How to Cite:
Vanitha, P., Pandiyan, P., Senthilkumar, B., & Sudha, V. (2022). A stochastic model for generalized Rayleigh distribution using three sources. *International Journal of Health Sciences*, *6*(S8), 4750–4756. https://doi.org/10.53730/ijhs.v6nS8.13293

A stochastic model for generalized Rayleigh distribution using three sources

P. Vanitha
Research scholar, Department of Statistics, Annamalai University, Annamalai Nagar, Tamil Nadu, India

P. Pandiyan
Department of Statistics, Annamalai University, Annamalai Nagar, Tamil Nadu, India

B. Senthilkumar
Assistant Professor, Thanthai Periyar Government Arts and Science College (Autonomous) Tiruchitappali

V. Sudha
Assistant Professor, Community Medicine Aarupadai Veedu Medical College & Hospital Puducherry
Corresponding author email: subashstat@gmail.com

***Abstract*---**Of all Cancers, Oral cancer is one of the most typical cancers in the world. Oral cancer has delayed clinical detection, poor prognosis, without specific biomarkers, and expensive therapeutic methods. This study aims to explain the basic aspects of oral cancers and focuses on squamous cell carcinoma of the oral cavity. The duration of cancer detection depends on the amount and duration of tobacco usage of the individual. In this study, we have presented stochastic models for squamous cell carcinoma onset and progression of oral cancer-affected individuals.

***Keywords*---**oral cancers, stochastic model, tobacco usage, expected life time.

Introduction

Globally, cancers have become a public health problem. Squamous cell carcinomas are the majority of cancers of the upper aero-digestive tract excluding the nasopharynx. Several risk factors are common for cancers of the lip, tongue, oral cavity, and oropharynx and also have similar biology. The developing countries contribute to the pattern of oral and oropharyngeal cancer incidences

that have considerable variation in different parts of the globe. The higher incidence rate of intra-oral cancers is seen in the population with high consumption of tobacco.

Due to the cancer being discovered late in its development, the death rate associated with the cancer is particularly high. Human papilloma virus16 also contributes more to the incidence rate of oral cancers, particularly in the posterior part of the mouth, which many times does not produce visible lesions or discoloration and can be discovered in history taking and oral examination.Mostly oral cancer is diagnosed when it is metastasized, mostly in the lymph nodes of the neck. The diagnosis at this stage is significantly worse than when it is diagnosed in a localized primary site.

The stochastic model and mathematical model are developed in areas of biology, medicine, and engineering. The mathematical models are being used in the study of tumor cell growth also in other diseases. Tobacco usage is the main risk factor for oral cancer. The use of these models helps to arrive at the conclusion and developed taking into consideration the problems in real life. The aim of the study is to find out the early detection of oral cancers and to analyze a stochastic model make it potential to predict it.

These mathematical models can be readily converted into a computer project that evaluates the effects of the disease and the potential spread of the disease. The threshold beyond which the human immune system cannot withstand is represented as the sum of two random variables. The Generalized Rayleigh distribution, which is a special case of both exponential and Wake by distribution, has good potential for the analysis of flood peaks because of its inherent properties. The point after which radical changes are likely to occur is called the threshold level.

Assumption of the model

This assumption is somewhat artificial, but it is made on the one hand due to the lack of detailed information from the real world and on the other hand to describe the production.

- Cigarette smoking, chewing, and alcohol consumption habits are the source of growth for cancer.
- The threshold for each individual is a random variable. If the total damage exceeds the threshold value Y, which in itself is a random variable, habit occurs and a person is recognized as infected.
- The arrival time between successive events, the consequences of failure, and the threshold are independent of each other.
- If the total damage exceeds the threshold value Y, which in itself is a random variable, habit occurs and a person is recognized as infected.
- The damages due to the events namely Cigarette smoking, chewing and alcohol products are statistically independent
- If the cumulative damage due to successive events crosses the antigenic diversity threshold level seroconversion takes place. The inter-arrival times

between habits, Cigarette smoking, chewing, and alcohol are statistically independent.

Notations

X_i: A continuous random variable denoting the amount of contribution to the antigenic diversity due to the alcohol use and smoking in the i^{th} habits, in other words the damage caused to the tumor growth in the i^{th} habits, with p.d.f $g(.)$ and c.d.f $G(.)$.

Y_1, Y_2, Y_3 A continuous random variable denoting the threshold level having Generalized Rayleigh distribution.

$g(.)$: The probability density functions of X_i

$g^*(.)$: Laplace transform of g (.)

$g_k(.)$: The k- fold convolution of g (.) i.e., p.d.f. of $\sum_{j=1}^{k} X_i$

$g_k^*(.)$: Laplace transform of $g_k(.)$.

$h(.)$: The p.d.f. of random threshold level which has Generalized Rayleigh distribution and $H(.)$ is the corresponding c.d.f.

U_i: a random variable denoting the inter-arrival times between habits with c.d.f. $F_i(.)$, $i = 1,2,3 \dots k$.

$f(.)$: p.d.f. of random variable denoting between successive habits with the corresponding c.d.f. $F(.)$

$F_k(.)$: The k-fold convolution functions of $F(.)$

$S(.)$: The survivor function, i.e.$P[T > t]$

$L(t)$: $1 - S(t)$.

$V_k(t)$: Probability that there are exactly k habits

Model description

In this paper having the threshold which follows Generalized Rayleigh distribution is discussed with the shape parameter α=n is been considered. The expected time and variance are obtained. The three -parameters generalized Rayleigh distribution is a particular member of the generalized Weibull distribution, originally proposed by Mudholkar and Srivastava (1993).

$$F(x) = \left[1 - e^{-\left(\frac{x-d}{b}\right)}\right]\left[1 - e^{-\left(\frac{x-d}{b}\right)}\right]\left[1 - e^{-\left(\frac{x-d}{b}\right)}\right]$$

$$= 1 - e^{-\left(\frac{x-d}{b}\right)} - 2e^{-\left(\frac{x-d}{b}\right)} + 2\left[e^{-\left(\frac{x-d}{b}\right)}\right]^2 + 2\left[e^{-\left(\frac{x-d}{b}\right)}\right]^2 - \left[e^{-\left(\frac{x-d}{b}\right)}\right]^3$$

$$= 1 - 3e^{-\left(\frac{x-d}{b}\right)} + 3e^{-2\left(\frac{x-d}{b}\right)} - e^{-3\left(\frac{x-d}{b}\right)}$$

$$\bar{H}(x) = 1 - \left[1 - 3e^{-\left(\frac{x-d}{b}\right)} + 3e^{-2\left(\frac{x-d}{b}\right)} - e^{-3\left(\frac{x-d}{b}\right)}\right]$$

$$= 1 - 1 + 3e^{-\left(\frac{x-d}{b}\right)} - 3e^{-2\left(\frac{x-d}{b}\right)} + e^{-3\left(\frac{x-d}{b}\right)}$$

$$= 3e^{-\left(\frac{x-d}{b}\right)} - 3e^{-2\left(\frac{x-d}{b}\right)} + e^{-3\left(\frac{x-d}{b}\right)}$$

$$\bar{H}(x) = 3e^{\left(\frac{d-x}{b}\right)} - 3e^{2\left(\frac{d-x}{b}\right)} + e^{3\left(\frac{d-x}{b}\right)} \quad \dots (1)$$

$$P(X_i < Y) = \int_0^\infty g_k^*(x)\bar{H}(x)dx \quad \dots (2)$$

$$P(X_i < Y) = \int_0^\infty g^*(x)\bar{H}(x)dx$$

$$P\big((X_i < Y)\big) = \int_0^\infty g^*(x)3e^{\left(\frac{d-x}{b}\right)} - 3e^{2\left(\frac{d-x}{b}\right)} + e^{3\left(\frac{d-x}{b}\right)}dx$$

By taking Laplace transform, we get

$$= \int_0^\infty g_k(x)3e^{\left(\frac{d-x}{b}\right)} - 3e^{2\left(\frac{d-x}{b}\right)} + e^{3\left(\frac{d-x}{b}\right)}dx \qquad \text{By convolution theorem}$$

Now the threshold Y is such that it has two habits namely $Y_1$and Y_2. Transfer of Immune system from Y_1 to Y_2 is also possible. We have the threshold level of seroconversion is given by $Y = max(Y_1, Y_2, Y_3)$.

$P[max(Y_1, Y_2, Y_3)] = P[(Y_1 < y) \cap (Y_2 < y) \cap (Y_3 < y)] = P[Y_1 < y]P[Y_2 < y]P[Y_3 < y]$

Now that, Y_1 and Y_2Y_2follow Generalized Rayleigh distribution with parameterb, d *and* α .

$$P\left(\sum_{i=1}^{k} X_i < Y\right) = 3\int_0^\infty g^*(x)e^{-\left(\frac{x-d}{b}\right)} - 3\int_0^\infty g^*(x)e^{\left(\frac{2x-2d}{b}\right)} + \int_0^\infty g^*(x)e^{\left(\frac{3x-3d}{b}\right)}dx$$

$$= 3\left[g^*\left(\frac{1-d}{b}\right)\right]^k - 3\left[g^*\left[2\left(\frac{1-d}{b}\right)\right]\right]^k + \left[g^*\left[3\left(\frac{1-d}{b}\right)\right]\right]^k \qquad \ldots(3)$$

Probability that the total damage does not cross the threshold level till time t.

$S(t) = P(T > t)$ = Probability that the total damage survives beyond t

$$= \sum_{k=0}^{\infty} \text{P \{there are exactly k habits in } (0, t] * \text{P (the total cumulative } (0, t]\}$$

$$S(t) = P(T > t) = \sum_{k=0}^{\infty} V_k(t)P\big(X_i < max(Y_1, Y_2, Y_3)\big) \qquad \ldots(4)$$

It may happen that successive shocks become increasingly effective in causing damage, even though they are independent. This means that $V_k(t)$ is the probability and that there are exactly k occasions of damage in (o,t).

P(exactly k habits in $(0, t]) = F_k(t) - F_{K+1}(t)$ with $F_0(t) = 1$

$$= \sum_{k=0}^{\infty}[F_k(t) - F_{k+1}(t)]\left[g^*\left(\frac{1-d}{b}\right)\right]^k + \sum_{k=0}^{\infty}[F_k(t) - F_{k+1}(t)]\left[g^*\left(\frac{1-d}{b}\right)\right]^k - \sum_{k=0}^{\infty}[F_k(t) - F_{k+1}(t)]\left[g^*\,2\left(\frac{1-d}{b}\right) + \left(\frac{1-d}{b}\right)\right]^k \qquad \ldots(5)$$

$L(t) = 1 - S(t)$, Taking laplace transform of $L(t)$, We get

$$L(t) = 1 - \left\{3\sum_{k=0}^{\infty}[F_k(t) - F_{k+1}(t)]\left[g^*\left(\frac{1-d}{b}\right)\right]^k - 3\sum_{k=0}^{\infty}[F_k(t) - F_{k+1}(t)]\left[g^*\left(\frac{2-2d}{b}\right)\right]^k + 3\sum_{k=0}^{\infty}[F_k(t) - F_{k+1}(t)]\left[g^*\left(\frac{3-3d}{b}\right)\right]^k\right\} \qquad \ldots(6)$$

On simplification we get,

$$L(t) = 3\left[1 - g^*\left(\frac{1-d}{b}\right)\right]\sum_{k=1}^{\infty} F_k(t)\left[g^*\left(\frac{1-d}{b}\right)\right]^{k-1} - 3\left[1 - g^*\left(\frac{2-2d}{b}\right)\right]\sum_{k=1}^{\infty} F_k(t)\left[g^*\left(\frac{2-2d}{b}\right)\right]^{k-1} + \left[1 - g^*\left(\frac{3-3d}{b}\right)\right]\sum_{k=1}^{\infty} F_k(t)\left[g^*\left(\frac{3-3d}{b}\right)\right]^{k-1} \quad \ldots(7)$$

By taking Laplace-Stieltjes transform, it can be shown that

$$l^*(s) = \frac{3\left[1 - g^*\left(\frac{1-d}{b}\right)\right]f^*(s)}{\left[1 - g^*\left(\frac{1-d}{b}\right)f^*(s)\right]} - \frac{3\left[1 - g^*\left(\frac{2-2d}{b}\right)\right]f^*(s)}{\left[1 - g^*\left(\frac{2-2d}{b}\right)f^*(s)\right]} + \frac{\left[1 - g^*\left(\frac{3-3d}{b}\right)\right]f^*(s)}{\left[1 - g^*\left(\frac{3-3d}{b}\right)f^*(s)\right]} \quad \ldots(8)$$

Let the random variable U denoting inter arrival time which follows exponential with parameter$c$. Now $f^*(s) = \left(\frac{c}{c+s}\right)$, substituting in the above equation (8) we get

$$= \frac{3\left[1 - g^*\left(\frac{1-d}{b}\right)\right]\left(\frac{c}{c+s}\right)}{\left[1 - g^*\left(\frac{1-d}{b}\right)\left(\frac{c}{c+s}\right)\right]} - \frac{3\left[1 - g^*\left(\frac{2-2d}{b}\right)\right]\left(\frac{c}{c+s}\right)}{\left[1 - g^*\left(\frac{2-2d}{b}\right)\left(\frac{c}{c+s}\right)\right]} + \frac{\left[1 - g^*\left(\frac{3-3d}{b}\right)\right]\left(\frac{c}{c+s}\right)}{\left[1 - g^*\left(\frac{3-3d}{b}\right)\left(\frac{c}{c+s}\right)\right]} \quad \ldots(9)$$

$$E(T) = -\frac{d}{ds}l^*(s) given\ s = 0, \qquad E(T^2) = \frac{d^2}{ds^2}l^*(s)\ given\ s = 0$$

From which variance $V(T) = E(T^2) - [E(T)]^2$can be obtained

$$\mathrm{E}(T) = \frac{3}{c\left[1 - g^*\left(\frac{1-d}{b}\right)\right]} - \frac{3}{c\left[1 - g^*\left(\frac{2-2d}{b}\right)\right]} + \frac{1}{c\left[1 - g^*\left(\frac{3-3d}{b}\right)\right]} \quad \text{on simpilification}$$

$$\mathrm{E}(T) = \frac{3b}{c[(1-d+\mu b)]} - \frac{3b}{c[2(1-d)+\mu b]} + \frac{b}{c[3(1-d)+\mu b]} \quad \text{on simpilification}$$

$$\mathrm{E}(T^2) = \frac{3}{c^2\left[1 - g^*\left(\frac{1-d}{b}\right)\right]^2} - \frac{3}{c^2\left[1 - g^*\left(\frac{2-2d}{b}\right)\right]^2} + \frac{1}{c^2\left[1 - g^*\left(\frac{3-3d}{b}\right)\right]^2}$$

$$\mathrm{E}(T^2) = \frac{3b^2}{c^2[(1-d+\mu b)]^2} - \frac{3b^2}{c^2[2(1-d)+\mu b]^2} + \frac{b^2}{c^2[3(1-d)+\mu b]^2}$$

$V(T)=E(T^2)-(E(T))^2$

$$V(T) = \frac{18b^2}{C^2[1-d+\mu b][2(1-d)+\mu b]} + \frac{6b^2}{C^2[2(1-d)+\mu b][3(1-d)+\mu b]} - \frac{6b^2}{C^2[1-d+\mu b][3(1-d)+\mu b]} - \frac{6b^2}{C^2[1-d+\mu b]^2} - \frac{6b^2}{C^2[2(1-d)+\mu b]^2}$$

Table: 1 infected person's stage wise

C	μ= (FEV1)	λ= (FVC)	E(T)	V(T)
1	84	88	1.01085	1.022
2	76	83	0.50552	0.256
3	72	80	0.33708	0.114
4	72	76	0.25312	0.064

5	61	72	0.20235	0.041
6	61	65	0.16907	0.029
7	58	55	0.14560	0.021
8	55	53	0.12745	0.016
9	50	50	0.11333	0.013
10	43	49	0.10179	0.010

Fig.1: The Chart for infected person's Expected time E(T)

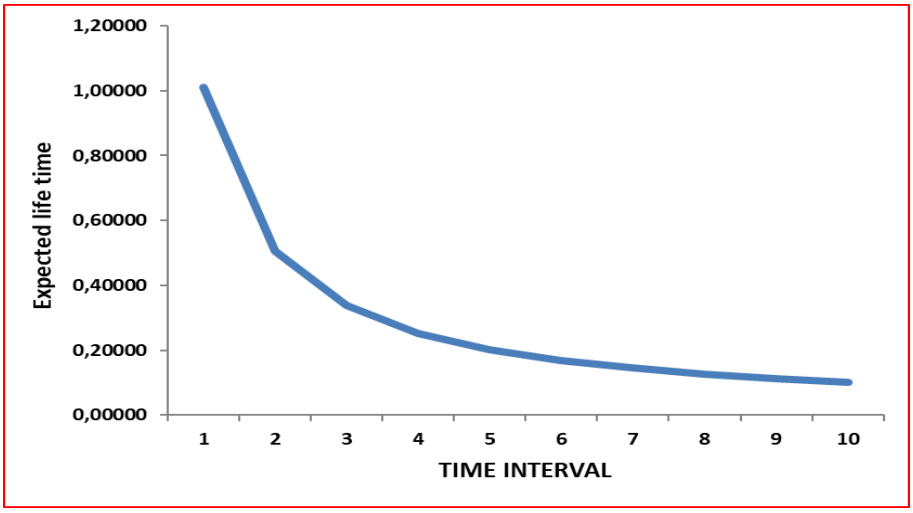
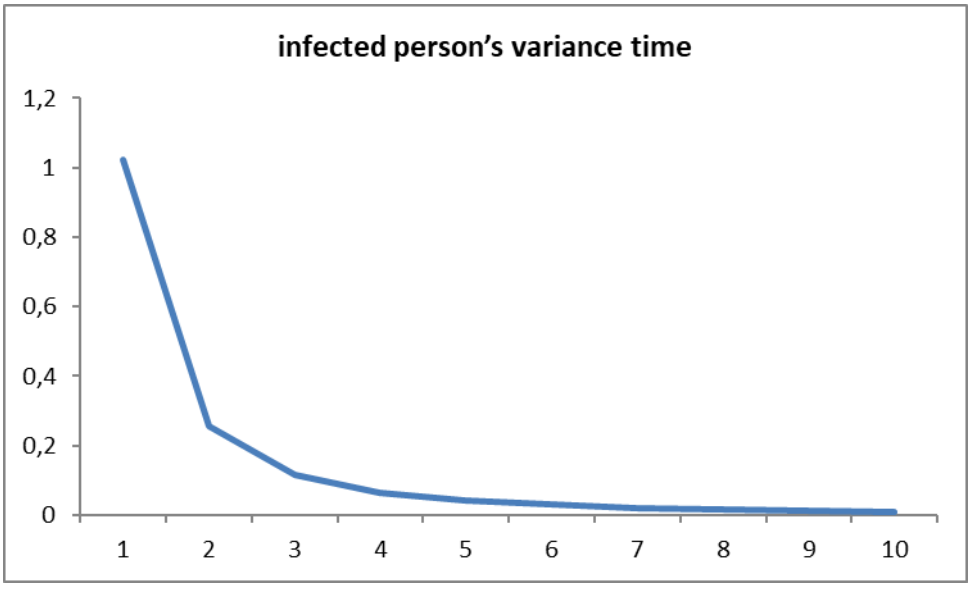


Fig – 2: The Chart for infected person's variance time



Conclusion

In the present study, people who use tobacco, especially smokeless tobacco will have lesser life expectancy than non-users. In this study, we used a technique to model the tobacco usage behaviour of the study participants. We simulate transient and permanent event one after other. These risk factors like smoking tobacco, smokeless tobacco usage have completely different effects on these three breakup states. In this mathematical modelling, we recommend treatment for which lesions statistically located but not diagnosed. With the shock model approach, the cumulative damage is found which exploits the upgrade method. Through this distribution, the expected time is obtained through the model with the data. In oral cancer, the length of the time interval depends upon the duration of tobacco usage of the cancer affected person. The model shows that, the immune system breaks down once the person get cancer. The life span can be extended with proper treatment and follow-up

References

1. Ammar Sarhan and Zaindin M. Modified Weibull distribution, Applied Sciences, 11: 123–136, (2009).
2. P. Pandiyan, S. Koventhan, G. Subash Chandra bose Estimating The Survival Time Of cancer Drinker Patients Using Stochastic Model, Indo – Asian Journal of Multidisciplinary Research Volume – 4(4); 1464 – 1467 ,(2018)
3. Pandiyan P and Koventhan S. Estimating of survival time of cancer Patients, International Journal of Recent Scientific Research, 9(1), 23533 – 23534, (2018).
4. Pandiyan P. Calculating Human Immune Virus Infected Person Survival Time Through Stochastic Model, Bio- Science Research Bulletin, 29(1): 1 - 4(2013).
5. S. Jothimanickam & P. Pandiyan (2019), Calculating The Survival Time Of Cancer Patients Through Exponentiated Weibull Distribution, Journal of International and Computational Science,Volume 9, Issue 12.
6. S. Jothimanickam, P. Pandiyan, E. Susiganeshkumar & T. Vignesh (2020), Stochastic Model for Expected Time Using Generalized Exponential Distribution,Journal Of Xi'an University Of Architecture &Technology, Volume XII, Issue IX, September 2020.
7. Yadavalli, V.S.S., Moremi Morire OreOluwapo Labeodan and Swaminathan Udayabaskaran, (2011). “A Stochastic point process model of the incubation period of a HIV infected individual”, *Journal of South African Statistic*s, Vol.45, pp.293-314