Int. J. Math. And Appl., 9(2)(2021), 219–222
ISSN: 2347-1557
Available Online: http://ijmaa.in/

International Journal of *Mathematics And its Applications*

Renewal Process for Finding the Occurrence Rate of an Event

V. G. Rajalekshmi[1,*]

1 Department of Mathematics, S. D. Collge, Alappuzha, Kerala, India.

Abstract: Data mining is the technique of mining pattern from data .One of the technique in data mining is prediction. There are many models in various subjects which help to predict the occurrence of certain events in the data base. In this paper we use the Renewal process for finding occurrence rate events in the transactional data base.

Keywords: Order of an event, Occurrence rate of an event, Generating Function, Recursion equation for moments.

1. Introduction

A data base is a collection of information that is organized so that it can be easily accessed, managed and updated. A transactional database is a data base management system (DBMS). Examples includes sales orders, invoices purchase orders, shipping documents, insurance claims etc. These data are typically grouped into transactional records. Transactional data base is the collection of some subset of a universal item set. In this paper we consider the occurrence of an item in a transaction as an event. Here we consider a one-one correspondence with the transactional data in the data base and the set of all natural numbers. Order of the first event is the number of the first transactional data which contain the item A or the first transactional data in which the event occurs. It is denoted by n_1. Here we are going to find the rate of occurrence of event with respect to the number of data.

2. Related Work

Renewal theorem and its theory can be applied in all fields of life. The theory stochastic renewal process and the renewal theorem have been fundamental to the development of risk-based asset management models [2], in bridge management the renewal theory has been applied [5]. Although the renewal process has been discussed in many mathematical treaties [1, 3] the concepts are not amenable to the engineering community. A conceptually simple and intuitive interpretation of the renewal process with applications are given in [4].

* *E-mail: rajisreeramam@gmail.com*

3. Model Description

Here we consider the set of transactional data in the database has a one to one correspondence with natural numbers. That is, we can name the collection of transactional data by $T_1, T_2, \ldots$. Let A be an arbitrary item in the item set. We have to find those transactional data which contains the item 'A'. Such transactional data are called event. The i$^{\text{th}}$ event is the transactional data containing i$^{\text{th}}$ occurrence of the item. Order of the first event is the number of the first transactional data which contain the item A or the first transactional data in which the item A occurs. It is denoted by n_1. The order of the i$^{\text{th}}$ event is the total number of transactional data between the $(i-1)^{\text{th}}$ event and i$^{\text{th}}$ event including the transactional data containing the i$^{\text{th}}$ event. It is denoted by n_i, $i = 1, 2, 3, \ldots$, $0 \leq n_i$ for all i. Here occurrence of an event is considered as a renewal occurrence. The order of the event which we have defined is called the renewal period of the process. Corresponding to every transactional data we can define a random variable

$$X_i, i = 1, 23, \ldots.$$

Such that

$$X_i = \begin{cases} 1, & \text{if } A \in T_i; \\ 0, & \text{if } A \notin T_i. \end{cases}$$

The order of event (n) is a sequence of random variables. The probability distributes of order of event is given by

$$P(n = k) = p_k$$

Let X_i be the random variable representing the occurrence of the event.

$$X_j = \begin{cases} 1, & \text{if the j}^{\text{th}} \text{ data contain the item;} \\ 0, & \text{otherwise.} \end{cases}$$

Data which contain the item is denoted by $S_k \ \forall \ k$ therefore. We can write

$$X_j = \sum_{k=1}^{j} 1_{(S_k = j)}, \quad j \geq 1$$

S_k is the number of transactional data in which the k$^{\text{th}}$ event occur. We have

$$1_{(S_k=j)} = \begin{cases} 1, & \text{if the condition } S_k = j \text{ satisfies;} \\ 0, & \text{otherwise.} \end{cases}$$

S_k's represent the number transactional data in which the item present. If the j$^{\text{th}}$ data does not contain the item then

$$X_j = \sum_{k=1}^{l} 1_{(S_k=j)} = \sum 0 = 0$$

The probability of occurrence of the event at the j$^{\text{th}}$ transactional data is denoted by u_j. Therefore

$$u_j = P(X_j = 1); j \geq 1 \tag{1}$$

(1) can be expressed as

$$u_j = \sum_{k=1}^{j} P(S_k = j); j \geq 1 \tag{2}$$

$$\begin{aligned} E(X_j) &= 1P(X_j = 1) + 0P(X_j = 0) \\ &= P(X_j = 1) \\ &= u_j \end{aligned}$$

$$i.e.,\ u_j = E(X_j), \quad j = 1, 2, \ldots$$

Take $u_0 = 1$. The occurrence of the event among the data can be derived using the generating function.

$$U_{(z)} = u_0 + u_1 z + u_2 z^2 + \cdots = \sum_{j=0}^{\infty} u_i z^i \tag{A}$$

p_k be the probability of the order of event. That is from k consecutive data the item A contains only in the last data. Its generating function is

$$P_{(z)} = \sum_{k=0}^{\infty} p_k z^k \tag{B}$$

We have $p_0 = 0$, A & B are related by

$$U(z) = 1 + P(z)U(z)$$

The following recursion equation can be used to calculate u_k.

$$u_n = \sum_{k=1}^{n} p_k u_{n-k} \quad (n = 1, u_0 = 1) \tag{3}$$

The mean and the factorial second moments can be expressed in terms of the sequences u_j as

$$\begin{aligned} E(N_m) &= E(X_1) + \ldots + E(Xm) \\ &= u_1 + u_2 + \ldots + u_m \\ &= \sum_{j=1}^{m} u_j \end{aligned} \tag{4}$$

$N_m =$ be the total transactional data containing the item A among the first m transactional data. The value of X_j is either 1 or 0. Therefore

$$X_j^2 = X_j \ \ \forall\ j$$

$$\begin{aligned} (N_m \ \ (N_m - 1)) &= N_m^2 - N_m \\ &= (X_1 + X_2 + \cdots + X_m)^2 - (X_1 + X_2 + \cdots + X_m) \\ &= 2 \sum_{i \neq j} X_i X_j \end{aligned}$$

From the above idea we can calculate the second moment of the number of data containing the item as

$$E\left(N_m^2\right) = EN_m\left(N_m - 1\right) + E\left(N_m\right)$$

The renewal rate or the rate of data occur which contain our item can be directly obtained from the generating function of the distribution of the number of occurrence of the event. Then we can compute the moments.

4. Conclusion

Renewal theory and Data mining, both the topics have great research opportunities. Occurrence rate of an event is great information about the collection of data. Also from this we can predict the nature of the data contain a particular event.

References

[1] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, Second Edition, San Diego: Academic Press, (1975).

[2] R. Rackwitz, *Optimizing systematically renewed structures*, Reliability Engineering and System Safety, 73(3)(2001), 269-279.

[3] Stephen Breen, Michael S. Waterman and Ning Zhang, *Renewal Theory for Several Patterns*, J. Appl. Prob., 22(1985), 228-234.

[4] J. A. M. Van der Weide, M. D. Pandey and J. M. Van Noortwijk, *A Conceptual Interpretation of the Renewal Theorem with applications*, Risk, Reliability and Societal Safety-Aven & Vinnen (eds), Taylor & Francis group, London, (2007).

[5] J. M. Van Noortwijk and H. E. Klatter, *The use of lifetime distributions in bridge maintenance and replacement modeling*, Computers and Structures, 82(13-14)(2004), 1091-1099.