BAYESIAN MODELING OF INFECTIOUS DISEASES

Raymond G. Hoffmann, Purushottam W. Laud, Ana Johnson-Masotti, Mathew Hayat
Raymond G. Hoffmann, The Medical College of Wisconsin, PO Box 26509, Milwaukee, WI 53226.
hoffmann@mcw.edu

Keywords: MCMC ; AIDS ; Nonlinear Modeling

ABSTRACT

Many models of infectious disease are based on systems of differential equations for the specification of the transmission probabilities. Estimation of the parameters is difficult since many of the parameters are only partially specified or poorly specified in the available literature. Prior distributions will be used to specify the partially known parameters together with the available data. Examples will include HIV transmission.

INTRODUCTION

Infectious disease modeling requires a stochastic model for the transmission of the disease. Usually, the model requires information about parameters that are, at best, only poorly known. For some diseases, like HIV, the time of infection is substantially earlier than the time of developing clinical symptoms. Consequently, designing an intervention to reduce the probability of transmission of the disease requires a model for the dynamics of the model. Only some parts of the dynamic model may be observable during the course of the study because of the time delay of infection to disease onset.

Our aim has been to determine the cost effectiveness of different HIV prevention strategies (Johnson-Masotti, et al, 2001). Since we can’t wait for symptoms to occur, we model the risk of transmission and the behavior of the subjects in the study in order to identify changes in the risk of transmission based on changes in behavior.

EXAMPLE: SEXUAL TRANSMISSION OF HIV

The Bernoulli model for HIV transmission risk (Pinkerton and Abramson, 1998) is based on observable (self-report or diary) sexual behavior. (McAuliffe and DeFrancisco, 2001) It is based on the following data:

1. the number of partners, and
2. the number of protected and unprotected sexual acts
   - receptive vaginal intercourse
   - insertive vaginal intercourse
   - receptive anal intercourse
   - insertive anal intercourse
   - oral sex
   - mutual masturbation.

By using a Bayesian model we can incorporate imprecise information on
- HIV prevalence
- condom effectiveness
- probability of transmission for each type of sexual activity

Gender affects the possibility of certain types of acts, but not the risk associated with them. Oral sex is of such low risk that it is usually not included in models of transmission.

The details for the model of the risk of infection are:

\[
\text{Pr(infection for person } i \mid \text{behavior } j) = 1 - \prod_{\text{partner } j} (1 - p_{ij})
\]

where

\[
p_{ij} = \pi (1 - (1 - \alpha_k))^n_k (1 - (1 - \varepsilon)\alpha_k)^c_k
\]

and

- \(\pi\) = prevalence
- \(\alpha_k\) = risk of behavior \(k\)
- \(1 - \varepsilon\) = condom efficacy
- \(n_k\) = Number of unprotected acts of type \(k\)
- \(c_k\) = Number of acts of type \(k\) with a condom
\( \pi, \varepsilon, \) and the \( \alpha_k \) are based on point estimates and confidence intervals from the literature on the prevalence, condom efficacy (about 95\%) and the risk of transmission for each type and occurrence of each act.

They are described with an informative prior, for example, condom efficacy uses a beta distribution, 

\[
1 - \varepsilon \sim \text{Beta}(\alpha, \beta) \text{ with mean and variance based on the literature (mean } = 0.95) \]

The probabilities of infection, \( \alpha_k \), are described with a logit link function and a normal error model to match the confidence intervals obtained in the literature

\[
\text{logit}(\alpha_k) \sim \text{Normal}(\mu_k, \sigma_k^2) \text{ with estimates based on the literature}
\]

- 0.0006 for vaginal insertive
- 0.0014 for vaginal receptive
- 0.0014 for anal insertive
- 0.0200 for anal receptive

Prevalence of HIV is modeled the same:

\[
\text{logit}(\pi) \sim \text{Normal}(\nu, \sigma_\pi^2) \]

based on the estimates for the Milwaukee area (mean 0.025\%)

On the other hand the observed data, number of partners and numbers of acts of sexual intercourse of each type, are described with a Poisson prior and a non-informative gamma hyperprior on the Poisson parameter for each type of act.

To estimating the posterior, we use a full conditional approach with informative priors based on the literature and non-informative (or weakly informative) priors for the behavioral data.

We have used WINBUGS for the MCMC simulations as well as SAS/IML because of the simplicity of the data management and the simplicity of distributions that we are using.

We are currently using this model to evaluate the effect of differences in the way that data is collected on the estimates of the posterior risk. For example, many HIV intervention trials aggregate sexual behavior across all the partners rather than for each partner as it was in the above model. We are also using this model to determine cost-effectiveness of different behavioral interventions.
EXAMPLE: INJECTION DRUG USE TRANSMISSION OF HIV – BEHAVIORAL MODEL

The Bernoulli model can also be used to formulate a similar type of model for HIV transmission by needle sharing (Pinkerton 1993, Lurie and Reingold, 1993). Since there is only one risky behavior, the model simplifies to

\[ p_{ij} = \omega_i \omega_S \left( 1 - \left( 1 - \pi \right) + \pi \left( 1 - \alpha \left( 1 - \omega_B \epsilon \right) \right)^n \right)^m \]

Where
- \( \omega_i = \text{pr(injecting drugs)} \)
- \( \omega_S = \text{pr(share | inject drugs)} \)
- \( \pi = \text{prevalence of HIV in IDU partners} \)
- \( \alpha = \text{pr(HIV infection | HIV+ needle)} \)
- \( \omega_B = \text{pr(bleached needle used)} \)
- \( \epsilon = \text{efficacy of bleaching the needle} \)
- \( n = \text{number of injections/partner} \)
- \( m = \text{number of partners} \)

As in the previous example, we use logit link functions with normal priors to match the literature values. For example, for the efficacy of bleach to disinfect the needles, we use a beta distribution

\[ 1 - \epsilon \sim \text{Beta}(\alpha, \beta) \text{ with mean and variance based on the literature (mean } = 0.80) \]

for prevalence, we use

\[ \text{logit}(\pi) \sim \text{Normal}(\nu, \sigma_k^2) \]

estimated mean is 0.075

and for the other probabilities we use the logit/normal model with means

\[ \omega_i = 0.080 \]
\[ \omega_S = 0.117 \]
\[ \alpha = 0.0030 \]
\[ \omega_B = 0.584 \]

The data consists of

- \( n = \text{number of injections/partner} \)
- \( m = \text{number of partners} \)

Each of these is modeled with a Poisson distribution with a gamma hyper-parameter on the Poisson mean.

EXAMPLE: INJECTION DRUG USE TRANSMISSION OF HIV – NEEDLE CIRCULATION MODEL

An alternative model for the transmission of HIV by sharing of infected needles has been developed by Kaplan, 1994. This model uses two stochastic models:

- one for the “two” state Markov process for the transition of a needle from
  - infected to uninfected – based on injection frequency, prevalence of HIV, the use of bleach to clean needles and the dilution effect of “uninfected blood on the needle”
  - uninfected to infected – based on injection frequency, prevalence of HIV and the conditional probability that a needle becomes infected if it is used by an HIV positive person.
  - (there is also an absorbing state of the needle no longer being used)
- a second stochastic model for the time that a needle is in circulation.

The first piece of the model is based on a differential equation for the probability that a needle transitions between the states (HIV+/HIV-):

\[ \frac{d\pi(t)}{dt} = \lambda - (\lambda + \mu) \bullet \pi(t) \]

where

- \( \pi(t) \) is the probability that the needle is infected at time \( t \)
- \( \lambda \) is the rate of transition from uninfected to infected and
- \( \mu \) is the rate of transition from infected to uninfected

and for the other probabilities we use the logit/normal model with means

\[ \omega_i = 0.080 \]
\[ \omega_S = 0.117 \]
\[ \alpha = 0.0030 \]
\[ \omega_B = 0.584 \]
This leads to an equation for the mean probability of a needle being infected in terms of the time from its first use.

\[
\pi(t) = \frac{\lambda}{\lambda + \mu} (1 - e^{-(\lambda + \mu)t}) + \pi_0 e^{-(\lambda + \mu)t}
\]

The second part of the model is the needle survival process which is described by an exponential regression survival distribution

\[
f(t_i, z_i) = e^{\beta z} \exp(-e^{\beta z} t_i)
\]

The regression model represents the effect of the introduction of clean needles at an exchange rate of \(z\). The mean survival with no needle exchange has a parameter (mean) \(\tau = \exp(-\beta_0)\) which has a gamma prior. Assuming that the removal rate is linear in terms of the exchange rate \(v\) needles per client per unit time gives

\[
\beta z = \beta_0 + \beta_1 v
\]

The term \(\beta_0\) represents the non-zero removal rate that obtains even when there is no needle exchange program, namely \(1/\tau\).

The likelihood that a random needle is HIV+ is based on the posterior mean of the joint density of the two processes

\[
\pi(t) \cdot e^{\beta z} \exp(-e^{\beta z} t_i)
\]

\(\lambda\) and \(\mu\) in \(\pi(t)\) and \(\beta_0\) and \(\beta\) in the survival time distribution are also exponential with values estimated from the literature and the observed data using capture-recapture methodology applied to the distribution and return of marked needles. The inverse gamma distributions with \(\phi\), \(\delta\), \(\eta\) and \(\phi\) as the hyper-parameters are will also be used as hyperpriors.

**REFERENCES**


Holtgrave DR, Pinkerton SD, Jones TS, Lurie P, Vlahov D. Cost and cost-effectiveness of increasing access to sterile syringes and needles as an HIV prevention intervention in the United States. *Journal of Acquired Immune Deficiency Syndromes. 18 Suppl 1:S133-8, 1998*

Pinkerton SD, Abramson PR. Not all behavior change is equivalent. [letter; comment]. *American Journal of Public Health. 88(4):684, 1998*


Activity Number: 132 Type: Contributed
Date/Time: Monday, August 6, 2001 : 2:00 PM to 3:50 PM Sponsor: Section on Statistics in Epidemiology*
Abstract - #301590 Title: Bayesian Modeling of Infectious Diseases
Author(s): Raymond Hoffmann*+
Affiliation(s): Medical College of Wisconsin
Address: 8701 Watertown Plank Rd., Milwaukee, Wisconsin, 53226, Keywords: MCMC ; AIDS ; Nonlinear Modeling
Abstract: Many models of infectious disease are based on systems of differential equations for the specification of the transmission probabilities. Estimation of the parameters is difficult since many of the parameters are only partially specified or poorly specified in the available literature. Prior distributions will be used to specify the partially known parameters together with the available data. Examples will include HIV transmission.