\,\mathrm{erf}^{-1}(2u-1)$, $b = \sqrt{2}\,\mathrm{erf}^{-1}(2v-1)$, and r is the empirical correlation coefficient of the (l_j, m_j) pairs. We then compute the $w_{l_j m_j}$ via Equation (2) and solve for the underlying Bernoulli odds ratio w_j via Equation (3). We then calculate the bivariate Bernoulli distribution using Equation (4). By treating the national-level data as another "jurisdiction", we also estimate the national-level bivariate Bernoulli distribution and use that to estimate the multivariate joint distribution, which we further use in a national-level simulation model as discussed below.

2.3 Validation of copula approach

We validated the proposed continuous to discrete copula method using individual-level American Community Survey (ACS) data for poverty and employment from IPUMS [30]. We estimated the joint distribution between these features for each county, and used a chi-squared test, which showed a good fit for approximately 95% of the samples compared to actual bivariate distributions. We also applied our method to a synthetic dataset having high inter-feature dependencies, which similarly indicated a good fit.

2.4 Numerical application for care behavioral model

We chose VLS as a proxy for overall care behavior as VLS is achieved by linking-to-care at diagnosis and consistent retention-in-care and treatment. We included the social conditions and intermediary variable as listed in Figure 1. We applied the proposed copula method to those associations where we only had the jurisdictional data available to estimate the pairwise joint distributions and used them along with other associations available from the literature to then estimate the multivariate joint distribution using an MRF graphical model by solving through maximum entropy. Using the estimated joint distribution, we can estimate the conditional probabilities $\Pr(VLS \mid SDH)$ for any combination of SDH. As expected, estimates indicate that the probability of VLS is lower for individuals with at least one socially disadvantaged condition compared to those with no social conditions. We evaluated a hypothetical intervention assuming that the chance of VLS for people with social needs could be increased to become equal to that among those with no social needs, to determine the corresponding impact of such an improvement. Two scenarios were simulated in the PATH 4.0 [31], which is a national simulation model of HIV in the United States; a baseline scenario using status-quo marginal distributions $\Pr(VLS)$ and an intervention scenario using $\widehat{\Pr}(VLS)$.

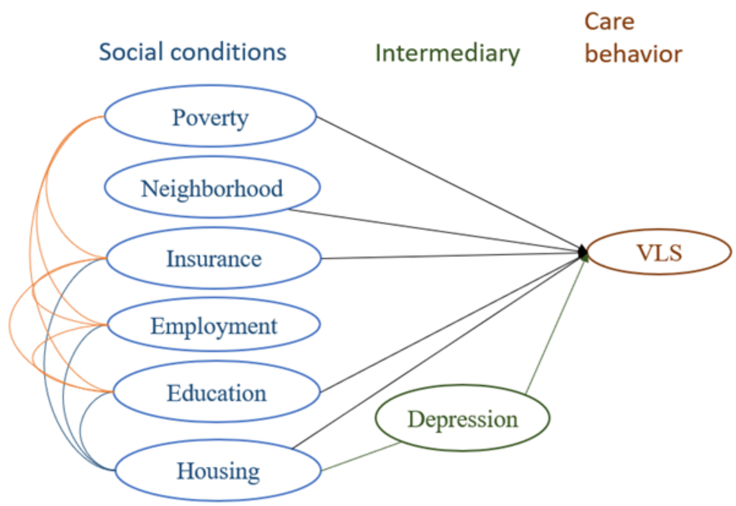

Figure 1: Illustration of associations between social conditions and care behavior. Note: all variables are binary. The orange links relate to pairwise associations calculated using the proposed method. All the other associations are calculated using literature data.

3 Results

The estimated joint distribution indicated that 78% of all people with HIV had at least one social burden. These estimates were higher in people not in HIV care (87%) compared those in HIV care (73%). From the simulation runs, a hypothetical 100% efficacious intervention addressing all social needs is projected to increase VLS among people with diagnosed HIV (PWDH) from a baseline of 65.5% to 79%, leading to an estimated 29% (20% to 41%) reduction in cumulative national HIV incidence over a 10-year simulation period.

4 Conclusions

We presented a new method combining copulas with graphical models to estimate joint distributions from dependent marginal samples. Using this method, we enhanced the prior PATH simulation framework for analysis of structural interventions alongside behavioral and pharmaceutical interventions, thereby providing an improved decision analysis tool to inform public policy.

5 Declaration

This work was supported by the National Science Foundation, United States, under NSF 1915481. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

References

[1] Amir Khosheghbal, Peter J Haas, and Chaitra Gopalappa. Mechanistic modeling of social conditions in disease-prediction simulations via copulas and probabilistic graphical models: HIV case study. *Health Care Manag. Sci.*, 28(1):28–49, March 2025. URL https://doi.org/10.1007/s10729-024-09694-3.

[2] HIV Treatment, Division of HIV Prevention, National Center for HIV, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Https://Www.Cdc.Gov/Hiv/Basics/Livingwithhiv/Treatment.Html [Last Accessed 2/14/2023], .

[3] PrEP Effectiveness, Division of HIV Prevention, National Center for HIV, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Https://Www.Cdc.Gov/Hiv/Basics/Prep/Prep-Effectiveness.Html [Last Accessed 2/14/2023], .

[4] Centers for Disease Control and Prevention. Estimated HIV Incidence and Prevalence in the United States, 2015–2019. HIV Surveillance Supplemental Report 2021;26(No. 1). Http://Www.Cdc.Gov/Hiv/Library/Reports/Hiv-Surveillance.Html. Published May 2021. Accessed [11/20/2023], .

[5] Adrienna Bingham, Ram K. Shrestha, Nidhi Khurana, Evin U. Jacobson, and Paul G. Farnham. Estimated Lifetime HIV–Related Medical Costs in the United States. *Sexually Transmitted Diseases*, 48 (4):299–304, April 2021. ISSN 1537-4521, 0148-5717. doi: 10.1097/OLQ.0000000000001366. URL https://journals.lww.com/10.1097/OLQ.0000000000001366.

[6] Jaih B. Craddock, Anamika Barman-Adhikari, Katie Massey Combs, Anthony Fulginiti, and Eric Rice. Individual and Social Network Correlates of Sexual Health Communication Among Youth Experiencing Homelessness. *AIDS and Behavior*, 24(1):222–232, January 2020. ISSN 1090-7165, 1573-3254. doi: 10.1007/s10461-019-02646-x. URL http://link.springer.com/10.1007/s10461-019-02646-x.

[7] Benjamin F. Henwood, Harmony Rhoades, Brian Redline, Eldin Dzubur, and Suzanne Wenzel. Risk behaviour and access to HIV/AIDS prevention services among formerly homeless young adults living in housing programmes. *AIDS Care*, 32(11):1457–1461, November 2020. ISSN 0954-0121, 1360-0451. doi: 10.1080/09540121.2019.1699643. URL https://www.tandfonline.com/doi/full/10.1080/09540121.2019.1699643.

[8] Diane Santa Maria, Saumali S. Daundasekara, Daphne C. Hernandez, Wei Zhang, and Sarah C. Narendorf. Sexual risk classes among youth experiencing homelessness: Relation to childhood adversities, current mental symptoms, substance use, and HIV testing. *PLOS ONE*, 15 (1):e0227331, January 2020. ISSN 1932-6203. doi: 10.1371/journal.pone.0227331. URL https://dx.plos.org/10.1371/journal.pone.0227331.

[9] Jennifer P. Edidin, Zoe Ganim, Scott J. Hunter, and Niranjan S. Karnik. The Mental and Physical Health of Homeless Youth: A Literature Review. *Child Psychiatry & Human Development*, 43 (3):354–375, June 2012. ISSN 0009-398X, 1573-3327. doi: 10.1007/s10578-011-0270-1. URL http://link.springer.com/10.1007/s10578-011-0270-1.

[10] Ya-Lin A. Huang, Emma L. Frazier, Stephanie L. Sansom, Paul G. Farnham, Ram K. Shrestha, Angela B. Hutchinson, Jennifer L. Fagan, Abigail H. Viall, and Jacek Skarbinski. Nearly Half Of US Adults Living With HIV Received Federal Disability Benefits In 2009. *Health Affairs*, 34(10): 1657–1665, October 2015. ISSN 0278-2715, 1544-5208. doi: 10.1377/hlthaff.2015.0249. URL http://www.healthaffairs.org/doi/10.1377/hlthaff.2015.0249.

[11] Centers for Disease Control and Prevention. Behavioral and Clinical Characteristics of Persons with Diagnosed HIV Infection—Medical Monitoring Project, United States, 2018 Cycle (June 2018 May 2019). HIV Surveillance Special Report 25. Https://Www.Cdc.Gov/Hiv/Library/Reports/Hiv Surveillance.Html. Published May 2020. Accessed 10/11/2024., .

[12] Centers for Disease Control and Prevention. Behavioral and Clinical Characteristics of Persons with Diagnosed HIV Infection—Medical Monitoring Project, United States, 2021 Cycle (June 2021–May 2022). HIV Surveillance Special Report 32. Https://Stacks.Cdc.Gov/View/Cdc/149083. Published August 2023. Accessed 11/14/2024., .

[13] Adaora A Adimora and Judith D Auerbach. Structural Interventions for HIV Prevention in the United States. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 55(Supplement 2): S132–S135, December 2010. ISSN 1525-4135. doi: 10.1097/QAI.0b013e3181fbcb38. URL `https://journals.lww.com/00126334-201012152-00016`.

[14] Theresa Ann Sipe, Terrika L. Barham, Wayne D. Johnson, Heather A. Joseph, Maria Luisa Tungol-Ashmon, and Ann O'Leary. Structural Interventions in HIV Prevention: A Taxonomy and Descriptive Systematic Review. *AIDS and Behavior*, 21(12):3366–3430, December 2017. ISSN 1090-7165, 1573-3254. doi: 10.1007/s10461-017-1965-5. URL `http://link.springer.com/10.1007/s10461-017-1965-5`.

[15] Eleanor E. Friedman, Hazel D. Dean, and Wayne A. Duffus. Incorporation of Social Determinants of Health in the Peer-Reviewed Literature: A Systematic Review of Articles Authored by the National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention. *Public Health Reports®*, 133(4):392–412, July 2018. ISSN 0033-3549, 1468-2877. doi: 10.1177/0033354918774788. URL `https://journals.sagepub.com/doi/10.1177/0033354918774788`.

[16] Kim M. Blankenship, Sarah J. Bray, and Michael H. Merson. Structural interventions in public health:. *AIDS*, 14:S11–S21, June 2000. ISSN 0269-9370. doi: 10.1097/00002030-200006001-00003. URL `http://journals.lww.com/00002030-200006001-00003`.

[17] K. M. Blankenship, S. R. Friedman, S. Dworkin, and J. E. Mantell. Structural Interventions: Concepts, Challenges and Opportunities for Research. *Journal of Urban Health*, 83(1): 59–72, March 2006. ISSN 1099-3460, 1468-2869. doi: 10.1007/s11524-005-9007-4. URL `http://link.springer.com/10.1007/s11524-005-9007-4`.

[18] Thomas R. Frieden. A Framework for Public Health Action: The Health Impact Pyramid. *American Journal of Public Health*, 100(4):590–595, April 2010. ISSN 0090-0036, 1541-0048. doi: 10.2105/AJPH.2009.185652. URL `https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2009.185652`.

[19] Compendium of Evidence-Based Interventions and Best Practices for HIV Prevention, Centers for Disease Control and Prevention, Https://Www.Cdc.Gov/Hiv/Research/Interventionresearch/Compendium/Si/Index.Html [Last Accessed August 2022]., .

[20] Joseph W. Hogan, Noya Galai, and Wendy W. Davis. Modeling the Impact of Social Determinants of Health on HIV. *AIDS and Behavior*, 25(S2):215–224, November 2021. ISSN 1090-7165, 1573-3254. doi: 10.1007/s10461-021-03399-2. URL `https://link.springer.com/10.1007/s10461-021-03399-2`.

[21] Robson Bruniera De Oliveira, Felipe Alves Rubio, Rodrigo Anderle, Mauro Sanchez, Luis Eugenio De Souza, James Macinko, Ines Dourado, and Davide Rasella. Incorporating social determinants of health into the mathematical modeling of HIV/AIDS. *Scientific Reports*, 12 (1):20541, November 2022. ISSN 2045-2322. doi: 10.1038/s41598-022-24459-0. URL `https://www.nature.com/articles/s41598-022-24459-0`.

[22] Marie C. D. Stoner, Jessie K. Edwards, Daniel Westreich, Kelly Kilburn, Jennifer Ahern, Sheri A. Lippman, F. Xavier Gómez-Olivé, Kathleen Kahn, and Audrey Pettifor. Modeling Cash Plus Other Psychosocial and Structural Interventions to Prevent HIV Among Adolescent Girls and Young Women in South Africa (HPTN 068). *AIDS and Behavior*, 25(S2):133–143, November 2021. ISSN 1090-7165, 1573-3254. doi: 10.1007/s10461-021-03158-3. URL `https://link.springer.com/10.1007/s10461-021-03158-3`.

[23] Kate Shannon, Steffanie A Strathdee, Shira M Goldenberg, Putu Duff, Peninah Mwangi, Maia Rusakova, Sushena Reza-Paul, Joseph Lau, Kathleen Deering, Michael R Pickles, and Marie-Claude Boily. Global epidemiology of HIV among female sex workers: influence of structural determinants. *The Lancet*, 385(9962):55–71, January 2015. ISSN 01406736. doi: 10.1016/S0140-6736(14)60931-4. URL `https://linkinghub.elsevier.com/retrieve/pii/S0140673614609314`.

[24] Davide Rasella, Gabriel Alves De Sampaio Morais, Rodrigo Volmir Anderle, Andréa Ferreira Da Silva, Iracema Lua, Ronaldo Coelho, Felipe Alves Rubio, Laio Magno, Daiane Machado, Julia Pescarini, Luis Eugênio Souza, James Macinko, and Inês Dourado. Evaluating the impact of social determinants, conditional cash transfers and primary health care on HIV/AIDS: Study protocol of a retrospective and forecasting approach based on the data integration with a cohort of 100 million Brazilians. *PLOS ONE*, 17(3):e0265253, March 2022. ISSN 1932-6203. doi: 10.1371/journal.pone.0265253. URL `https://dx.plos.org/10.1371/journal.pone.0265253`.

[25] Kevin P. Murphy. *Machine learning: a probabilistic perspective*. Adaptive computation and machine learning series. MIT Press, Cambridge, MA, 2012. ISBN 978-0-262-01802-9.

[26] Stephen E. Fienberg. An Iterative Procedure for Estimation in Contingency Tables. *The Annals of Mathematical Statistics*, 41(3):907–917, June 1970. ISSN 0003-4851. doi: 10.1214/aoms/1177696968. URL `http://projecteuclid.org/euclid.aoms/1177696968`.

[27] Yvonne Millicent Mahala Bishop, Stephen E. Fienberg, and Paul W. Holland. *Discrete multivariate analysis: theory and practice*. MIT Press, Cambridge, Mass., 6. printing edition, 1980. ISBN 978-0-262-02113-5 978-0-262-52040-9.

[28] Gery Geenens. Copula modeling for discrete random vectors. *Dependence Modeling*, 8 (1):417–440, January 2020. ISSN 2300-2298. doi: 10.1515/demo-2020-0022. URL `https://www.degruyter.com/document/doi/10.1515/demo-2020-0022/html`.

[29] RB Nelson. *An Introduction to Copulas*. Springer Series in Statistics. Springer New York, New York, NY, 2006. ISBN 978-0-387-28659-4. doi: 10.1007/0-387-28678-0. URL `http://link.springer.com/10.1007/0-387-28678-0`.

[30] Steven Ruggles, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rodgers, and Megan Schouweiler. IPUMS USA: Version 15.0, 2024. URL `https://www.ipums.org/projects/ipums-usa/d010.V15.0`.

[31] Sonza Singh, Anne Marie France, Yao-Hsuan Chen, Paul G. Farnham, Alexandra M. Oster, and Chaitra Gopalappa. Progression and transmission of HIV (PATH 4.0)-A new agent-based evolving network simulation for modeling HIV transmission clusters. *Mathematical Biosciences and Engineering*, 18(3):2150–2181, 2021. ISSN 1551-0018. doi: 10.3934/mbe.2021109. URL `http://www.aimspress.com/article/doi/10.3934/mbe.2021109`.