Jñānābha, Vol. 30, 2000

MATHEMATICAL MODELS ON EPIDEMIOLOGICAL PROBLEMS WITH SPECIAL REFERENCE TO AIDS AND ITS REDUCTION STRATEGIES - USE OF COMPOUND BINOMIAL DISTRIBUTION.

By

S.S. Mishra, D.N. Pandey and A.S. Mishra

Department of Mathematics and Statistics, Dr. R.M.L. Avadh University, Faizabad-224001, U.P., India

(Received : February 15, 1999)

ABSTRACT

In this paper, an attempt has been made to make use of compound binomial distribution in dealing with problems of AIDS. Among methods include interval estimation, appoximation through Gaussian distribution and M.L.E. method etc. Finally, graphical approach has been used to demonstrate the comparative study of incidence infection rate corresponding to various sources of infection.

1. Introduction. When an outcome of the phenomenon is surrounded by uncertainties, it seems to be much difficult to quantify the amount of an outcome of that phenomenon. In order to effectively deal with probalilistic uncertainty, probability distribution theory has been developped to solve the above problem. Broadly categorised in two types-discret distribution and continuous distribution theories have tremendous potential for application in various wide ranging areas such as physical sciences, life sciences, epidemiology, management business, commerce, psychology and sociology etc. In 1974, Bachev and Petkova [1] have presented a mathematico-statistical analysis of the cycle recurrence in the droplet infections of communicable diseases. They have used dispersion analysis in the study of cycle recurrence or cyclicity.

In 1978, similarly, Bender [2] has conceptualised mathematical modelling in the areas of sex preference and sex ratio and problem of choices etc. Nevertheless, he has also studied Monte Carlo simulation technique which are applicable to a doctor's waiting room and sediment volume etc. Probabilistic model has also been used by him in the study of radio active decay, optimal facility location and distribution of particle size etc. Now a days, epidemiological problems have seriously attracted the attention of scientists involved therein. Different types of problems connected with epidemiology have been modelled mathematically and statistically.

In 1989, Edwards and Hamson [3] have also presented mathematical (statistical) modelling based on random variables. They have suggested various discrete and continuous random variables and their corresponding distributions to model different scientific and real life situations. They have especially discussed binomial, Poisson, exponential and normal distributions and their applications to various real life situations. Their approach to statistical modelling based on random varibles has undoubtedly added a new dimension to its study. Further more, in 1994 Kapur [5] has also described various aspects of mathematical modelling and its application in differnt areas such as physical and life sciences including epidemiology.

Under the category of epidemiological problems, recently, it is none but a problem of AIDS which has shaken the root of developed and developing societies and consequently gripped so badly that it has posed a great challenge before the epidemiologists. Unless, we effectively quantify the seriousness of disease, we cannot adopt any efficient reduction and prevention strategies to contain the HIV infection. Various mathematical modellings carried out previously could not have been able to precisely revealed such a fact which could be instrumental to adopting any effective reduction strategy. Recently Kotia and Srivastava [4] have made an effort to analyse impact of various sources of infection in order to seriously understand the problems of AIDS, so that infection rate could be contained effectively. Their study has made a hallmark in the investigations of problems of AIDS based on various sources of infection in the Indian perspective. Of late, Flanders, W.D. and Kleinbaum, D.G. [6] have presented a study on basic models for disease occurrence in Epidemiology. They have analysed disease occurrence with special reference to smoking deaths due to lung cancer in UK. More recently, Misra [7] has made an attempt in analysing the seriousness of AIDS infection by making use of Poisson distribution. However, his analytical approach could remain quite theoretical and couldnot reveal the process of data-fit without which purpose of investigation could not be fulfilled completely.

In the present paper an effort has been made by us to make use of compound binomial distribution in dealing with problems of AIDS. Here, it is a natural inquisitiveness that if any infected person of HIV transmits his or her HIV in several other persons with certain amount of probability along with the probability with which an individual infection is developing then what will be the probability or probabilitic rate that how many persons with how much probability would be infected in future out of total number of infected persons due to an individual HIV positive patient. Here we have been contemplating this situation of AIDS infection by making use of compound binomial distribution. Here the present paper is exclusively devoted to exhibit various analytical aspects such as theoretical, numerical and graphical in order to reach some index as a probabilistic rate of occurrence of AIDS infection which may be subsequently found helpful in formulating reduction as well as prevention strategies. Among methods include interval estimation, approximation through Gaussian distribution and MLE method etc. Finally, graphical aproach has been used to demonstrate the comparative study of incidence infection rate corresponding to various sources of infection.

2. Assumptions and Preliminaries. Before we use the compound binomial distribution in statistically analysing the problems of AIDS we have some underlined assumptions under which bionomial distribution can be applied in the afore said analysis :-

(i) AIDS disease can Independently occur among the group of people.

(ii) Risk of desease is equal with the change of time.

(iii) The occurrence of disease is rare and random in its beginning but subsequently it can cover the globly for the given geographical region.

Here we define compound binomial distribution as follows :

Let us suppose that $X_1, X_2, X_3, \ldots$ are indentically and independently distributed Bernouli variates with $P[X=0]=p$ and $P[X=1]=q=1-p$. For a fixed n, the random variable $X = X_1 + X_2 + X_3 + \ldots + X_n$ is a binomial variate with parameters n and p and probability function :

$$p\,(x=r) = \binom{n}{r} p^r q^{n-r},\ r = 0,1,2, \ldots, n.$$

which gives the probability of r successes in n independent trials with constant probability p of success for each trial.

Now suppose that n, instead of being regarded as a fixed constant, is viewed as a random variable following Poisson law with Parameter λ.Then,

$$P\,(n=k) = \frac{e^{-\lambda}\lambda^k}{k!}\ ;\ k = 0,\ 1,2, \ldots$$

In such a case x is said to have compound binomial distribution. The joint probability funciton of x and n is given by

$$P\,(X=r \cap n=k) = P\,(n=k)\,P\,(x=\frac{r}{n}=k) = \frac{e^{-\lambda}\lambda^k}{r} - \binom{k}{r} p^r q^{k-r}.$$

Since $P\,(X=r/n=k)$ is the probability of r successes in k trials. Obviously

$$r < k \Rightarrow k > r.$$

The marginal distribution of x is given by :

$$P(x = r) = \sum_{k=r}^{\infty} P(X = r \cap n = k)$$

$$= e^{-\lambda} p^r \sum_{k=r}^{\infty} \binom{k}{r} \frac{\lambda^k q^{k-r}}{k!}$$

$$= \frac{e^{-\lambda}(\lambda p)^r}{k!} \sum_{k=r}^{\infty} \frac{(\lambda p)^{k-r}}{(k-r)!}$$

$$= \frac{e^{-\lambda}(\lambda p)^r}{k!} \sum_{j=0}^{\infty} \frac{(\lambda p)^j}{(k-r)!} ; j = k-r.$$

$$= \frac{e^{-\lambda}(\lambda p)^r}{k!} e^{\lambda q},$$

which is a probability function of a Poission Variate with parameter λp. Hence $E(x) = \lambda p$ and variate $(x) = \lambda p$.

Now we estimate λp by making use of maximum likelihood estimation *(MLE)* method these after use interprete it in the term of disease incidence after relating it to the Person-time.

3. Statstical Analysis. Source Wise Appromimation of Probabilities by Gaussian Distribution for 95% Confidence Limit. If *MLE* of λp is C then incidence rate of the disease is represented by C/PT. Here it may be of interest to note that if group of cases is smaller the value of probability can be estimated accurately but when group of cases becomes larger then it can be approximated with the help of Gaussian distribution which is shortly demonstrated in the next discussion.

We know that there are several sources of infection through which *HIV* of *AIDS* are spreading in the society. Here we have been made available the statistics corresponding to each and every source of infection from different sero surviellance centres established in various corner of India. This is being displayed by table given as under.

3.1 Table

S.No.	*Sources of Infection*	*No. of Cases*
1.	Sexual Promiscuous	4483
2.	Blood Donors	1682
3.	Recipients of blood	209
4.	Drug users	1647
5.	Others	2247

Source : Ministry of Health, India, 1994

1. **Sexual Promiscuous**

   $4363 \leq x \leq 4603$

   $0.0194646 < r < 0.0205353$

2. **Blood Donors**

   $1602 \leq x \leq 1762$

   $0.0158739 \leq r \leq 0.0174593$

3. **Recipients of blood**

   $181 \leq x \leq 237$

   $0.0133235 \leq r \leq 0.0174457$

4. **Drug Users**

   $1567 \leq x \leq 1727$

   $0.0190285 \leq r \leq 0.0209969$

5. **Others**

   $2157 \leq x \leq 2337$

   $0.0159991 \leq r \leq 0.0173342$

3.2 Table

S.No.	Sources of Infection	Rate of Occurrence
1.	Sexual Promiscuous	$0.0194646 \leq r \leq 0.0205353$
2.	Blood Donors	$0.0158739 \leq r \leq 0.0174593$
3.	Recipients of blood	$0.0133235 \leq r \leq 0.0174457$
4.	Drug users	$0.0190285 \leq r \leq 0.0209969$
5.	Others	$0.0159991 \leq r \leq 0.0173342$

4. Graphical Representation.

Here we plot a graph corresponding to each and every source of infection and subsequently we conduct the comparative study of probabilistic occurrence of this disease. From graph we can approximate the following informations :

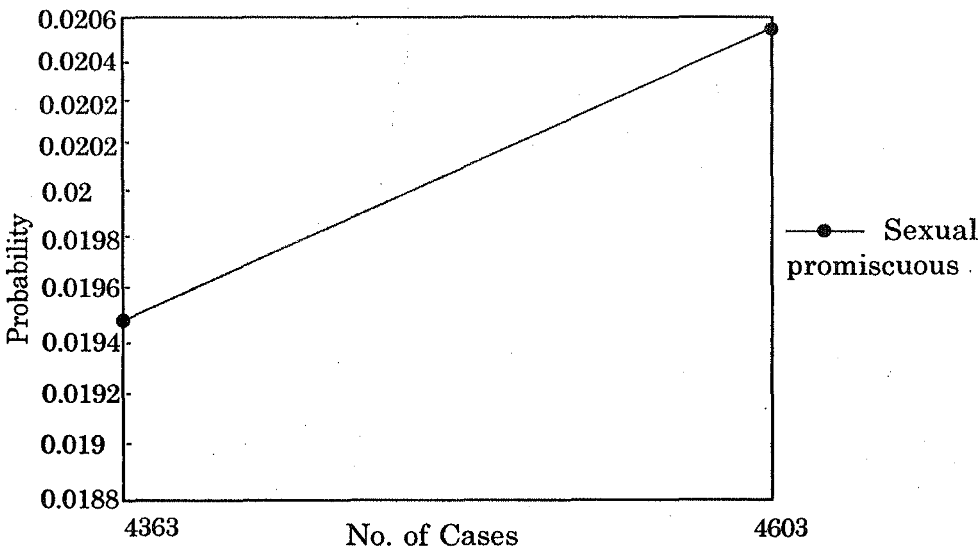
1. From graphical representation it is clear that at $X = 0$ probability becomes zero in all cases.
2. The behaviour of graph is of type $r = r_0\ ekx$ i.e. as value of x increases the value of r also increases without limit unless whole population gets infected by this disease.
3. If any successful cure of prevention srtategy is invoked or adopted the trend of the graph may change reverse.
4. This graph resembles with a famous exponential growth and will continue so long as any curve is not found.

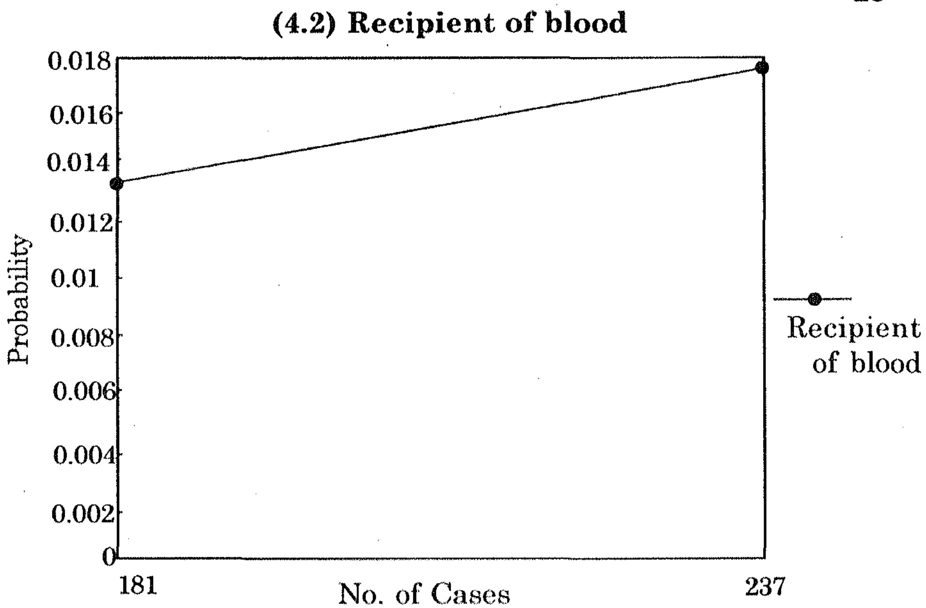
5. Chances of occurrence of disease through recipients of blood source infected are being reflected higher than other sources of infection of a given number of cases.
6. 5% increase in the case will result in an increase of 0.0001 in case of sexual promiscuous, 0.0001 in case of blood donors, 0.0006 in case of recipients of blood, 0.0002 in case of drug users, 0.0001 in case of others.

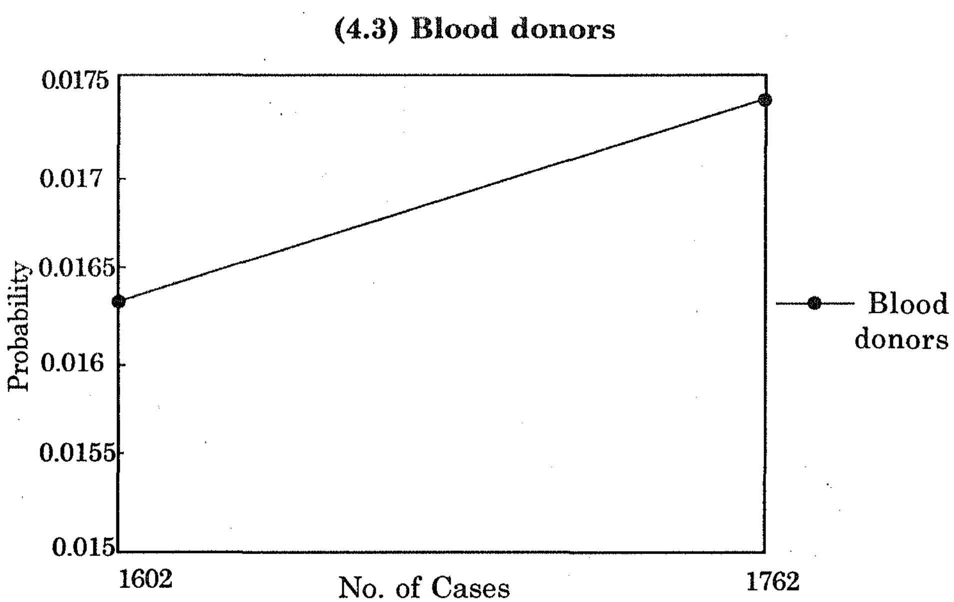
In view of above facts, reduction and prevention strategies may be chalked out efficiently and we may bring a better level of awareness among the people to reduce the infection rate. In the lack any vaccine or cure, we think preventive strategy is only sought cure which can be formulated effectively with the above set of informations.

As we have already mentioned that in case of larger group, probabilities are approximated through Gaussiam distribution with certain degree of confidence limit 95% and we describe the approximation corresponding to each and every source of infection as stated in the above table.

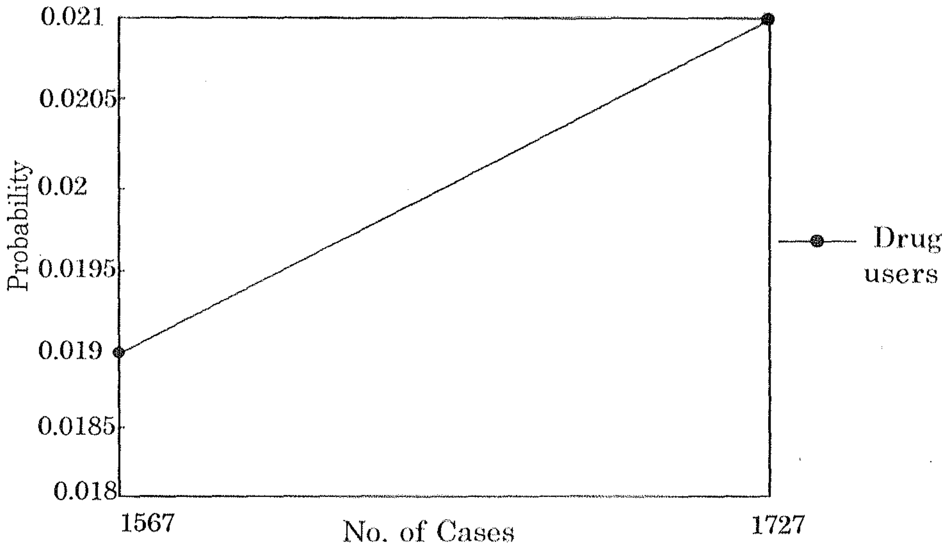
(4.1) Sexual Promiscuous



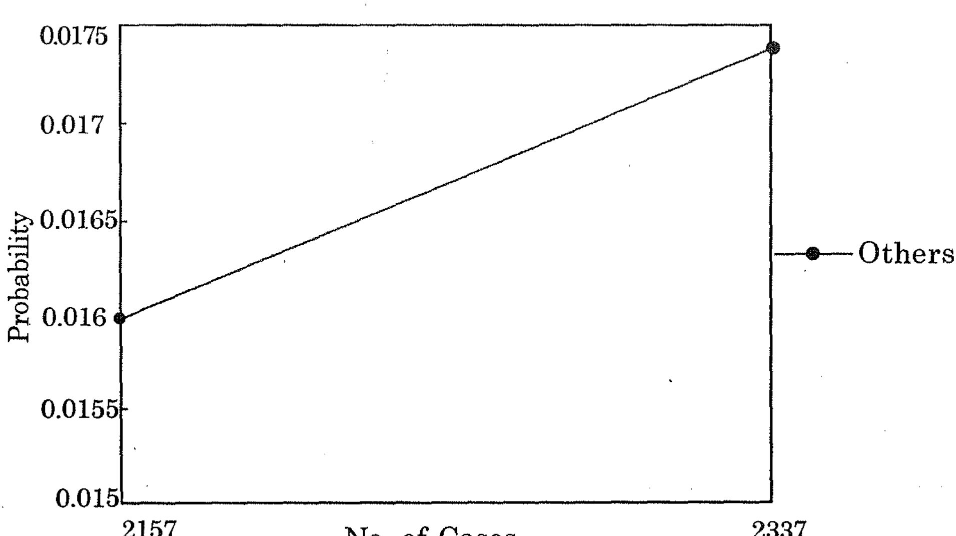

(4.2) Recipient of blood
0.018
0.016
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
Probability
181
No. of Cases
237
Recipient of blood


(4.3) Blood donors
0.0175
0.017
0.0165
0.016
0.0155
0.015
Probability
1602
No. of Cases
1762
Blood donors

(4.4) Drug users



(4.5) Others



5. Conclusion.

We can easily conclude that we have quantified the probabilitic rate of ocurrence of disease (*AIDS*) which is proven as an instrument to make the people understand and get acquainted at which rate *AIDS* is spreading in the society through different sources of infection and how intervariance is taking place dure to different sources of infections. This thing is a sole idea behind adopting the redution and prevention strategies in order to contain it. Nevertheless compound binomial distribution can also be applied in the study of various wide ranging problems as cancer, cardiovasenlar disease and Hepatitis *B* etc.

REFERENCES

[1] S. Bachev and Petkova, *Mathematico Statistical Analysis of the Cyclic Recurrence in the Droplet Infections of Communicable Diseases.* North Holland Publishing Company 1974.

[2] A. Bender Edward, *An Introduction to Mathematical Modelling*, A Wiley Interscience Publication John Wiley and Sons, New York, 1978.

[3] D. Edward and M. Hamson, *Mathematical Modelling*, Macmillan Education Ltd. Hound Mills, Basing stoke, Hampshire, London, 1989.

[4] J.N. Kapur, *Mathematical Modelling*, H.S. Poplai for Wiley Eastern Limited, 4835/ 24, Ansari Road Daryaganj, New Delhi, 1994.

[5] A. Kotia and G. Srivastava, *AIDS Diagnosis Treatment and Prevention*, S.K. Jain Publishers and Distributiors 485, Jain Bhawan, B.N. Bhawan, Shahadra, Delhi, 1994.

[6] W. Dana Flanders and G.K. David, Basic Models for disease occurrence in Epidemiology, *International Journal of Epidemiology*, **24 No. 1**, (1995) 1-7.

[7] S.S. Misra (1996), An estimation of disease occurrence in Epidemilogy with special refefrence to *AIDS*, *Vij. Par. Anu. Pat.*, **393** (1996), 161-163.