**Scholars Journal of Applied Medical Sciences (SJAMS)** **ISSN 2320-6691 (Online)**
**Sch. J. App. Med. Sci., 2017; 5(11A):4284-4292** **ISSN 2347-954X (Print)**

(An International Publisher for Academic and Scientific Resources)
www.saspublishers.com DOI:10.36347/sjams.2017.v05i11.001

# A Stochastic Model with Antigenic Diversity Threshold Using Largest Order Statistics

**Kannan R, Iyappan M**
Department of Statistics Annamalai University, Annamalainagar- 608 002

**Original Research Article**

***Corresponding author**
*Kannan R*

**Article History**
*Received: 19.10.2017*
*Accepted: 06.11.2017*
*Published: 30.11.2017*

**Abstract:** In the study of HIV infection and AIDS, one of the important aspects of investigation is the estimation of time to cross antigenic diversity threshold of HIV infected. The estimation of time to cross antigenic diversity threshold over the time interval (0, t] is an important aspect which help medical intervention. We propose the stochastic model assuming the intercontact time between successive contact from a largest order statistics and threshold distribution is Generalized Rayleigh distribution. The expected time to seroconversion and its variance are derived and numerical illustration is also given using simulated data.
**Keywords:** Antigenic Diversity, Threshold, Acquired Immuno Deficiency Syndrome, Human Immunodeficiency Virus, Largest Order Statistics, Seroconversion

## INTRODUCTION

Acquired Immunodeficiency Syndrome (AIDS) is caused by the Human Immunodeficiency Virus (HIV), which crossed from primates into humans. Infection with HIV has produced one of the most dramatic epidemics of the twentieth century. It has spread worldwide, leaving no region of the world unaffected. Before effective therapies were developed, infection with HIV meant an inexorable decline in health until death was a welcome relief. Now that the capability to decrease viral replication has been achieved, those who can afford this expensive treatment survive by keeping the infection dormant not by eliminating the virus altogether. Unfortunately, the antiviral reagents available come with serious side effects, and resistance to these agents develops readily for a consistent percentage of patients. At the same time that therapies with specific antiviral agents have decreased morbidity and mortality, they have resulted in relaxation of appropriate public and private health measures, which threatens a recrudescence of epidemic infection.

For the study of HIV infected and AIDS the mathematical tools in combination with biological aspects gives a scientific orientation. Isham [1] has given a review of the mathematical modeling of the transmission dynamics of HIV infection and AIDS. The breakdown human immune system is very much based on the diversity of the antigen namely HIV accrues in successive sexual contact. The antigenic diversity threshold which mean the antigenic diversity crosser a particular level, then the human immune system collapse and seroconversion take place immediately. The antigenic diversity and estimation has been discussed by Stiliankis *et al*. [2], Nowak and May [3] and Kirschner *et al*. [4], Sathiyamoorthi and Kannan [5] used the shock model and cumulative damage process evolved by Esary *et al.* [6] to estimate the expected time to cross the antigenic diversity threshold.

Ratchagar *et al.* [7] have derived a model for the estimation of expected time to seroconversion of HIV infected using order statistics. Kannan *et al.* [8-12] have obtained a stochastic model for estimation of expected time to seroconversion of HIV infected using order statistics and threshold follows Gamma, Erlang-2, Exponentiated Exponential, Exponentiated Modified Weibull, Exponential Geometric distribution. In this paper, it is assumed that the threshold follows Generalized Rayleigh distribution and inter-arrival times form an order statistics and so they are not independent. The expected time to cross antigenic diversity threshold and its variance are obtained talking the interarrival times between contacts distributed as the largest order statistics. This is due to the fact that if the largest order statistics is taken it implies that if the largest order statistics is taken it implies that the interarrival times are becoming the larger. Hence, frequent contacts would not be possible which will have its impact on the time to seroconversion. In this study, the theoretical results are substantiated using numerical data.

**Assumptions of the model**

The following are the assumptions understanding in the model developed here

- Sexual contacts are the only source of HIV infection.
- An individual is exposed to a damage process acting on the immune system and damage is assumed to be linear and cumulative.
- If the total damage caused when crosses a threshold level Y which itself is a random variable, the seroconversion occurs and a person is recognized as infected.
- The interarrival times between successive contacts are identically and independently distributed random variables.
- The sequence of successive contacts and threshold level are independent.
- From the collection of large number of interarrival times between successive contacts of a person, random samples of 'k' observation are taken.

**Notations**

The notations used in this model are as follows

| | | |
|---|---|---|
| $X_i$ | : | a random variable denoting the amount of damage arising due to $i^{th}$ contact $X_i^{'s}$ are identically and independently distributed with p.d.f. $g(.)$ and c.d.f. $G(.)$. |
| $Y$ | : | a random variable representing the antigenic diversity threshold which follows Generalized Rayleigh distribution with parameter $'\mu'$ and $'\lambda'$ the p.d.f. $h(.)$ and c.d.f. $H(.)$. |
| $U_{(k)}$ | : | a random variable representing the interarrival time between the contacts which follows largest order statistics with p.d.f. $f_{u_{(k)}}(t)$ and c.d.f. $F_{u_{(k)}}(t)$. |
| $g_k(.)$ | : | 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**

It can be shown that

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

Let Y ~ Generalized Rayleigh Distribution $(\alpha, \lambda, \mu)$

$$\overline{H}(x) = 1 - \left[1 - e^{-\lambda(x-\mu)^2}\right] = e^{-\lambda(x-\mu)^2}$$

Hence

$$P\left[\sum_{i=1}^{k} X_i < Y\right] = \int_0^{\infty} g_k(x)\left[e^{-\lambda(x-\mu)^2}\right]$$

$$= \left[g^*\lambda(1-\mu)^2\right]^k$$

The survival function S(t) is

S(t) = P[T > t]

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

$$= \sum_{k=0}^{\infty} \left[ F_k(t) - F_{k+1}(t) \right] \left[ g^*\lambda(1-\mu)^2 \right]^k$$

$$= \sum_{k=0}^{\infty} \Pr\{\text{there are exactly k contacts in } (0,t]\}$$

* Pr{the cumulative total of antigenic diversity < Y}

L(t) = 1 – S(t)

$$= 1 - \left\{ \sum_{k=0}^{\infty} \left[ F_k(t) - F_{k+1}(t) \right] \left[ g^*\lambda(1-\mu)^2 \right]^k \right\}$$

$$L(t) = \left[ 1 - g^*\left(\lambda + \lambda\mu^2 - 2\lambda\mu\right) \right] \sum_{k=1}^{\infty} F_k(t) \left[ 1 - g^*\left(\lambda + \lambda\mu^2 - 2\lambda\mu\right) \right]^{k-1}$$

On simplification

$$l(t) = 2\left[ 1 - g^*\left(\lambda + \lambda\mu^2 - 2\lambda\mu\right) \right] \sum_{k=1}^{\infty} \left[ f_k(t) \right] \left[ 1 - g^*\left(\lambda + \lambda\mu^2 - 2\lambda\mu\right) \right]^{k-1}$$

Now Taking Laplace Stieltjes transform of l(t),which is denoted by l*(s), we have

$$l^*(s) = \frac{\left[ 1 - g^*\left(\lambda + \lambda\mu^2 - 2\lambda\mu\right) \right] f^*(s)}{\left[ 1 - g^*\left(\lambda + \lambda\mu^2 - 2\lambda\mu\right) f^*(s) \right]}$$

On simplification … (1)

Now, the random variables $U_i$ which denote the time intervals between contacts is taken to be the random variable $U_k$ which denotes the largest order statistics. Under the assumptions of the model case (i) which has been discussed as case (i), statistical measures for time to seroconversion are derived here under the assumption that $U_k$ denote the highest order statistics.

The time intervals between contacts $U_1, U_2, .....U_N$ are i.i.d. random variables and $U_{(1)} < U_{(2)} < ..... < U_{(N)}$ from k order statistics which are random variables that are not independent.

The probability density function of the largest order statistics is $f_{u(k)}(t) = K\left[F(t)\right]^{k-1} f(t)$

The Laplace Stieltjes transformation of the above equation is given by

$$f^*_{u(k)}(s) = \int_0^{\infty} e^{-st} K\left[F(t)\right]^{k-1} f(t)\,dt$$

Assuming that f(t) follows exp(c). It can be shown that

$$f^*_{u(k)}(t) = \frac{K!c^k}{(c+s)\ (2c+s)\ (3c+s)\cdots(kc+s)} \quad \text{… (2)}$$

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

$$l^*(s) = \frac{\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\right]\dfrac{k!c^k}{(c+s)(2c+s)(3c+s)....(kc+s)}}{\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\dfrac{k!c^k}{(c+s)(2c+s)(3c+s)....(kc+s)}\right]}$$

$$= \frac{\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\right]k!c^k}{\left[\{(c+s)(2c+s)(3c+s)....(kc+s)\}-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)k!c^k\right]}$$

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

$$= \frac{k!c^k\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\right]}{\left[\{(c+s)(2c+s)(3c+s)....(kc+s)\}-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)k!c^k\right]^2}$$

At s = 0

$$= \frac{k!c^k\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\right]}{\left[\{(c)(2c)(3c)....(kc)\}-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)k!c^k\right]^2}$$

$$E(T) = \frac{1}{k!c^k\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\right]}$$

On simplification

… (3)

Let $g^*(.) \square \exp(\alpha)$

$$g^*(\lambda) = \frac{\alpha}{\alpha+\lambda}, g^*2(\lambda\mu) = \frac{\alpha}{\alpha+2\lambda\mu}, g^*\left(\lambda\mu^2\right) = \frac{\alpha}{\alpha+\lambda\mu^2}$$

… (4)

Then (4) in (3), we can get

$$E(T) = \frac{1}{k!c^k\left[1-\left\{\dfrac{\alpha}{\alpha+\lambda}+\dfrac{\alpha}{\alpha+\lambda\mu^2}-\dfrac{\alpha}{\alpha+2\lambda\mu}\right\}\right]}$$

$$E[T] = \frac{\alpha^3+\alpha^2\lambda+\alpha^2\lambda\mu^2+2\alpha^2\lambda\mu+\alpha\lambda^2\mu^2+2\alpha\lambda^2\mu^3+2\alpha\lambda^2\mu+2\lambda^3\mu^3}{k!c^k\lambda\left[\alpha^2+\alpha^2\mu^2-2\alpha^2\mu+2\alpha\lambda\mu^2+2\lambda^2\mu^3\right]}$$

On simplification

….. (5)

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

$$= \frac{d}{ds}\left\{k!c^k\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\right](-1)\left[\begin{matrix}\{(c+s)(2c+s)(3c+s)....(kc+s)\}\\ -g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)k!c^k\end{matrix}\right]^{-2}\right\}$$

$$= \frac{2\left(k!c^k\right)\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\right]}{\left[\left\{(c+s)(2c+s)(3c+s)....(kc+s)\right\}-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)k!c^k\right]^3}$$

At s = 0

$$E(T^2) = \frac{2}{\left[k!c^k\right]^2\left[1-g^*\left(\lambda+\lambda\mu^2-2\lambda\mu\right)\right]^2} \quad \ldots (6)$$

Let $g^*(.) \square \exp(\alpha)$

$$g^*(\lambda) = \frac{\alpha}{\alpha+\lambda}, g^*2(\lambda\mu) = \frac{\alpha}{\alpha+2\lambda\mu}, g^*\left(\lambda\mu^2\right) = \frac{\alpha}{\alpha+\lambda\mu^2}$$

$$E\left(T^2\right) = \frac{2\left(\begin{array}{l}\alpha^3+\alpha^2\lambda+\alpha^2\lambda\mu^2+2\alpha^2\lambda\mu+\alpha\lambda^2\mu^2 \\ \qquad +2\alpha\lambda^2\mu^3+2\alpha\lambda^2\mu+2\lambda^3\mu^3\end{array}\right)^2}{\left(k!c^k\right)^2\lambda^2\left(2\lambda^2\mu^3+2\alpha\lambda\mu^2+\alpha^2+\alpha^2\mu^2-2\alpha^2\mu\right)^2}$$

On simplification

Hence, the variance of time to seroconversion is V(T) = E(T2) – [E(T)]2

$$V[T] = \frac{\left(\begin{array}{l}\alpha^3+\alpha^2\lambda+\alpha^2\lambda\mu^2+2\alpha^2\lambda\mu+\alpha\lambda^2\mu^2 \\ \qquad +2\alpha\lambda^2\mu^3+2\alpha\lambda^2\mu+2\lambda^3\mu^3\end{array}\right)^2}{\left(k!c^k\right)^2\lambda^2\left(2\lambda^2\mu^3+2\alpha\lambda\mu^2+\alpha^2+\alpha^2\mu^2-2\alpha^2\mu\right)^2}$$

On simplification

**Numerical Illustrations**

**Table – 1: Seroconversion time**

| $k$ | $\mu = 0.5, \alpha = 0.2, \lambda = 0.3$ $c = 0.1$ | |
|---|---|---|
| | **Mean** | **Variance** |
| 1 | 00036.66667 | 000001344.44444 |
| 2 | 00183.33333 | 000033611.11111 |
| 3 | 00611.11111 | 000373456.79012 |
| 4 | 01527.77778 | 002334104.93827 |
| 5 | 03055.55556 | 009336419.75309 |
| 6 | 05092.59259 | 025934499.31413 |
| 7 | 07275.13228 | 052927549.62067 |
| 8 | 09093.91534 | 082699296.28230 |
| 9 | 10104.35038 | 102097896.64481 |
| 10 | 10104.35038 | 102097896.64481 |

**Table-2: Seroconversion time**

| $\mu$ | $\alpha = 0.2, \lambda = 0.3, k = 2$ $c = 1$ | |
|---|---|---|
| | **Mean** | **Variance** |
| 0.5 | 1.83333333 | 3.36111111 |
| 1 | 1.11111111 | 1.23456790 |
| 1.5 | 0.90375587 | 0.81677467 |
| 2 | 0.83333333 | 0.69444444 |
| 2.5 | 0.80481405 | 0.64772566 |
| 3 | 0.79234973 | 0.62781809 |
| 3.5 | 0.78697572 | 0.61933078 |
| 4 | 0.78502415 | 0.61626292 |
| 4.5 | 0.78481477 | 0.61593422 |
| 5 | 0.78551390 | 0.61703209 |

**Table – 3: Seroconversion time**

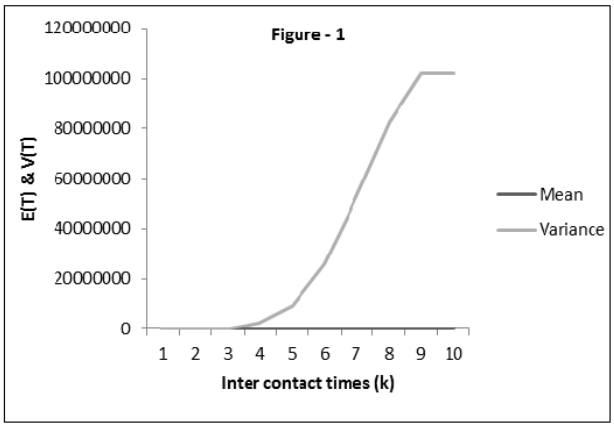
| λ | $\mu = 0.5, \alpha = 0.2, k = 2$ $c = 1$ | |
|---|---|---|
| | **Mean** | **Variance** |
| 0.5 | 1.30000000 | 1.69000000 |
| 1 | 0.90000000 | 0.81000000 |
| 1.5 | 0.76666667 | 0.58777778 |
| 2 | 0.70000000 | 0.49000000 |
| 2.5 | 0.66000000 | 0.43560000 |
| 3 | 0.63333333 | 0.40111111 |
| 3.5 | 0.61428571 | 0.37734694 |
| 4 | 0.60000000 | 0.36000000 |
| 4.5 | 0.58888889 | 0.34679012 |
| 5 | 0.58000000 | 0.33640000 |

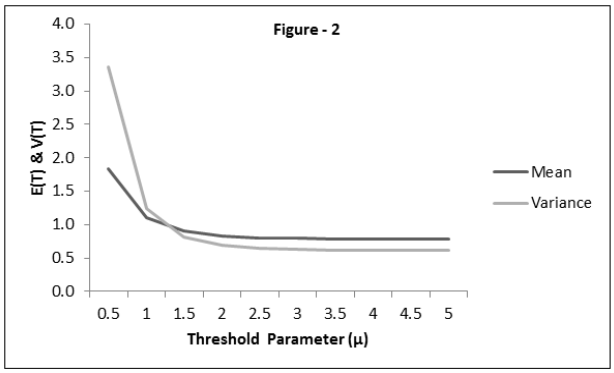
**Table – 4: Seroconversion time**

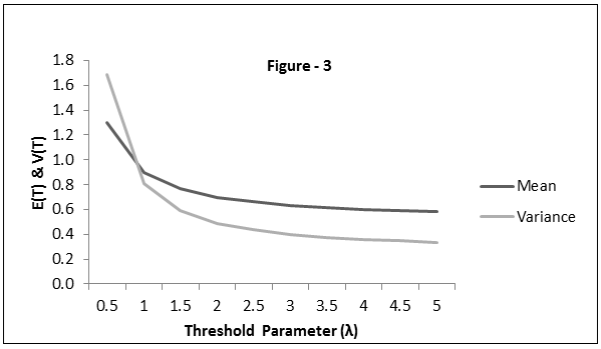
| α | $\mu = 0.2, \lambda = 0.3, k = 2$ $c = 1$ | |
|---|---|---|
| | **Mean** | **Variance** |
| 0.1 | 1.16666667 | 1.36111111 |
| 0.2 | 1.83333333 | 3.36111111 |
| 0.3 | 2.50000000 | 6.25000000 |
| 0.4 | 3.16666667 | 10.02777778 |
| 0.5 | 3.83333333 | 14.69444444 |
| 0.6 | 4.50000000 | 20.25000000 |
| 0.7 | 5.16666667 | 26.69444444 |
| 0.8 | 5.83333333 | 34.02777778 |
| 0.9 | 6.50000000 | 42.25000000 |
| 1 | 7.16666667 | 51.36111111 |

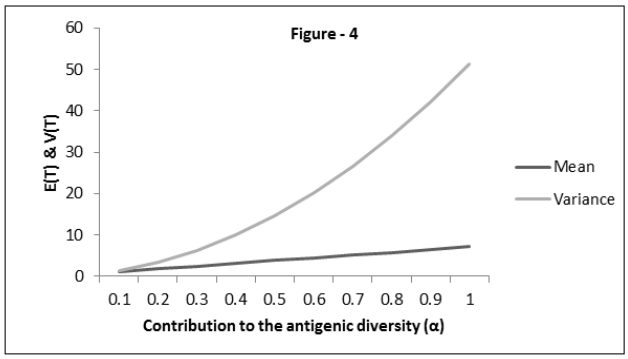
**Table – 5: Seroconversion time**

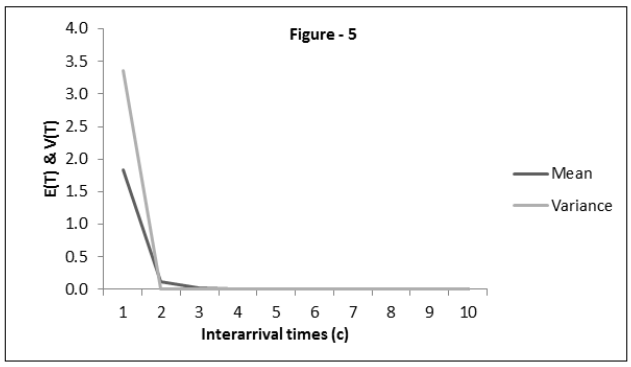
| $c$ | $\mu = 0.5, \alpha = 0.2, \lambda = 0.3$ $k = 2$ | |
|---|---|---|
| | **Mean** | **Variance** |
| 1 | 1.83333333 | 3.36111111 |
| 2 | 0.11458333 | 0.01312934 |
| 3 | 0.02263374 | 0.00051229 |
| 4 | 0.00716146 | 0.00005129 |
| 5 | 0.00293333 | 0.00000860 |
| 6 | 0.00141461 | 0.00000200 |
| 7 | 0.00076357 | 0.00000058 |
| 8 | 0.00044759 | 0.00000020 |
| 9 | 0.00027943 | 0.00000008 |
| 10 | 0.00018333 | 0.00000003 |






## CONCLUSIONS

- The value of both E(T) and V(T) increases with an increase in 'k' (Table -1) namely the number of contacts. If 'k' becomes larger than the corresponding $U_k$ the mean and variance of time to seroconversion also become larger thereby implying that it is the largest of the interarrival times in such a case the inter contact times are elongated thereby having a delayed time to seroconversion.
- If $\mu$, which is parameter of threshold which follows Generalized Rayleigh distribution, increases then the expected time to seroconversion decreases. This is due to the fact that E(T) decreases if $\mu$ increases. Hence the average threshold level is less; and hence it takes less time to cross the same. Hence the variance of seroconversion also decreases as indicated in Table -2 and Fig-2.

The behavior of E(T) for $k, \alpha, c$ and $\mu$ but with variation in $\lambda$, is such that an increase in $\lambda$ which is the parameter of Generalized Rayleigh distribution of threshold increases then the expected time to seroconversion and its variance are on the decrease as indicated in Table -3 and Fig-3.

- As the value of 'α' which is namely the parameter of the random variable $X_i$ denoting contribution to the antigenic diversity increase then it is seen that mean time to seroconversion and variance time to seroconversion are increase as indicated in Table – 4 and Fig-4.
- As the value of parameter of 'c' which is namely parameter of the distribution of the interarrival times between the contacts increase it means that the average interarrival times which is given by $E(U) = \frac{1}{c}$ since U~ Exp(c), therefore interarrival times between the contacts become smaller and hence the mean time to seroconversion and its variance also decreases as indicated in Table -5 and Fig-5.

## REFERENCES

1. Isham V. Mathematical modelling of the transmission dynamics of HIV infection and AIDS: a review. Journal of the Royal Statistical Society. Series A (Statistics in Society). 1988 Jan 1:5-49.
2. Stilianakis NI, Schenzle D, Dietz K. On the antigenic diversity threshold model for AIDS. Mathematical biosciences. 1994 Jun 1;121(2):235-47.

3. Nowak MA, May RM, Anderson RM. The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease. Aids. 1990 Nov 1;4(11):1095-103.
4. Kirschner DE, Webb GF, Cloyd M. Model of HIV-1 disease progression based on virus-induced lymph node homing and homing-induced apoptosis of CD4+ lymphocytes.
5. Sathiyamoorthi R, Kannan R. A stochastic model for time to seroconversion of HIV transmission. Journal of Kerala Statistical Association. 2001 Dec 23;12:23-39.
6. Esary JD, Marshall AW. Shock models and wear processes. The annals of probability. 1973 Aug 1:627-49.
7. Kannan R, Karthi R. A stochastic approach to determine the statistical measure for time to seroconversion of hiv infected using largest order statistics. International Journal of Mathematical Archive (IJMA) 2017 Feb 15;8(1).
8. 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.
9. Kannan R, Ganesan A, Thirumurugan A. Stochastic model for estimation of expected time to seroconversion of HIV infected using largest order statistics. Int. J. Agricult. Stat. Sci. 2009 Dec 1;5(2):397-404.
10. Kannan R, Kavitha S, Sathiyamoorthi R. A stochastic approach to determine the expected time to seroconversion of hiv infected using geometric process. International Journal of Mathematical Archive (IJMA) 2013 Feb 27;4(1).
11. Kannan R, Chandrasekar K. A stochastic model for estimation of expected time to seroconversion of hiv infected using alertness. International Journal of Mathematical Archive. 2015 May 27;6(4).
12. Kannan R, Karthi R. A stochastic approach to determine the statistical measure for time to seroconversion of hiv infected using largest order statistics. International Journal of Mathematical Archive (IJMA) 2017 Feb 15;8(1).