ISSN Print: 2394-7500
ISSN Online: 2 394-5869
Impact Factor: 5.2
IJAR 2016; 2(8): 36-41
www.allresearchjournal.com
Received: 08-06-2016
Accepted: 09-07-2016

Kannan R
Department of statistics,
Annamalai University,
Annamalai Nagar,
Tamil Nadu, India

Karthi R
Department of statistics,
Annamalai University,
Annamalai Nagar,
Tamil Nadu, India

Correspondence
Kannan R
Department of statistics,
Annamalai University,
Annamalai Nagar,
Tamil Nadu, India

A stochastic approach to determine the statistical measure for time to seroconversion of HIV infected using smallest order statistics

Kannan R and Karthi R

Abstract
The study of HIV infection, transmission and spread of AIDS is quite common by the use of stochastic model. The estimation of statistical measure for time to seroconversion of HIV infected over the time interval (0,t] is an important aspect which help medical intervention. We propose a stochastic model under the assumption that, the interval time between the contacts forms an order statistics and threshold follows exponential- geometric distribution. In developing such a stochastic model, the concept of shock model and cumulative damage process are used. Numerical illustration provided for different combination of parameter involved in the distribution of the random variable used in this model.

Keywords: Human immunodeficiency virus (HIV), acquired immunodeficiency syndrome (AIDS), smallest order statistics, antigenic diversity threshold, seroconversion

Introduction
Human Immunodeficiency Virus (HIV) causes the viral infection that leads to Acquired Immunodeficiency Syndrome (AIDS). AIDS is a multifaced problem having medical, social, legal, ethical and cultural dimension and only comprehensive solution covering all these aspects can succeed. The statistical modeling on AIDS epidemic has played a vital role in understanding. The application of mathematical sciences has been found be very useful in the study of HIV transmission. In the study of HIV infection, the estimation of likely time at which the seroconversion takes place is an important aspect. In the estimation of statistical measure to seroconversion, a stochastic model is conceptualized, the taking into consideration the relevant random variable which contribute to seroconversion. The successive contact of an individual with infected partner result in acquiring more and more of antigenic diversity which otherwise can be called antigenic diversity threshold which means the antigenic diversity exceed the particular level then it collapse the immune system and leads to seroconversion. For a detailed study of antigenic diversity threshold and its estimation, one can refer to Nowak and May (1991) [6] and Stilliankis *et al.*, (1994)[8].
The time intervals between successive contacts have a major role to play in the determination of statistical measures for seroconversion of HIV transmission. Ratchagar *et al.*, (2003)[7] has derive a model for the estimation of statistical measure for seroconversion of HIV infected using order statistics. Kannan *et al.*, (2008,2010,2013 and 2015)[2, 3, 4,5] have obtain a stochastic model for estimation of statistical for seroconversion of HIV infected using order statistics and threshold follows Gamma, Erlang-2, Exponentiated exponential and exponentiated modified Weibull distribution. In developing such a stochastic model the concept of shock model and cumulative damage process are developed by Esary *et al.*, (1973)[1] is used. In this paper, it is assume that, the threshold follows Exponential- geometric distribution and time interval between successive contacts forms an order statistics. The statistical measure for seroconversion are obtained taking time interval between successive contacts distributed as smallest order statistics. This is due to the fact that, if the smallest order statistics is taken implies that, inter - arrival time are becoming the smaller. Hence frequent contact would be possible which interval will have its impact on the time to seroconversion. Numerical illustrations are provided using simulated data.

Assumptions of the model

1. The transmission of HIV is only through sexual contacts.
2. An uninfected individual has sexual contacts with HIV infected partner and a random number of HIV is getting transmitted, at each contact.
3. An individual is exposed to a damage process acting on the immune system and the damage is assumed to be linear and cumulative.
4. The inter–arrival times between successive contacts are taken to be identically and independently distributed random variables.
5. The sequence of successive contacts and threshold level are independent.
6. From the collection of large number of inter-arrival times between successive contacts of a person, a random sample of 'k' observations are taken.

Notations

X_i	A random variable denoting the increase in the antigenic diversity arising due to the HIV transmitted during the i^{th} contact $X_1, X_2,\ldots, X_k$ are continuous i.i.d. random variables, with p.d.f. g(.) and c.d.f. G(.).
Y	A random variable representing antigenic diversity threshold and follows exponential-geometric distribution with parameters β and p, the p.d.f. being h(.) and c.d.f. H(.).
$U_{(1)}$	A continuous random variable denoting the inter-arrival times between the contacts follows smallest order statistics with p.d.f. $f_{u_{(1)}}(t)$ and c.d.f. $F_{u_{(1)}}(t)$
$g_k(.)$	The p.d.f of the random variable $\sum_{i=1}^{k} X_i$.
$F_k(.)$	The k^{th} convolution of F (.).
T	A continuous random variable denoting the time to seroconversion with p.d.f. *l*(.) and c.d.f. L(.).
$V_k(t)$	Probability of exactly k contacts in (0, t].
*l**(s)	The Laplace Stieltjes transform of *l*(t).
f *(s)	The Laplace Stieltjes transform of *f* (t).

Results

The survival function S(t)is

$$S(t) = P(T > t)$$

$$= \sum_{k=0}^{\infty} Pr\{there\ are\ exactly\ k\ contacts\ in\ (0,t]\} \times Pr\{the\ cumulative\ total\ of\ antigenic\ diversity < Y\}$$

$$S(t) = \sum_{k=0}^{\infty} V_k(t)\, P\left[\sum_{i=1}^{k} X_i < y\right] \quad \ldots (1)$$

It can be shown that,

$$P\left[\sum_{i=1}^{k} X_i < Y\right] = \int_0^{\infty} g_k(x)\bar{H}(x)dx$$

Where, $\bar{H}(x) = 1 - H(x)$

The probability density function of exponential-geometric distribution is,

$$h(y) = \beta(1-p)e^{-\beta y}(1-pe^{-\beta y})^2$$

and its distribution function is

$$H(y) = \left(1 - e^{-\beta y}\right)(1-pe^{-\beta y})^{-1}$$

And

$$\bar{H}(y) = \frac{e^{-\beta y}(1-p)}{(1-pe^{-\beta y})}$$

Since Y is taken to be exponential-geometric distribution (β, p).

Hence

$$P\left[\sum_{i=1}^{k} X_i < y\right] = \theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right] \quad \ldots (2)$$

Substitute equation (2) in equation (1), we get,

$$S(t) = \sum_{k=0}^{\infty}[F_k(t) - F_{k+1}(t)]\left[\theta(-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right]$$

L(t) = 1 – S(t)

$$= 1 - \left\{\sum_{k=0}^{\infty}[F_k(t) - F_{k+1}(t)]\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right]\right\}$$

$$l(t) = -\sum_{k=0}^{\infty}[f_k^*(t) - f_{k+1}^*(t)]\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right]$$

Taking Laplace transform of l(t) is,

$$l^*(s) = -\sum_{k=0}^{\infty}[f_k^*(s) - f_{k+1}^*(s)]\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right] \quad \ldots (3)$$

Where $f^*(s)$ is the Laplace transform off(.).
The inter–arrival times $U_1, U_2, U_3,\ldots, U_k$ are i.i.d random variables. Now arranging $U_1, U_2, U_3,\ldots,U_k$ in the increasing order of magnitude we get,

$U_{(1)} \leq U_{(2)} \leq \ldots \leq U_{(k)}$.

Now, we Consider the first order statistics, $U_{(1)}$
The p.d.f of the smallest order statistics is

$$f_{u_{(1)}}^*(t) = K[1 - F(t)]^{k-1} f(t)$$

The Laplace Stieltjes transform the same is given by

$$f_{u_{(1)}}^*(t) = \int_0^{\infty} e^{-st}\, K[1 - F(t)]^{k-1} f(t)dt$$

Assuming that *f*(t) follows exp(c), it can be shown that

$$f_{u_{(1)}}^*(s) = \frac{kc}{kc + s} \quad \ldots (4)$$

Consider equation (3),

$$l^*(s) = -\sum_{k=0}^{\infty}[[f^*(s)]^k - [f^*(s)]^{k+1}]\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right]$$

$$l^*(s) = -\{[f^*(s)]^k[1 - f^*(s)]\}\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right] \quad \ldots (5)$$

Substitude eqn (4) in eqn (5), we get,

$$l^*(s) = -\left\{\left[\frac{kc}{kc + s}\right]^k\left[1 - \frac{kc}{kc + s}\right]\right\}\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right]$$

$$l^*(s) = -\left[\frac{(kc)^k s}{(kc + s)^{k+1}}\right]\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right]$$

$$E(T) = -\left.\frac{dl^*(s)}{ds}\right|_{s=0}$$

$$= \left[\frac{(kc + s)^{k+1}(kc)^{\mathrm{k}} - (kc)^{\mathrm{k}}\,\mathrm{s}\,(\mathrm{k}+1)(kc + s)^k}{(kc + s)^{2k+2}}\right] \times \left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right]$$

Put s=0,

$$= \left[\frac{(kc)^{k+1}(kc)^{\mathrm{k}}}{(kc)^{2k+2}}\right] \times \left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta + i\beta}\right]\right]$$

$$E(T) = -\frac{dl^*(s)}{ds}\Big|_{s=0} = \frac{1}{kc}\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta+i\beta}\right]\right] \quad \ldots (6)$$

(On simplification)

$$E(T^2) = \frac{d^2l^*(s)}{ds^2}\Big|_{s=0}$$

$$\frac{d^2l^*(s)}{ds^2} = \left[\frac{[(kc+s)^{k+2}(kc)^{\mathrm{k}}(k)] - [(kc)^{\mathrm{k}}(sk-kc)(k+2)(kc+s)^{k+1}]}{(kc+s)^{2k+4}}\right] \times \left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta+i\beta}\right]\right]^2$$

$$E(T^2) = \frac{d^2l^*(s)}{ds^2}\Big|_{s=0} = \frac{2k+2}{(kc)^2}\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta+i\beta}\right]\right]^2 \quad \ldots (7)$$

(On simplification)

Hence,

$$V(T) = E(T^2) - [E(T)]^2 \quad \ldots (8)$$

Substitute equation (6) and (7) in equation (8), we get,

$$V(T) = \frac{2k+2}{(kc)^2}\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta+i\beta}\right]\right]^2 - \left[\frac{1}{kc}\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta+i\beta}\right]\right]\right]^2$$

$$V(T) = \frac{2k+1}{(kc)^2}\left[\theta(1-p)\sum_{i=1}^{k}\left[\frac{p^{i-1}}{\theta+i\beta}\right]\right]^2 \quad \ldots (9)$$

(On simplification)

Numerical Illustrations

Table 1

K	$\boldsymbol{\theta = 0.2, p = 0.5, \beta = 0.4, c = 1}$	
	E(T)	V(T)
1	0.142857	0.061224
2	0.071429	0.025510
3	0.047619	0.015873
4	0.035714	0.011480
5	0.028571	0.008980
6	0.023810	0.007370
7	0.020408	0.006247
8	0.017857	0.005421
9	0.015873	0.004787

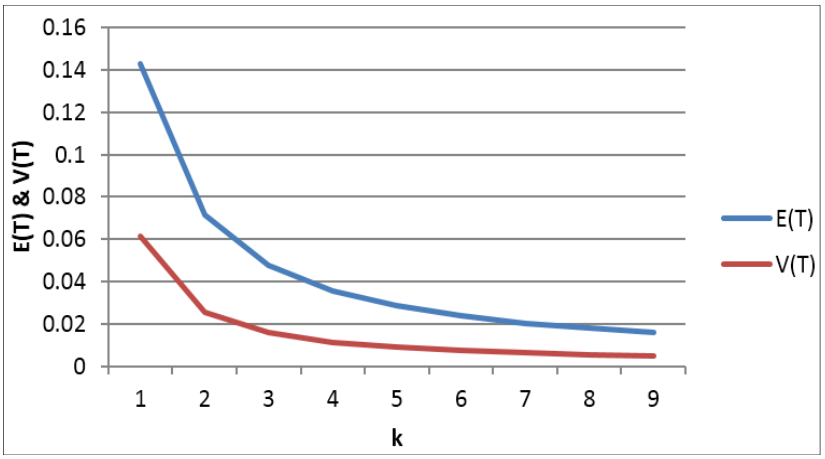

Fig 1

Table 2

c	$\theta = 0.4, p = 0.5, \beta = 0.3, k = 1$	
	E(T)	V(T)
1	0.222222	0.148148
2	0.111111	0.037037
3	0.074074	0.016461
4	0.055556	0.009259
5	0.044444	0.005926
6	0.037037	0.004115
7	0.031746	0.003023
8	0.027778	0.002315
9	0.024691	0.001829

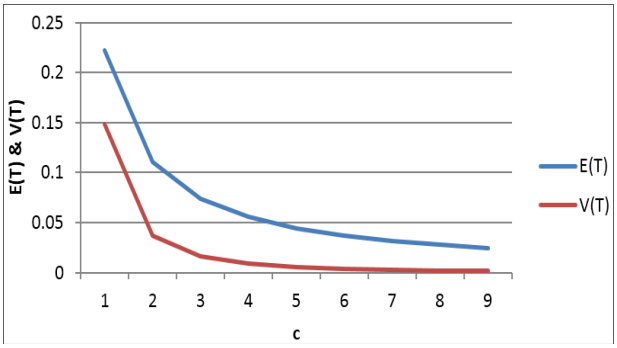

Fig2

Table 3

θ	$p = 0.5, \beta = 0.4, k = 1, c = 1$	
	E(T)	V(T)
0.1	0.083333	0.020833
0.2	0.142857	0.061224
0.3	0.187500	0.105469
0.4	0.222222	0.148148
0.5	0.250000	0.187500
0.6	0.272727	0.223140
0.7	0.291667	0.255208
0.8	0.307692	0.284024
0.9	0.321429	0.309949

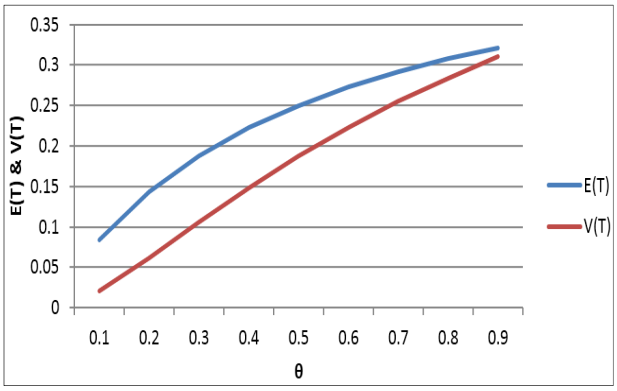

Fig 3

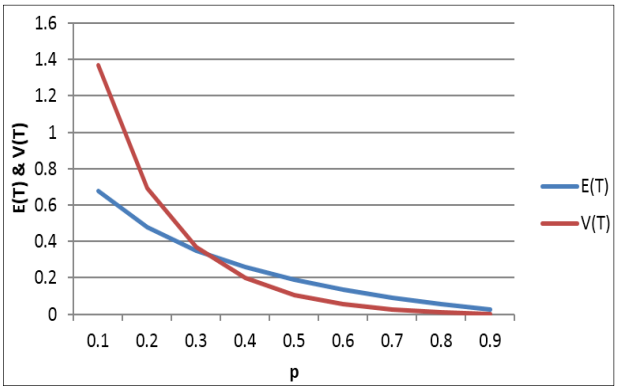

Fig 4

Table 4

p	$\theta = 0.3, \beta = 0.4, k = 1, c = 1$	
	E(T)	V(T)
0.1	0.675000	1.366875
0.2	0.480000	0.691200
0.3	0.350000	0.367500
0.4	0.257143	0.198367
0.5	0.187500	0.105469
0.6	0.133333	0.053333
0.7	0.090000	0.024300
0.8	0.054545	0.008926
0.9	0.025000	0.001875

Table 5

β	$\theta = 0.2, p = 0.5, k = 1, c = 1$	
	E(T)	V(T)
0.1	0.142857	0.061224
0.2	0.041667	0.005208
0.3	0.014706	0.000649
0.4	0.005682	9.68E-05
0.5	0.002315	1.61E-05
0.6	0.000977	2.86E-06
0.7	0.000422	5.35E-07
0.8	0.000186	1.04E-07
0.9	8.31E-05	2.07E-08

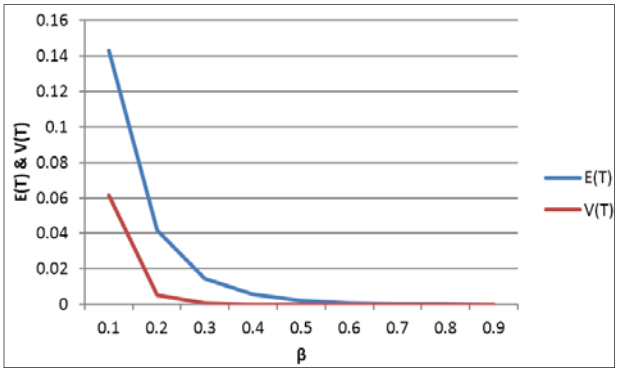

Fig 5

Conclusions

As the value of 'k' increases, the statistical measures for time to seroconversion decrease. This is due to the fact that, as 'k' namely, number of contact in (0, t] increases it means that, the contact are more frequent. Hence, it takes less time to cross the threshold. It is easily seen from the table-1 and figure-1.

As the value of the parameter 'c' is the distribution of times interval between successive contacts shown an increase it means that, the average time interval between successive contacts, which is given by$E(U) = \frac{1}{C}$, Since U ~ exp(c), therefore time interval between successive contacts becomes smaller and hence the statistical measure for time to seroconversion decreases, as indicated in table-2 and figure-2.

As the value of 'θ' which is namely, the parameter of the random variable X_i denoting contribution to the antigenic diversity increases, then it seen that, the mean time to seroconversion and variance time to seroconversion both increases, as indicated in table-3 and figure-3.

In the table-4, the variation in statistical measure that consequences to the change in parameter 'p' is noted as the parameter of the threshold parameter p increases, then the statistical measure for the time to seroconversion decreases, as shown in figure-4.

It is observed from the table-5, and also the figure-5, as the value of 'β' which is the parameter of the exponential-geometric distribution of the threshold increases, the expected time to seroconversion as well as variance time to seroconversion are decreases.

Acknowledgement

We are immensely grateful to Dr. R. Sathiyamoorthi, former Professor and Head, Department of Statistics, Annnamalai University, Chidambaram and Dr. G. S. Harisekharan, WIPRO, Chennai for their invaluable constructive suggestions, guidance and intellectual support.

References

1. Esary JD, Marshall AW, Proschan F. Shock models and Wear processes, Ann. Probability, 1973; 1(4):627-649.
2. Kannan R, Chandrasekar K. A stochastic model for the estimation of expected time to seroconversion of hiv infected using order statistics'. International Journal of World Research. 2015; 1(23):1-8.
3. Kannan R, Ganesan A, Sathiyamoorthi R, Malarvizhi G. A Stochastic Model for the Estimation of Time to Seroconversion of HIV Infected using Order Statistics, Bio-Science Research Bulletin. 2008; 24(1):1-7.
4. Kannan R, Kavitha S, Sathiyamoorthi R. A stochastic approach to determine the expected time to seroconversion of HIV infected using order statistics, Journal of the Indian Academy of Mathematics. 2013; 35:131-141.
5. Kannan R, Thirumurugan A, Sathiyamoorthi R, Malarvizhi G. A stochastic model with the antigenic diversity threshold of HIV transmission using order statistics'. Bio-Science Research Bulletin. 2010; 26(1):45-55.
6. Nowak MA, May RM. Mathematical Biology of HIV Infections: Antigenic Variation and Diversity Threshold', Mathematical Biosciences. 1991; 106:1-21.
7. Ratchagar NP, Vijaya S, Sathiyamoorthi R, Kannan R. A stochastic model for estimation of expected time to seroconversion of HIV infected using order statistics, Proceedings of the National Conference on Mathematical and Computational Models, PSG College of Technology, Coimbatore, 2003, 37-42.
8. Stilianakis N, Schenzle D, Dietz K. On the Antigenic Diversity Threshold Model for AIDS, Mathematical Biosciences. 1994; 121:235-247.