

Likelihood-Based Inference for Stochastic Models of Sexual Network Formation

Mark S. Handcock
Center for Statistics and the Social Sciences
Department of Statistics
University of Washington, Seattle, WA

James Holland Jones¹
Center for Studies in Demography and Ecology
Center for Statistics and the Social Sciences
Center for AIDS and Sexually Transmitted Diseases
University of Washington, Seattle, WA

29 January 2003
Working Paper #29
Center for Statistics and the Social Sciences
University of Washington
Box 354322
Seattle, WA 98195-4322, USA

¹Corresponding Author: Center for Statistics and the Social Sciences, Box 354320, Seattle, WA 98195-4320, email: jameshj@stat.washington.edu, Phone: 206-221-6874, Fax: 206-221-6873

Abstract

Sexually-Transmitted Diseases (STDs) constitute a major public health concern. Mathematical models for the transmission dynamics of STDs indicate that heterogeneity in sexual activity level allow them to persist even when the typical behavior of the population would not support endemicity. This insight focuses attention on the distribution of sexual activity level in a population. In this paper, we develop several stochastic process models for the formation of sexual partnership networks. Using likelihood-based model selection procedures, we assess the fit of the different models to three large distributions of sexual partner counts: (1) Rakai, Uganda, (2) Sweden, and (3) the USA. Five of the six single-sex networks were fit best by the negative binomial model. The American women's network was best fit by a power-law model, the Yule. For most networks, several competing models fit approximately equally well. These results suggest three conclusions: (1) no single unitary process clearly underlies the formation of these sexual networks, (2) behavioral heterogeneity plays an essential role in network structure, (3) substantial model uncertainty exists for sexual network degree distributions. Behavioral research focused on the mechanisms of partnership formation will play an essential role in specifying the best model for empirical degree distributions. We discuss the limitations of inferences from such data, and the utility of degree-based epidemiological models more generally.

1 Introduction

Sexually-Transmitted Diseases, including HIV/AIDS, constitute a major global public health concern. UNAIDS estimates that there were 40 million adults and children living with HIV/AIDS in 2001. In addition to being a major humanitarian calamity, the AIDS pandemic represents a substantial barrier to economic development in many resource-poor settings throughout the world.

STDs other than HIV have been labeled a “hidden epidemic” by the National Institute of Medicine (1), who estimate that the annual economic cost of STDs other than HIV/AIDS in the United States alone to be \$16.4 billion (1). While the prevalence of some of the traditional bacterial STDs such as Gonorrhea (*Neisseria gonorrhoeae*) and Syphilis (*Treponema pallidum*) have undergone steady decline over the last twenty years, others, such as Chlamydia (*Chlamydia trachomatis*) have increased in prevalence (2). Furthermore, some STDs on the verge of elimination have made dramatic reversals, re-establishing themselves as endemic infections (3).

Non-fatal viral STDs have extremely high prevalence globally and may be a significant co-factor in HIV infection. Genital herpes has a prevalence of up to 22% in the general population of the United States, (4) and can reach as high as 60% prevalent in some developing countries (5).

In addition to the morbidity caused by STDs themselves, the persistence of endemic bacterial STDs raises concerns for emerging infections and the evolution of anti-microbial resistance. Antibiotic-resistant strains of STD microbes pose a significant emerging health risk (1). For example, penicillinase-producing *Neisseria gonorrhoeae* (PPNG) and chromosomally mediated penicillin- and tetracycline-resistant *Neisseria gonorrhoeae* (CMRNG) have led to these drugs being abandoned. Furthermore, fluoroquinolone resistance is becoming increasingly common in both the developing and developed world, leaving dangerously few therapeutic options for the treatment of once curable diseases (6; 7; 8; 9). The maintenance of an endemic infection in which selection is able to favor resistant strains of microbes is a great concern for emerging infections generally because of the unusual mechanisms of selection in prokaryotes (10).

The control and eventual eradication of STDs is an important public health goal. Both mathematical and statistical models of infectious disease processes have proven to be invaluable tools for infectious disease epidemiology (11; 12). However, developing useful models for STDs presents a number of challenges. Prominent among these is characterizing population heterogeneity in sexual behavior. The average behavior of most populations is not sufficient either to allow an epidemic or maintain an endemic STD infection. Mathematical formalizations of STD infection dynamics indicate that heterogeneity in sexual behavior allows STDs, which would otherwise fade out given average behavior, to persist (13; 14; 15). The importance of behavioral heterogeneity has focused attention on the properties of the distribution of sexual partner number. In this paper, we will present an analysis of the statistical properties of empirical sexual partnership distributions. These distributions play a key role in the mathematical theory of STD transmission, and may hold the key to their control (16; 17).

1.1 Mathematical Epidemiology of STDs

In the standard theory of infectious disease transmission dynamics, (18; 14) the force of infection is an increasing function of the size of the population, and populations will exhibit threshold sizes below which an epidemic is impossible. These behaviors are a consequence of the traditional assumption of mass action. However, Anderson and colleagues (14) note that there is no reason to assume that the number of intimate contacts will increase with increasing population size.

In addition to threshold population sizes for epidemics, the dynamics of an epidemic are governed by a threshold parameter, R_0 , the basic reproductive number. R_0 represents the expected number of secondary cases produced by a single index case in a population of susceptibles. In the case of an unstructured population, R_0 is simply the product of three quantities: (1) the transmissibility of the pathogen, (2) the duration of infectiousness, and (3) the contact rate between susceptible and infectious individuals. For more complexly structured models, the calculation of R_0 can be generalized in a fairly straightforward manner (19), though the interpretation becomes more difficult.

Given the relatively few sexual contacts the typical person has, it was perplexing that STDs could remain endemic in a population. Epidemiological insights into the frequent re-infection of the same group of people with curable bacterial STDs led to the formulation of the core-group concept in STD epidemiology. (13) While the value of the basic reproduction number is below threshold for the majority of the population, it exceeds threshold for the core, where the infection can persist and occasionally infect people outside of the core. This concept is analogous to the “rescue effect” in metapopulation biology, whereby a global population remains viable through the net growth and subsequent dispersal of the population in a few “source” local populations (20).

For non-curable STDs such as HIV, the core-group concept as developed by Hethcote and Yorke (13) does not strictly apply since there is no role for re-infection. However, behavioral heterogeneity still clearly plays an important role in the transmission dynamics.

Behavioral heterogeneity has been incorporated into the formulation of R_0 by Anderson and colleagues (21). Assuming random mixing with respect to the degree distribution in a population structured by sexual activity, they show that R_0 increases linearly with the variance of the degree distribution of the population sexual contact graph.

Since surveys of sexual behavior reveal that the great majority of people have one partner or fewer in the last year, (22; 23) the driving factor for STDs is clearly the tail of the degree distribution, and this is where the emphasis for inference typically focuses (24; 25).

1.2 Social Networks

By definition, socially communicable diseases are transmitted from person to person. An intuitive and mathematically convenient means of representing social contacts is a graph (26). The nodes of the graph represent individual people and the edges represent contact. The number of edges adjacent to a particular node is its degree, and the collection of nodal degrees is the degree distribution of the population.

Graph-theoretic network models have been used to describe a wide variety of relational data including friendship networks among children, scientific collaboration networks, (27) social and

economic exchange networks, and, more recently, contact networks for the spread of infectious disease (28).

1.2.1 Models for Degree Distributions

Empirical degree distributions for sexual partnership networks are highly skewed (15). The modal yearly degree is $k = 1$ for nearly all large representative surveys (e.g., 22; 23; 29; 30; 31). The tremendous skew of sexual degree distributions has, through analogy to a variety of physical systems, suggested the possibility of power-law scaling (32). Networks exhibiting power-law scaling have been referred to as “scale-free” networks in (33) and subsequent publications. This attribution is associated with properties of the implicit underlying stochastic mechanism, and is often used loosely.

Let K be the degree of a randomly sampled person from the population. Recent empirical work (33; 25) has claimed that some sexual network degree distributions have a probability mass function (PMF) for network degree of the form, $P(K = k) \approx k^{-\rho}, k \gg 1$, where $P(K = k)$ is the probability of observing exact degree k and ρ is referred to as a scaling parameter. Let f and g be two functions with support the whole numbers. We take $f(k) \asymp g(k)$ to mean that there exist constants c_1, c_2 such that $0 \leq c_1 < f(k)/g(k) \leq c_2 < \infty$ for $k = 1, \dots$. We then say that $P(K = k)$ has *power-law behavior* if $P(K = k) \asymp k^{-\rho}$.

Inference on the scaling parameter ρ of a power-law model typically involves fitting a regression line through the apparently linear region of a plot of the survival function of the degree distribution plotted against the distribution on double logarithmic axes (33; 25). The measurement of uncertainty is then taken as the standard error of the estimated slope. This methodology is inappropriate for the inference problem, yielding (1) biased estimates of the scaling parameter, and (2) greatly underestimated model uncertainty (15).

Even for very large surveys (such as NHSLS in the USA (22)), the number of observations in the tail of the distribution is very small. The information contained in the tail of a degree distribution is therefore low. Consequently, the precision of inferences on the tail is extremely low.

The degree distribution, of course, was generated by the behavior of the individual actors. The distribution at any time point is probably best thought of as representing a dynamic equilibrium of underlying social and epidemiological processes, rather than simply a static pattern of sexual behavior (34). Specifying a plausible stochastic process by which the observed data can have been generated provides a better strategy for investigating the properties of sexual networks than considering mathematical distributions with weak proximate mechanisms. The equilibria of stochastic models of network formation can be fit to empirical data using likelihood techniques, allowing both the estimation of parameters and the assessment of goodness-of-fit to the data.

1.3 Stochastic Models for Network Formation

We will discuss three general classes of stochastic process model for sexual network formation: (1) nonhomogenous Poisson, (2) preferential attachment, and (3) “vetting” models.

Non-homogeneous Poisson models Consider the population of individuals with at least one partner in a given time period. Suppose that the number of additional partners $K - 1$ that the person has in the time period follows a Poisson distribution with expected value λ . There are many proximate mechanisms for this (for example, the partners are accumulated at a constant rate in time).

The assumption that all individuals have identical propensities to form partnerships is unrealistic. Individuals differ by gender, age, marital status, attractiveness, and other fundamental characteristics that greatly influence partnership formation. To model within-population heterogeneity, we can represent the individual expected values λ as independent draws from a distribution $P(\lambda)$.

There are myriad reasonable models for $P(\lambda)$. One flexible choice is the Gamma distribution. A Poisson distribution with Gamma-distributed rate parameter λ is a classical hierarchical model, with marginal distribution of negative binomial (35). Suppose that $P(\lambda)$ is a Gamma distribution with mean μ and standard deviation τ . Let $\lambda_i + 1$ be the expected number of partners for person i . The model can be written:

$$K = k | \lambda_i \sim \text{Poisson}(\lambda_i) + 1 \quad (1)$$

$$\lambda_i \sim \text{Gamma}(\text{mean} = \mu, \text{s.d.} = \tau) \quad (2)$$

or more formally,

$$P(K = k | \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^{k-1}}{\Gamma(k)} \quad (3)$$

$$P(\lambda_i) = \frac{e^{-\lambda_i/\beta} (\lambda_i/\beta)^{\alpha-1}}{\beta \Gamma(\alpha)} \quad \lambda_i > 0 \quad (4)$$

where $\alpha = \mu^2/\tau^2$ and $\beta = \tau^2/\mu$, so that $P(K = k) = \int_0^\infty P(K = k | \lambda) P(\lambda) d\lambda$ $k = 1, \dots$ is a negative binomial distribution shifted to $k = 1, 2, \dots$. We call the resulting partner distribution the *shifted negative binomial distribution*. The shifted negative binomial distribution can have quite a long left tail. Nonetheless, the variance is finite. The negative binomial distribution is flexible enough to model a variety of shapes of degree distributions and is widely used in ecology and epidemiology (36; 37).

Preferential attachment models In preferential attachment models, the probability that a contact is made with any particular individual is a function of that individual's current degree. Two models for preferential attachment are (1) the Yule distribution (38; 15) and (2) the Waring distribution (39).

The underlying stochastic model motivating the partnership distributions under the Yule begins with a network of r connections. Assume that (1) there is a constant probability $(\rho - 2)/(\rho - 1)$ that the $r + 1$ st partnership in the population will be initiated from a randomly chosen person to a previously sexually inactive person, and (2) otherwise the probability that the $r + 1$ st partnership will be to a person with exactly k partners is proportional to $kf(k|r)$, where $f(k|r)$ is the frequency of nodes with exactly k connections out of the r total links in

the population. Simon (38) called the limiting partnership distribution of this process the *Yule distribution*, following the pioneering work of Yule (40).

Often a measure of passing time is associated with the growth model. In some variants a partnership is associated with each time increment, and in others a person is associated with each time increment. However these formulations are equivalent in their essential representation of the network.

This stochastic process has been rediscovered, apparently without awareness, by several research groups (41; 42; 43) and been used to characterize Internet growth. The special case with $\rho = 3$ has been proposed by (41). The stochastic process underlying the Yule distribution has been used by (41; 42) to characterize Internet growth. It is also implicit in the analysis of the scaling properties of sexual contact networks (25). It has been described as the “rich get richer” model, as people in the network with many partners tend to accumulate partners faster than those with less. The resulting distribution under this model has the desired property that most people have very few partners, while a very few have many sexual partners. The PMF of the Yule distribution (35) is:

$$P(K = k) = \frac{(\rho - 1)\Gamma(k)\Gamma(\rho)}{\Gamma(k + \rho)}, \quad k = 1, 2, \dots \quad (5)$$

where $\Gamma(\rho)$ is the Gamma function of ρ . The Yule distribution has power-law behavior as $P(K = k) \asymp k^{-\rho}$.

The *Waring distribution* is a natural generalization of the Yule proposed by Irwin. (39) The motivating stochastic process is identical to that of Simon, with the exception that the probability that the $r + 1$ st partnership in the population will be initiated from a randomly chosen person to a previously sexually inactive person is $\frac{\rho-2}{\rho+\alpha-1}$. This model allows for the probability of *non-preferential* partnerships to be a separate parameter from that governing the preferential process. The limiting distribution is given by

$$P(K = k) = \frac{(\rho - 1)\Gamma(\rho + \alpha)}{\Gamma(\alpha + 1)} \cdot \frac{\Gamma(k + \alpha)}{\Gamma(k + \alpha + \rho)}, \quad \alpha > -1. \quad (6)$$

The Waring distribution has apparently been derived independently by (44) in the context of modeling growth of the Internet.

Vetting models The general idea underlying the vetting models is that people form sexual partnerships based on a two-stage process. First, they generate an acquaintance list from which they, second, choose their sexual partners. This class of model is extremely flexible in that practically any probability distribution can be specified for both of these processes. This process focuses attention on the stopping rules that people employ when forming sexual partnerships.

The Yule-vetting models are generalizations of the Yule distribution that recognize that the formation of sexual partnerships is not cost-less. Suppose that sexual partners are chosen from a pool of acquaintances. First, individuals form a random number A acquaintances from a PMF $P(A = a)$. In many situations the process of acquaintance formation may be a relatively cost-less process and $P(A = a)$ may have power-law behavior. This process may represent social networking, geographic, or other processes. Second, suppose the potential number L of

sex partners an individual has in the time period follows a distribution that is typically short-tailed. For example, it could be a geometric or negative binomial distribution. There are many proximate mechanisms for this (e.g., the partners are from a queue with constant rate). However the actual number of partners K is bounded by the number of acquaintances the person has, in the sense that it can not exceed that number: $K = \min(A, L)$. The distribution of the number of partners is then:

$$P(K = k) = \sum_{a=1}^{\infty} P(K = k|A = a)P(A = a) = P(A = k)P(L > k) + P(A \geq k)P(L = k) \quad k = 1, 2, \dots \quad (7)$$

This will resemble $P(L = k)$ (e.g., short-tailed) when L is stochastically much smaller than A . It will resemble $P(A = a)$ when L is stochastically much larger than A . These two situations correspond to relatively large and small acquaintance networks, respectively.

As before, the assumption that all individuals have identical propensities to form sex partnerships is unrealistic due to the individual characteristics such as gender, age, marital status, attractiveness. We can model this within-population heterogeneity, by independently drawing the individual expected values from a distribution $P(\lambda)$.

If $P(\lambda)$ is a Gamma distribution with mean μ and scale parameter τ , and that the acquaintance distribution is Yule with scale parameter ρ , then the potential partnership distribution is negative binomial and we call the resulting partner distribution the *negative binomial Yule distribution*. These potential partnership distributions have an alternative interpretation. Suppose an individual chooses partners from their acquaintances and each independently satisfy some criterion with a fixed probability. Then each individual chooses partners until they meet a quota of people at satisfy this criterion. An example of this would be people looking for a number of long-term partners among many casual partners. Many other interpretations are possible and interesting with context.

A simple member of this class has $\tau = 1$ so that the heterogeneity distribution is exponential. In this case the potential partnership distribution is geometric and the distribution is called the *geometric Yule distribution*. This model has a nice interpretation as people looking for a single long-term partner among many casual partners.

The negative binomial Yule has as special limiting cases the Yule and Shifted Negative binomial distributions. Furthermore, in the intermediate ranges, it can be shown that

$$P(K = k) \asymp e^{-k/\kappa} k^{-\alpha} \quad k = 1, 2, \dots \quad (8)$$

where $\kappa = 1/\log(1 + 1/\tau\mu)$. The right-hand-side of this equation represents a model that has been considered by various authors and for a variety of networks. (45; 33) Newman has argued this model applies to scientific collaboration networks (27), and has named the distribution the *truncated power law*. It is unclear if the Power-Exponential model has a stochastic mechanism that motivates its exact mathematical form. However it can be motivated as an *ad hoc* approximation to the geometric Yule distribution which does have a reasonable stochastic process underlying it.

A distribution that may prove useful for generating long-tailed acquaintance lists is the *discrete Pareto* distribution (35). Similarly, the acquaintance list distribution can be modeled

as Yule, Waring, or negative binomial.

Vetting models represent many other degree distributions as special cases. One particular case of interest lies in the intermediate parameter range of models with power-law list generators in which it can be shown that $P(K = k) \asymp e^{-k/\kappa} k^{-\alpha}$, where $\kappa = 1/\log(1 + 1/\tau\mu)$. This is the power-exponential distribution used by Newman (27) to model scientific collaboration networks and is frequently referred to as the *truncated power law* model.

Discrete Pareto Linguists have long been interested in the frequency of words in large sequences of text. They proposed an underlying stochastic mechanism called the *Zipf-Estoup law* (46; 47) analogous to that of Simon. This results in a power-law model for K the number of times a word appears in the text (35):

$$P(K = k) = \frac{1}{\zeta(\rho)k^\rho}, \quad k = 1, 2, \dots \quad (9)$$

where $\zeta(\rho)$, the Riemann zeta function of ρ . This is often referred to as the *discrete Pareto distribution* in analogy to the continuous version (which has been used as a model for degree distribution in (24)). It is also referred to as the *(Riemann) zeta distribution*. Two examples of this distribution are Zipf's law ($\rho = 1$) relating the relative frequency of words in a text to their rank, and Lotka's law ($\rho = 2$) for the number of authors making contributions.

For the same expected value, the discrete Pareto distribution will initially decay more rapidly than the Geometric, negative binomial or Yule, but then more slowly in the tail (35). The implication of this is that a negative binomial or Yule degree distribution contaminated with extreme outliers could mimic a discrete Pareto distribution.

The discrete Pareto distribution lacks a plausible stochastic mechanism for network formation, which limits its appeal as a model of sexual contact networks. While this lack of a motivating stochastic mechanism constrains the utility of the discrete Pareto distribution as a model for sexual networks, it can nonetheless be used as an acquaintance-list generator in a vetting model.

1.4 Differential Tail Behavior

It is probable that the behavior that governs the acquisition (and maintenance) of the first sexual partner a person has in a given time interval will differ from the process leading to the acquisition of further partners. For example, the process leading to the choice of marriage partner is likely to differ from the process leading to the extra-marital affairs (48).

To account for the possibility of substantial differences between the process at low and high network degree, we allowed for the inclusion of extra parameters to fit the low-degree observations (e.g., $k = 1, 2$). We evaluated the improvement of fit effected by the inclusion of such parameters in the model selection stage (see section 2.2.1). Network data collected on sexual activity of ever-active people for a relatively short interval, such as a single year, are likely to contain a fraction of people for whom $k = 0$. While it is traditional to discard these values before performing inference on the scaling behavior of the degree distribution (33; 25), the number of zeros in a population holds a tremendous amount of information about both the

distribution and the process that generated it. All models that we fit contain a parameter for the frequency of degree $k = 0$.

2 Methods

2.1 Data

There are a variety of forms of network data (49). Local network data result from collecting information on the number and attributes of a focal individual’s sexual partners ascertained using an epidemiological/sociological survey instrument. For our analysis, we use local network data gathered from men and women as a part of three large, representative surveys of sexual behavior. For all surveys, we used the reported number of sexual partners in the last year as the estimate of individual network degree.

Rakai Project Sexual Network Survey The Rakai district is an administrative unit of southern Uganda with a mature AIDS epidemic and an HIV/AIDS prevalence of approximately 16%. As with all of Sub-Saharan Africa, the primary mode of HIV transmission in Rakai is believed to be heterosexual. Data were collected as part of a large international collaborative research project.

Sex in Sweden Data from Sweden come from the 1996 “Sex in Sweden” survey based on a nationwide probability sample and financed by the Swedish National Board of Health (23).

National Health and Social Life Survey Data from the United States comes from the National Health and Social Life Survey (NHSLS) (22). NHSLS was a national probability sample of the sexual behavior of Americans. Sexual partner count was ascertained by two different techniques for the NHSLS. First, respondents were asked to indicate the number of sexual partners they had had in the last year (and in their lives) in face-to-face interviews. Second, respondents were asked the same questions on a written instrument in which partner numbers were binned into categories of $k = \{0, 1, 2 - 5, 5 - 10, 10 - 20, 20 - 100\}$.

There is considerable evidence that the latter method generally yields more reliable results (22; 23). We therefore use the responses to the written question for our model estimation. This introduces some complications to the likelihood function which are nonetheless easily handled (see below).

Neither Sweden nor the United States is characterized by a generalized HIV/AIDS epidemic with national prevalence for both countries less than 1%. Table 1 presents summary statistics for the three studies.

2.2 Statistical Inference

We adopt a likelihood framework to estimate the model parameters and compare the different models against each other. The likelihood framework provides a set of powerful tools for inference. The method of maximum likelihood estimation has been traditionally regarded as optimal based on asymptotic arguments (50). For most models the MLE is approximately normally

distributed (35) and for small sample sizes the uncertainty of the MLE can be quantified by the bootstrap (51). Here, we employ bootstrap methods to quantify the small sample properties of the MLEs and calculate confidence intervals.

Like any sample from a population, the samples obtained in the sexual surveys we analyze here are imperfect representations of their populations. Errors accrue due to sampling frame mis-specification, informant mis-report and non-response (52; 53; 54). Likelihood methods enjoy the tremendous advantage that the sampling design is “ignorable” for many standard (and non-standard) designs under the likelihood framework (55). That is, the likelihood only depends on data in the sample, and not on unknown missing data.

Consider fitting a PMF $P_\theta(K = k)$ to survey information where θ is the (possibly vector) parameter. For example, for the Yule model the parameter is ρ , the scaling exponent. Given a random sample of n individuals with reported degrees K_1, \dots, K_n the logarithm of the likelihood of the model is

$$\mathcal{L}(\theta|K_1 = k_1, \dots, K_n = k_n) \equiv \sum_{m=1}^n \log(P_\theta(K = k_m)). \quad (10)$$

A maximum likelihood estimator (MLE) for θ is a value $\hat{\theta}$ that maximizes $\mathcal{L}(\theta|K_1, \dots, K_n)$ as a function of θ .

Since our models allow for the decoupling of the behavior of the majority of observations and that of the tail, the complete data likelihood will be somewhat more complicated than a standard likelihood. (56) Define k_{min} as the degree above which the parametric model (e.g., the Yule) is fit and denote $P(K = k) = \pi_k$ for $k \leq k_{min}$. By the law of total probability, the probability of observing degree k is:

$$P(K = k) = P(K = k, k > k_{min}) + P(K = k, k \leq k_{min}) \quad (11)$$

$$= P(K = k|K > k_{min})P(K > k_{min})\mathcal{I}(k > k_{min}) + \pi_k\mathcal{I}(k \leq k_{min}), \quad (12)$$

$$= P(K = k|K > k_{min})\left(1 - \sum_{m=0}^{k_{min}} \pi_m\right)\mathcal{I}(k > k_{min}) + \pi_k\mathcal{I}(k \leq k_{min}),$$

where \mathcal{I} is the indicator function.

Given n total observations with observed frequencies n_0, n_1, \dots on network degree $k = 0, 1, \dots$, the log-likelihood of the data is

$$\begin{aligned} \mathcal{L}(\pi, \theta|K_1 = k_1, \dots, K_n = k_n) &= \sum_{m=0}^{k_{min}} n_m \log(\pi_m) + \left(n - \sum_{m=0}^{k_{min}} n_m\right) \log\left(1 - \sum_{m=0}^{k_{min}} \pi_m\right) + \\ &+ \sum_{m=k_{min}+1}^{\infty} n_m \log\left(P(K = m|K > k_{min})\right). \end{aligned} \quad (13)$$

The last term is the contribution to the likelihood of the observed values given that they are in the tail of the distribution. From the form of (14) the observed values in the tail are sufficient for the parameter θ . Similarly, the MLEs of π_m are given by the sample proportions, $\hat{\pi}_m = n_m/n, m = 0, 1, \dots, k_{min}$ (see Appendix A).

Special care was needed in calculating the likelihood for the NHSLs data, because the data on sexual partner count were grouped. We employed a mixed parametric likelihood approach to the problem (56). The full data likelihood for NHSLs is presented in the Appendix.

2.2.1 Model Selection

The different models we consider have different numbers of parameters. In general, a model with more parameters may be expected to fit data better than a model with few parameters. The likelihood \mathcal{L} provides a measure of the goodness-of-fit of a model to the data that does not adjust for the complexity of the model. If the models are nested, in the sense that they form a sequence with each model being a subset of a previous more complex one, then likelihood ratio tests can be used. Many solutions have been proposed for non-nested situations, such as we are faced with here. We adopt two different approaches: (1) the Akaike Information Criterion (AIC) (57; 58; 59) and (2) the Bayesian Information Criterion (BIC) (60). For a simple random sample of n people with data K_1, \dots, K_n , the AIC is defined as $AIC = -2\mathcal{L}(\hat{\theta}|K_1 = k_1, \dots, K_n = k_n) + 2d$, and $BIC = -2\mathcal{L}(\hat{\theta}|K_1 = k_1, \dots, K_n = k_n) + \log(n)d$.

A model is judged better than another model if it has a smaller AIC (or BIC) value. Both AIC and BIC have solid theoretical foundations: Kullback-Leibler distance in information theory (for AIC), and integrated likelihood in Bayesian theory (for BIC). The BIC approach has the benefit of incorporating model uncertainty and sample size into the decision. The AIC has the advantage of efficiency. That is, for large sample size, it is the best approximation to the “true” model (58). If the complexity of the true model does not increase with the size of the data set, the BIC is usually preferred, otherwise AIC is preferred. However, both criteria should be used for guidance and not used to unilaterally exclude models solely based on ranking.

3 Results

The results of the model fits are presented in table 2. We have listed the five best-fitting models, by AIC. Four out of the six networks were best fit by the shifted negative binomial model and one was best fit by the special case of the negative binomial, the geometric. For all networks but one, the best-fitting model had low k_{min} (i.e., 1 or 2), whereas for the Ugandan men, $k_{min} = 4$.

For all the networks but one, the top five best-fitting models had similar values of the likelihood (AIC/BIC). However, for the American men’s network, the negative binomial model fit dramatically better than the next best model.

For all networks in which a power-law model (i.e., Yule, Waring) appeared, the value of the scaling parameter $\rho > 3$, indicating that all networks were characterized by finite variance (15).

4 Discussion

A wide range of potential models for the degree distributions exist and empirical data are often limited. Frequently, many models will fit the empirical information approximately equally well, at least superficially. Thus sophisticated statistical methodology should be used to assess the quality of the fit of proposed models, and the plausibility of stochastic mechanisms underlying the

model must be weighed. M. Kendall emphasized this point in his 1960 inaugural address to the Royal Statistical Society (34). Kendall argued that for statistical modeling in the social sciences to mature as a scientific discipline, it must move beyond simple curve fitting exercises and into tests of process models. While revealing regular patterns in social systems is an important first step in scientific understanding, their existence does not entail a causal mechanism. Since the observed patterns in the social world were generated by the behavior of individual actors, specifying plausible, testable stochastic processes by which larger patterns emerge is essential in the scientific understanding of social phenomena. Unfortunately, this message has often been lost in the passage of time and the segmentation of scientific enterprise.

Our results indicate that no unitary stochastic process easily explains the formation of sexual networks in the populations we have examined. It has been suggested that sexual networks may display power-law behavior (25; 61). While some of the power-law models appear to fit the women’s networks, both the men’s and women’s networks were generally better fit by the negative binomial and variants. Preferential attachment is one stochastic mechanism for the evolution of networks characterized by power-law degree distributions (42; 62) Lloyd and May (32) suggested that such processes might be relevant for human sexual contact networks and this hypothesis found putative empirical support by (25) in the Swedish sexual networks analyzed in this paper. We found the support for preferential attachment to be mixed overall. While a preferential attachment model (i.e., Yule, Waring) fell into the top five best-fitting models for all networks but the American and Swedish men, it was the top model only in the case of American women.

Despite their appealing stochastic mechanism, the vetting models performed poorly relative to the simpler models.

The Importance of Heterogeneity The stochastic mechanism most clearly supported by the model selection procedure was individual heterogeneity in the propensity for forming partnerships. This is reflected in the negative binomial fitting best for five of the six networks, and in the magnitude of the likelihood criteria supporting this model for the American men.

One interpretation for such a model is that each individual person has a constant hazard of forming a new sexual partnership, but that this hazard varies from individual to individual. For the great majority of people, this hazard is very low.

Limitations of Exclusively Degree-Based Models Research on social networks with application to epidemiology has focused on degree distribution as the primary property of interest (28; 45; 25). This is a natural starting point, both because of the theoretical focus on heterogeneity in sexual activity (21; 63) and the relative availability of local network data. However, there are other features of sexual networks which have a bearing on epidemic processes. Two such features that have received a great deal of attention recently are clustering and minimum path length, the determinants of the so-called “small world” network effect. Small-world networks can arise from power-law models of the degree distribution (33). Small world graphs tend to have larger final epidemic sizes than other sparse networks because of their relatively high connectedness (28).

Nodal attributes, such as gender, ethnicity or marital status, are also of fundamental importance for the formation of networks (16). The probability that an actor will form a sexual

contact with an alter is a function of both the actor’s and the alter’s nodal attributes, separately from their degree (64). Differential selectivity in interaction is not easily accommodated in standard degree-based epidemic theory. The standard model of heterogeneity (14) assumes random mixing conditional on activity class. More recent formal treatments, which expressly deal with network structure, deploy similar assumptions (45).

In addition to the nodal attributes, other network properties, which are not necessarily related to degree, play a fundamental role in epidemics on networks. One property of particular importance is concurrency. Concurrency denotes the propensity to form simultaneous partnerships (65). In both simulation (65; 66) and empirical studies, (67) concurrency has been demonstrated to have significant effects on STD epidemics independent of the number of sexual partnerships.

An alternative to strictly degree-based network models that can accommodate all these features are the Exponential Random Graph Models (ERGMs), also known as p^* models (68). While the specification of such models presents a host of technical challenges, recent advancements in statistical computing have put them within reach. Specifically, Markov-Chain Monte Carlo (MCMC) simulation provides a means both for routinely calculating the likelihoods of large ERGMs and simulation networks with given nodal and structural characteristics. Recent work into the properties of MCMC solutions to ERGM models promises to overcome some of the traditional difficulties in calculating likelihood estimates (69).

Limitations Inference about the properties of the sexual network degree distribution is limited by data quality. There are three problematic features of available network data: (1) potentially sizable sampling errors associated with sexual history surveys (70), (2) data coarsening (particularly for NHSLs), (3) censoring of very high degree individuals due to AIDS mortality (King Holmes, personal communication).

While these factors all limit our ability to infer network properties, they are certainly not unique to the likelihood-based procedures we have employed. The likelihood framework, in fact, can deal with issues associated with measurement error and bias more effectively than previous methods used in this context (15). However, the intrinsic data limitations warrant caution for the interpretation of stochastic models. Any stochastic model will be a caricature of reality, and it is the modeler’s responsibility to decide to what degree of precision a useful model must conform.

A potentially greater problem for scientific progress in this field is the dearth of behavioral data on how sexual partnership networks are formed. In their critique of the common practice of using observational data to simultaneously choose a model and estimate its parameters, Burnham and Anderson (58) note that the *scientific* work in multi-model inference lies in the development of *a priori* mechanistic models. These models can be derived either empirically or deductively from theory. In developing alternative models for this paper, we have attempted to specify plausible behavioral mechanisms by which networks form, allowing us to confront existing models (e.g., 25) with alternatives. Progress on this front is clearly predicated on careful behavioral work such as that of Gorbach and colleagues on the formation of concurrent sexual partnerships (71).

Table 2 reveals substantial model uncertainty. Different models yield similar AIC values for the same data. This result suggests two important features of the inference for sexual networks.

First, it highlights the need both for models based on plausible stochastic processes and for behavioral data to allow us to distinguish models based on mechanism. Second, it indicates that model uncertainty should be accounted for when attempting to make predictions regarding the behavior of sexual networks or epidemics thereupon. This uncertainty can be dealt with through, for example, Bayesian model averaging (72).

Conclusion The results we have presented here indicate that a general behavioral model for the formation of human sexual contact networks is still lacking. Preferential attachment is one model that has been previously suggested, and the assumptions of preferential attachment have been incorporated into subsequent work. However, this mechanism does not perform especially well when confronted with alternative models. We suggest that incorporating actor heterogeneity and dependence is essential for future network models in epidemiology. Furthermore, if we are to move beyond *ad hoc* curve fits of network degree distributions and make real progress in understanding the stochastic mechanisms which generate empirical networks, two points are essential: (1) we must recognize that there is much more to sexual networks than degree distributions, and (2) collaboration between network modelers, epidemiologists, and behavioral scientists is essential.

Acknowledgements

We gratefully acknowledge the critical feedback and support we have received from Martina Morris, King Holmes, Julian Besag, Adrian Raftery, Steve Goodreau, Mark Newman, Richard Hayes, Roy Anderson, and Bob May. Virginia Rutter suggested the term “vetting model.” Frederik Liljeros generously made available the Swedish data used in (25) We especially wish to thank Dr. Bo Lewin, Professor of Sociology, Uppsala University and head of the research team responsible for the “Sex in Sweden” study for providing the data used in this study. This research supported by Grant 7R01DA012831-02 from NICHD and Grant 1R01HD041877 from NIDA.

Appendices

A Maximum likelihood estimator for π_m

Equation 14 presents the full data log-likelihood for the models we use. Here we verify that the maximum likelihood estimator $\hat{\pi}_m$ is simply the sample proportion of m , n_m/n .

The maximum likelihood estimator is given by,

$$\frac{\partial \mathcal{L}(\pi, \theta | K_1 = k_1, \dots, K_n = k_n)}{\partial \pi_k} = \frac{n_k}{\pi_k} - \frac{n - \sum_{m=0}^{k_{min}} n_m}{1 - \sum_{m=0}^{k_{min}} \pi_m},$$

which is zero, only if $\hat{\pi}_m = \frac{n_m}{n} \forall m = 0, \dots$ This partial is not defined for $\pi_k = 0$ or $\sum_{m=0}^{k_{min}} \pi_m = 1$. However in these special cases the likelihood is maximized (with probability 1) by the sample proportions.

The Hessian matrix of equation 14,

$$\left[\frac{\partial^2 \mathcal{L}(\pi, \theta | K_1 = k_1, \dots, K_n = k_n)}{\partial \pi_k \partial \pi_j} \right],$$

is negative definite indicating that the MLE is the unique global maximum.

B Likelihood for NHSLs data

The NHSLs data for annual sexual partner count were binned into categories $k = \{0, 1, 2 - 5, 6 - 10, 11 - 20, 21 - 100\}$. Suppose the observations C are a categories $\{1, \dots, C\}$. The probability of an observation in category c with inclusive range $[l_c, u_c]$ is

$$P_\theta(C = c) = P_\theta(l_c \leq K \leq u_c) = \sum_{m=l_c}^{u_c} P(K = m)$$

where $P_\theta(K = m)$ is the log-likelihood. The exact observation of the degree is the special case where the categories correspond to a single degree.

Suppose we observe categories $C_1 = c_1, \dots, C_n = c_n$. The full data likelihood for the grouped NHSLs data is then

$$\mathcal{L}(\theta | C_1 = c_1, \dots, C_n = c_n) = \sum_{m=1}^n \log(P_\theta(C = c_m)).$$

References

- [1] National Institute of Medicine (1997) *The Hidden Epidemic: Confronting Sexually Transmitted Diseases* (National Academy Press, Washington, DC).
- [2] Centers for Disease Control and Prevention (2001) Sexually transmitted disease surveillance 2000 supplement: Chlamydia prevalence monitoring project annual report, Technical report, Department of Health and Human Services, Centers for Disease Control and Prevention.
- [3] Williams, L. A., Klausner, J. D., Whittington, W. L., Handsfield, H. H., Celum, C., and Holmes, K. K. (1999) *American Journal of Public Health* **89**, 1093–7.
- [4] Holmes, K. K., Sparling, P., Mårdh, P.-A., Lemon, S., Stamm, W., Piot, P., and Wasserheit, J., eds. (1999) *Sexually Transmitted Diseases* (McGraw-Hill, New York), 3rd edition.
- [5] Gershengorn, H. B. and Blower, S. M. (2000) *AIDS Patient Care and STDS* **14**, 133–142.
- [6] Knapp, J. S., Fox, K. K., Trees, D. L., and Whittington, W. L. (1997) *Emerging Infectious Disease* **3**, 33–39.

- [7] Centers for Disease Control and Prevention (2000) *Journal of the American Medical Association* **284**, 1917–1919.
- [8] Rahman, M., Sultan, Z., Monira, S., Alam, A., Nessa, K., Islam, S., Nahar, S., Shama, A. W., Alam Khan, S., Bogaerts, J., Islam, N., and Albert, J. (2002) *Journal of Clinical Microbiology* **40**, 2037–40.
- [9] Aplasca De Los Reyes, M. R., Pato-Mesola, V., Klausner, J. D., Manalastas, R., Wi, T., Tuazon, C. U., Dallabetta, G., Whittington, W. L., and Holmes, K. K. (2001) *Clinical Infectious Disease* **32**, 1313–1318.
- [10] Levin, B. R. and Bergstrom, C. T. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6981–6985.
- [11] Anderson, R. M. and Garnett, G. P. (2000) *Sexually Transmitted Diseases* **27**, 636–643.
- [12] Foulkes, M. A. (1998) *Statistics in Medicine* **17**, 1–25.
- [13] Hethcote, H. W. and Yorke, J. A. (1984) *Lecture Notes in Biomathematics* **56**, 1–105.
- [14] Anderson, R. M. and May, R. M. (1991) *Infectious diseases of humans: Dynamics and control* (Oxford University Press, Oxford).
- [15] Jones, J. and Handcock, M. (2002) Epidemic thresholds exist in human sexual contact networks, Working Paper 23, Center for Statistics and the Social Sciences.
- [16] Morris, M. (1991) *Mathematical Biosciences* **107**, 349–377.
- [17] Morris, M. (1995) in *Epidemic Models: their structure and relation to data*, D. Mollison, ed. (Cambridge University Press, Cambridge), pp. 302–322.
- [18] Bailey, N. (1975) *The mathematical theory of infectious disease* (Hafner Press, New York).
- [19] Diekmann, O., Heesterbeek, J. A. P., and Metz, J. A. J. (1990) *Journal of Mathematical Biology* **28**, 365–382.
- [20] Kawecki, T. J. and Holt, R. D. (2002) *American Naturalist* **160**, 333–347.
- [21] Anderson, R., Medley, G., May, R. M., and Johnson, A. (1986) *IMA Journal of Mathematics Applied in Medicine and Biology* **3**, 229–263.
- [22] Laumann, E., Gagnon, J., Michael, T., and Michaels, S. (1994) *The social organization of sexuality: Sexual practices in the United States* (University of Chicago Press, Chicago).
- [23] Lewin, B., ed. (1996) *Sex in Sweden* (National Institute of Public Health, Stockholm).
- [24] May, R. M. and Lloyd, A. L. (2001) *Physical Review E* **64**, 066112.
- [25] Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E., and Åberg, Y. (2001) *Nature* **411**, 907–908.

- [26] Wasserman, S. and Faust, K. (1994) *Social network analysis: Methods and applications*, Structural analysis in the social sciences (Cambridge University Press, Cambridge).
- [27] Newman, M. E. J. (2001) *Proceedings of the National Academy of Sciences of the United States of America* **98**, 404–409.
- [28] Newman, M. E. J. (2002) *Physical Review E* **66**, art. no.–016128.
- [29] Hubert, M., Bajos, N., and Sandfort, T. (1998) *Sexual Behaviour and HIV/AIDS in Europe: Comparisons of National Surveys* (UCL Press, London).
- [30] Aral, S. O. (1999) *Sexually Transmitted Diseases* **26**, 262–264.
- [31] Youm, Y. and Laumann, E. (1999) *Sexually Transmitted Diseases* **29**, 689–697.
- [32] Lloyd, A. L. and May, R. M. (2001) *Science* **292**, 1316–1317.
- [33] Amaral, L. A. N., Scala, A., Barthélemy, M., and Stanley, H. E. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**, 11149–11152.
- [34] Kendall, M. (1961) *Journal of the Royal Statistical Society Series A-Statistics in Society* **124**, 1–16.
- [35] Johnson, N., Kotz, S., and Kemp, A. (1992) *Univariate discrete distributions*, Wiley series in probability and mathematical statistics (Wiley, New York), 2nd edition.
- [36] Martin, D. and Katti, S. (1965) *Biometrics* **21**, 34–48.
- [37] Alexander, N., Moyeed, R., and Stander, J. (2000) *Biostatistics* **1**, 453–463.
- [38] Simon, H. (1955) *Biometrika* **42**, 435–440.
- [39] Irwin, J. (1963) *Journal of the Royal Statistical Society Series A (General)* **126**, 1–45.
- [40] Yule, G. (1924) *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **213**, 21.
- [41] Barabási, A. L. and Albert, R. (1999) *Science* **286**, 509–512.
- [42] Albert, R. and Barabási, A. L. (2000) *Physical Review Letters* **85**, 5234–5237.
- [43] Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N. (2000) *Physical Review Letters* **85**, 4633–4636.
- [44] Levene, M., Fenner, T., Loizou, G., and Wheeldon, R. (2002) *Computer Networks* **39**, 277–287.
- [45] Newman, M. (2002) Random graphs as models of networks, Working Paper 02-02-005, Santa Fe Institute.
- [46] Estroup, J. (1916) *Les gammes sténographiques* (Institut Sténographique, Paris).

- [47] Zipf, G. (1949) *Human behavior and the principle of least effort* (Addison-Wesley, New York).
- [48] Drake, C. R. and McCabe, M. P. (2000) *Journal of Applied Social Psychology* **30**, 1421–1439.
- [49] Morris, M. (1997) *AIDS* **11**, S209–S216.
- [50] Casella, G. and Berger, R. (2002) *Statistical Inference* (Duxbury, Pacific Grove), 2nd edition.
- [51] Efron, B. and Tibshirani, R. J. (1993) *An introduction to the bootstrap* (Chapman and Hall, New York).
- [52] Rubin, D. (1987) *Multiple imputation for nonresponse in surveys* (Wiley, New York).
- [53] Morris, M. (1993) *Nature* **365**, 437–440.
- [54] Thompson, S. K. (1992) *Sampling* (Wiley, New York).
- [55] Thompson, S. K. and Seber, G. A. (1996) *Adaptive sampling* (Wiley, New York).
- [56] Groeneboom, P. and Wellner, J. (1992) *Information bounds and nonparametric maximum likelihood estimation* (Birkhäuser, Basel).
- [57] Akaike, H. (1974) *IEEE Transactions on Automatic Control* **AU-19**, 716–722.
- [58] Burnham, K. and Anderson, D. (2002) *Model selection and inference: A practical information-theoretic approach* (Springer, New York), 2nd edition.
- [59] Parzen, E., Tanabe, K., and Kitagawa, G., eds. (1998) *Selected papers of Hirotugu Akaike* (Springer, New York).
- [60] Raftery, A. E. (1995) *Sociological Methodology* **25**, 111–163.
- [61] Dezsó, Z. and Barabási, A. L. (2002) *Physical Review E* **65**, art. no. 055103.
- [62] Pastor-Satorras, R. and Vespignani, A. (2001) *Physical Review Letters* **86**, 3200–3203.
- [63] Gupta, S., Anderson, R., and May, R. (1989) *AIDS* **3**, 807–817.
- [64] Wasserman, S. and Pattison, P. (1996) *Psychometrika* **61**, 401–425.
- [65] Morris, M. and Kretzschmar, M. (1995) *Social Networks* **17**, 299–318.
- [66] Morris, M. and Kretzschmar, M. (1997) *AIDS* **11**, 641–648.
- [67] Potterat, J. J., Zimmerman-Rogers, H., Muth, S. Q., Rothenberg, R. B., Green, D. L., Taylor, J. E., Bonney, M. S., and White, H. A. (1999) *American Journal of Epidemiology* **150**, 1331–1339.

- [68] Frank, O. and Strauss, D. (1986) *Journal of the American Statistical Association* **81**, 832–842.
- [69] Handcock, M. (2002) Assessing degeneracy in statistical models of social networks, Working paper, Center for Statistics and the Social Sciences.
- [70] Brewer, D. D., Potterat, J. J., Garrett, S. B., Muth, S. Q., Roberts, J. M., Kasprzyk, D., Montano, D. E., and Darrow, W. W. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**, 12385–12388.
- [71] Gorbach, P. M., Stoner, B. P., Aral, S. O., Whittington, W., and Holmes, K. K. (2002) *Sexually Transmitted Diseases* **29**, 453–62.
- [72] Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (2000) *Statistical Science* **15**, 193–195.
- [73] Sewankambo, N. K., Gray, R. H., Ahmad, S., Serwadda, D., Wabwire-Mangen, F., Nalugoda, F., Kiwanuka, N., Lutalo, T., Kigozi, G., Li, C. J., Meehan, M. P., Brahmbatt, H., and Wawer, M. J. (2000) *AIDS* **14**, 2391–2400.

Survey	sex	n	mean	variance	HIV prevalence
Rakai	women	803	0.89	0.27	16%
	men	621	1.28	1.23	
Sweden	women	1335	1.01	0.88	0.08%
	men	1476	1.27	2.19	
USA	women	1919	1.09	0.69	0.6%
	men	1506	1.41	1.42	

Table 1: Descriptive statistics for the three sexual history surveys. Means and variances are for the number of reported sexual partners in the last year. HIV Prevalence data are for the year 2000 and come from (73) for Rakai and UNAIDS country fact sheets for Sweden and the United States.

population	sex	model	k_{min}	d	\mathcal{L}	AIC	BIC	ρ /mean	scale/s.d.
Uganda	F	nb	2	4	-525.17	1058.35	1077.10	0.27	1.90
		yule	3	4	-525.32	1058.64	1077.39	3.68	NA
		geo	3	4	-525.45	1058.91	1077.66	2.67	2.67
		nb	3	5	-524.89	1059.79	1083.23	1.88	0.11
		tpl	3	5	-525.28	1060.57	1084.01	2.59	14.84
	M	geo	4	5	-781.10	1573.10	1595.26	2.72	2.72
		zero-nb	4	6	-780.90	1573.80	1600.38	6.29	0.93
		tpl	4	6	-781.02	1574.03	1600.62	4.54	0.88
		nb	4	6	-781.22	1574.44	1601.03	3.58	0.52
		yule	2	3	-784.25	1574.50	1587.79	5.43	NA
Sweden	F	nb	1	3	-1068.45	2142.90	2158.27	0.38	3.62
		yule	2	3	-1068.64	2143.27	2158.64	4.23	NA
		waring	2	4	-1067.95	2143.91	2164.39	6.53	1.87
		g-y	2	4	1068.16	2144.31	2164.80	2.89	4.62
		yule	3	4	-1068.17	2144.33	2164.82	4.68	NA
	M	nb	1	3	-1509.17	3024.34	3040.07	0.66	2.47
		g-y	2	4	-1508.21	3024.43	3045.39	2.47	7.35
		tpl	2	4	-1508.22	3024.44	3045.41	1.82	6.75
		nb-y	1	4	-1508.27	3024.55	3045.51	1.51	1.23
		nb-y0	1	4	-1508.37	3024.55	3045.51	1.51	1.23
USA	F	yule	2	3	-1600.74	3207.48	3224.03	3.84	NA
		waring	1	3	-1601.34	3208.67	3225.23	3.11	-0.68
		tpl	4	6	-1598.64	3209.27	3242.38	33.07	0.14
		pois	4	5	-1599.69	3209.39	3236.98	0.69	NA
		yule	3	4	-1600.71	3209.42	3231.49	3.91	NA
	M	nb	1	3	-1599.14	3204.28	3220.08	0.78	0.26
		nb	2	4	-1605.31	3218.63	3239.70	1.43	1.75
		geo	2	3	-1608.33	3222.66	3238.46	1.39	1.39
		nb	3	5	-1606.58	3223.15	3249.49	1.65	2.09
		nb	4	6	-1607.11	3226.21	3257.81	1.95	2.67

Table 2: Top five best fitting models for each network ordered by the value of the AIC. k_{min} is the cutoff, d is the number of parameters of the model, $\log \mathcal{L}$ is the log-likelihood of the model, AIC is Akaike’s Information Criterion, BIC is the Bayesian Information Criterion, ρ is the scaling parameter for power distributions or the mean of the non-power distributions (e.g., negative binomial, geometric), scale is the second parameter of the distribution if applicable. Models: nb = shifted negative binomial, geo = geometric, tpl = truncated power law, zero-nb = zero negative binomial, g-y = geometric-yule, pois = Poisson, nb-y = negative binomial yule.

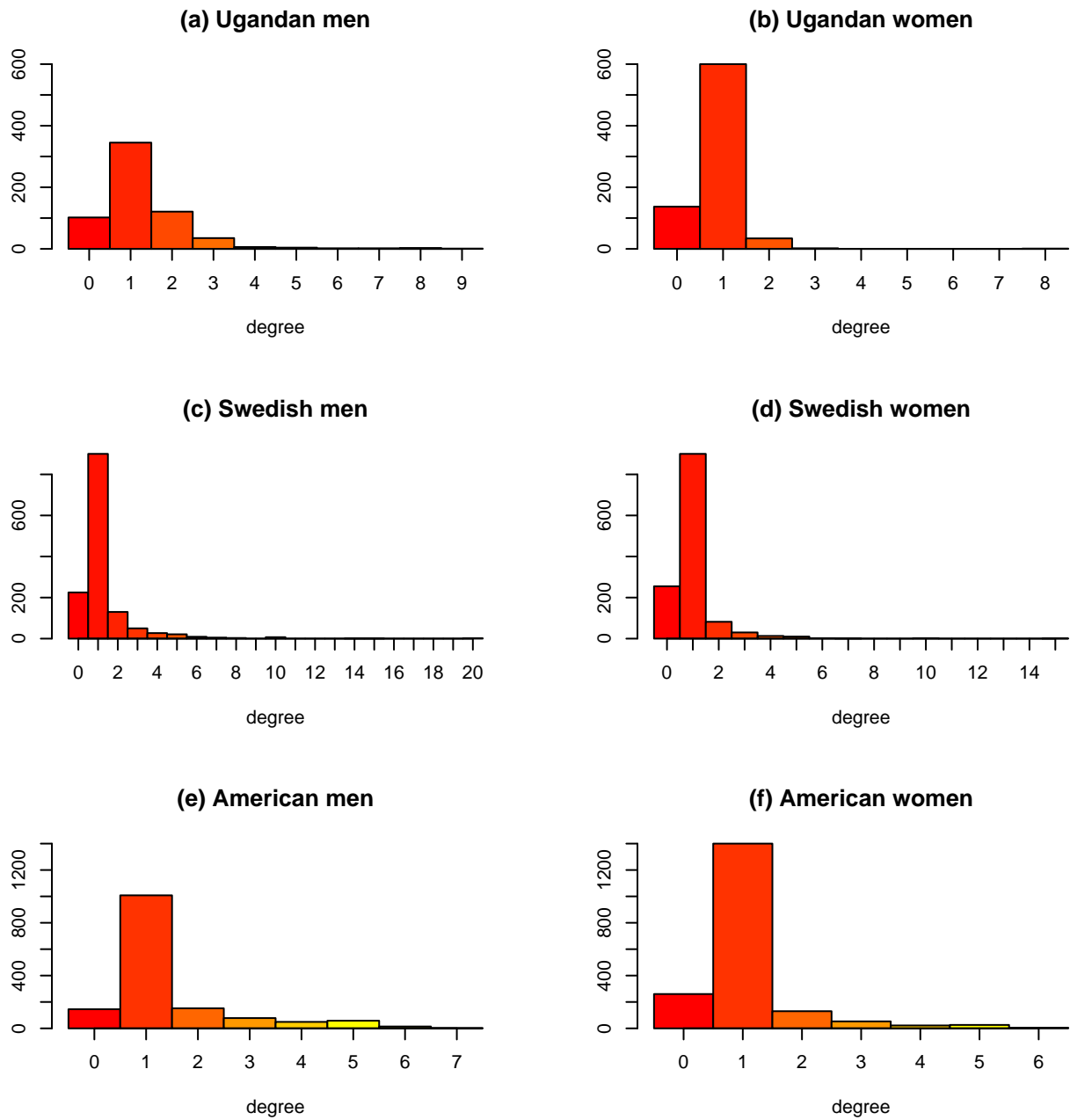


Figure 1: Reported degree distributions for men and women from the three population samples. The plots are histograms showing the absolute number of observed degree k (including zeros). By row, top to bottom, the figures are: Uganda, Sweden, USA.