Journal of the Nigerian Association of Mathematical Physics
Volume **25 (November, 2013), pp 281 – 286**

Probability Generating Function and Epidemology

Osaro Igbinosa

**Department of Mathematics,
University of Benin, Nigeria**

Abstract

A general Urn model in probability theory is used to handle an illustrative practical problem on the spread of the Human Immune Deficiency Virus (HIV) among Intravenous Drug Users (IVDU) who commonly used hypodermic needles. A probability generating function (pgf) is obtained from which a transition matrix is derived and subsequently used to obtain the expected number of (IVDU) that could be infected with HIV virus when a certain number of heterogeneous hypodermic needles are in circulation.

***Keywords*:** tuberculosis, homotopy analysis method, nonlinear equations, mathematical model, epidemics

1.0 Introduction

The classical random allocation model arises quite naturally in the context of needle sharing among intravenous drug users (IVDU) and was developed to model the growth of infective IVDUs [1 – 7]. Here, it is assumed that *i* hypodermic needles are used by IVDU who could be infected with a virus such as hepatitis or HIV. These needles are randomly shared among n susceptible IVDU infecting $1 \le s \le \min(i,n)$ of them. Gani and Yakowitz [7] studied the transmission of HIV in a group of IVDU which exchange needles repeatedly.

The model turns out to be identical to an Urn model with stochastic replacement considered earlier by Woodbury and Rutherford mentioned in Johnson et al [8]. In the Woodbury model, we have an Urn initially containing n black balls. Balls are drawn at random from the Urn, if a black ball is drawn it is replaced by a white ball while if a white ball is drawn it is returned directly to the Urn.

It follows that immediately after the (i+1)th trial $i \ge 0$ there will be s white balls , if at the ith trial there were s white balls and a white ball was drawn with probability s/n or if there were s-1 white balls and a black ball was drawn with probability (1-$\frac{s-1}{n}$) and replaced by a white one . Writing

p_s (I,n)=p{s white balls at trial i/n black balls initially}

We have the following recursive equation for the probability

$$p_s(i+1,n) = p_s(I,n)\, s/n + p_{s-1}(I,n)\left(1-\frac{s-1}{n}\right) \quad (1)$$

With s = 1, 2, , , , , , , , , , , , min(i+1,n) and

$$p_s(0,n) = \delta_{s,0}$$

In general the recursive equation for Urn model is given by

$$p_s(i+1,n) = p_s(I,n)\, h_s + p_{s-1}(i,n)\, g_{s-1} \quad (2)$$

Where $h_s + g_s = 1$

Corresponding author: ***Osaro Igbinosa,*** E-mail: osaroigbinosa@Gmail.com, Tel.: +2348036333456

Rutherford [9] studied the case where

$$h_s = p + cs$$

With $0 \angle c \leq \frac{1-p}{n}$ in which h_s may now be a simple linear function of s

Recall (2)

$$p_s(i+1,n) = p_s(i+1,n)h_s + p_{s-1}(i,n)g_{s-1} \quad (3)$$

Let

$$h_s = p + cs$$

And $g_{s-1} = q - c(s-1)$

Where 1-p=q

Substituting (3) and (4) in (2) we have

$$p_s(i+1,n) = p_s(i,n)[p+cs] + p_{s-1}(i,n)[q-c(s-1)] \quad (4)$$

2.0 Derivation of the Transition Matrix from the Probability Generating Function (PGF)

The simplest way to derive with this recursive algorithm (4) is through the probability generating function (pgf) $f_{i,n}(u)$ of the probabilities $p_s(i,u)$

i.e

$$f_{i,n}(u) = \sum_{s=0}^{\min(i,n)} p_s(i,n)u^s \quad (5)$$

And

$$f_{i+1,n}(u) = \sum_{s=0}^{\min(i+1,n)} p_s(i+1,n)u^s \quad (6)$$

We write (04) in the form

$$\sum_{s=0}^{\min(i+1,n)} p_s(i+1,n)u^s = \sum_{s=0}^{\min(i,n)} p_s(i,n)[p+cs]u^s + \sum_{s=0}^{\min(i,n)} p_{s-1}(i,n)[q-c(s-1)]u^s$$

This reduces to the form

$$f_{i+1.n}(u) = p\sum_{s=0}^{\min(i,n)} p_s(i,n)u^s + c\sum_{s=0}^{\min(i,n)} p_s(i,n)su^s + u\{q\sum_{s=1}^{\min(i,n)} p_{s-1}(i,n)u^{s-1} - c\sum p_{s-i}(i,n)[s-1]u^{s-1}\} \quad (7)$$

Substituting (5) and (6) in (7) yields

$$f_{i+1,n}(u) = p\ f_{i,n}(u) + cu\frac{\partial f_{i,n}(u)}{\partial u} + uq\ f_{i,n}(u) - u^2 c\frac{\partial f_{i,n}(u)}{\partial u}$$

This reduces to the form

$$f_{i+1,n}(u) = (p+qu)\ f_{i,n}(u) + cu(1-u)\frac{\partial f_{i,n}(u)}{\partial u} \quad (8)$$

where

$$\text{u}\frac{\partial f_{i,n}(u)}{\partial u} = \sum_{s=0}^{\min(i,n)} p_s(i,n)su^s$$

and $u\dfrac{\partial f_{i,n}(u)}{\partial u} = \sum_{s=1}^{\min(i,n)} p_{s-1}(i,n)[s-1]u^{s-1}$ Now we show (8) such that

$$f_{0,n}(u) = 1$$

$$f_{1,n}(u) = p + qu$$

$$f_{2,n}(u) = (p+qu)^2 + cqu(1-u)$$

$$f_{3,n}(u) = (p+qu)^3 + cqu(1-u)(p+qu) + cu(1-u)[2q(p+qu) + cq(1-2u)]$$

Respectively.To seek a recursive equation for the coefficients of the various powers of u.
we write

$$f_{i,n}(u) = a_0(i) + a_1(i)u + a_2(i)u^2 + \ldots\ldots + a_i(i)u^i \text{ For } n \geq i \tag{9}$$

inserting (9) in (8) we obtain

$$f_{i+1,n}(u) = (p+qu)[a_0(i) + a_1(i)u + a_2(i)u^2 + \ldots\ldots + a_i(i)u^i] + cu(1-u)$$
$$[a_1(i) + 2a_2(i)u + \ldots\ldots + ia_i(i)u^{i-1}] \tag{10}$$

Therefore

$$f_{i+1,n}(u) = (p+qu)a_0(i) + [(p+c)u + (q-c)u^2]a_1(i) +$$
$$[(p+2c)u^2 + (q-2c)u^3]a_2(i) + \ldots\ldots + [(p+ic)u^i + (q-ic)u^{i+1}]a_i(i) \tag{11}$$

Thus we put the coefficient of u in (10) in matrix form as follows

$$\text{A(i)} = \begin{bmatrix} a_0(i) \\ a_1(i) \\ a_2(i) \\ \cdot \\ \cdot \\ \cdot \\ a_{i-1}(i) \\ a_i(i) \end{bmatrix} = \begin{bmatrix} p & & & & & \\ q & p+c & & & & \\ & q-c & p+2c & & & \\ & & & p+(i-1)c & & \\ & & & q-(i-2)c & q-(i-1)c & p+ic \end{bmatrix} \begin{bmatrix} a_0(i-1) \\ a_1(i-1) \\ a_2(i-1) \\ \cdot \\ \cdot \\ \cdot \\ a_{i-1}(i-1) \\ 0 \end{bmatrix} \tag{12}$$

Setting p=0, q=1 and c= $\frac{1}{n}$ in (12) yields

$$\text{A(i)} = \begin{bmatrix} a_0(i) \\ a_1(i) \\ a_2(i) \\ \cdot \\ \cdot \\ \cdot \\ a_{i-1}(i) \\ a_i(i) \end{bmatrix} = \begin{bmatrix} 0 & & & & \\ 1 & \frac{1}{n} & & & \\ & 1-\frac{1}{n} & \frac{2}{n} & & \\ & & 1-\frac{(i-2)}{n} & 1-\frac{(i-1)}{n} & \\ & & & \frac{(i-1)}{n} & \frac{i}{n} \end{bmatrix} \begin{bmatrix} a_0(i) \\ a_1(i) \\ a_2(i-1) \\ \cdot \\ \cdot \\ \cdot \\ a_{i-1}(i-1) \\ 0 \end{bmatrix} \tag{13}$$

Equations (12) and (13) can be further expressed in matrix notation as

$$A(i) = B_i A(i-1) \quad (14)$$

The matrix B_i in equation (12) and ((13) are transition matrices of a markov chain process and

In equation (10) the pgf $f_{i,n}(u)$ written in power series form has coefficients $a_i(i)$ as

$$a(i) = q(q-c) \ldots\ldots\ldots\ldots\ldots \text{ [q-(i-1) c].} \quad (15a)$$

See Gani [6].This means that if n<I , then q must equal one of 0,c,….(I – 1)c for the distribution to be honest . It shows that the pgf $f_{i,n}(u)$ in (9) can be determined from (14) using the markov chain transition matrix of equation (11) which corresponds to the pgf $f_{i,+1,n}(u)$.

Consequently, Gani [6] obtained the expectation of $f_{i,n}(u)$ as

$$E(f_{i,n}) = n\left[1-\left(1-\frac{1}{n}\right)^{i}\right] \quad (15b)$$

Equation (15b) gives the expected number of people that will contact the HIV virus among the n susceptible IVDUs when i hypodermic needles are in use.In addition, we also derive (15b) absolutely using probability as follows:

The probability of any susceptible among n IVDUs contacting the HIV Virus is $\frac{1}{n}$ at each time an infested needle is in circulation. The probability of not contacting the virus at each time a needle is in use is $(1-\frac{1}{n})$. The probability of not contacting the virus when i infested needles are in circulation among the n IVDUs is $(1-\frac{1}{n})$. Then the probability of contacting the virus when i needles are in use is given by $[1-(1-\frac{1}{n})^{i}]$.Therefore the expected number of the IVDUs to be infected when i needles are in use among them is

$$E(n) = n\left[1-\left(1-\frac{1}{n}\right)^{i}\right] \quad (16)$$

Equation (16) is exactly the same as the one obtained by Gani [6] expressed in (15a) and (15b) .

3.0 Two Types Of Infectives

Suppose we now have two types of infective j and k from (16) then

$$E(n) = n\left[1-\left(1-\frac{1}{n}\right)^{j+k}\right] \quad (17)$$

Where i = j+k. Suppose j = cn – 1, k = 1, where c is some constant, then

$$E(cn-1, 1) = n\left[1-\left(1-\frac{c}{j+1}\right)^{j+1}\right] \quad (18)$$

$$E(j+1) = n\left[1-\left(1-\frac{c}{j+1}\right)^{j+1}\right] \quad (19)$$

$$E(j+1) = n\left(1-\ell^{-c}\right)$$

And

$$Var(j+1) = n\,\ell^{-c}\left(1-\ell^{-c}\right) \quad (20)$$

See Gani [6]. Recall Taylor series expansion $\ell^{Rt} \approx (1+\frac{Rt}{n})^{n}$

Even a small value of k, such as k = 1 makes a difference to the number of new type 1 infective after a single needle exchange.

We try to deduce the risk sensitivity of the above distribution (21). To deduce the minimum harmful consequence of risk we choose the outcome with mean/standard deviation pair (μ,σ) on the line of highest possible slope,

$$\mu = R_{\min} + \sigma Z,$$

where Z denotes z-score, see [1].

4.0 Numerical Illustration

In this section, we carry out some numerical experiments to demonstrate the application of the algorithm discussed in this paper.Table 1.1 shows the results of our experiments.

Table 1.1: Number of Needles and Expected Infectives, Variances and Standard Deviations

j	$M(j+1)$	$Var(j+1)$	STD
0	0.9	0.819	0.905
1	1.8	1.476	1.215
2	2.6	1.924	1.387
3	3.3	2.211	1.487
4	3.9	2.379	1.542
5	4.5	2.475	1.573
6	5	2.500	1.581
7	5.5	2.475	1.573
8	5.9	2.419	1.555
9	6.3	2.331	1.527
10	6.7	2.211	1.487

In Table 1.1 when j needles $(0 \le j \le 10)$ are in use with one infective the expected number of infective for 10 IVDUs is computed using equation (20). See Table 1.1.

From Table 1.1 it is clear that with each increase in the number of needles the expected number of infective increases while the standard deviation σ increases to a point then start decreasing. To deduce the risk sensitivity of this model we use the regression analysis and found that the equation

$$y = -6.2508 + 7.2692x$$
$$(2.7040), (1.8601) \qquad (21)$$

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-6.2508	2.70400	-2.3117	0.0461	-12.3677	-0.13392
X	7.2692	1.8601	3.9079	0.0035	3.06134	11.4771

The computed coefficient of variation from Table 1.1 is $R^2 = 0.6292$. Equation (22) could be expressed as

$$7.2692 = \frac{(\mu - (-6.2508))}{\sigma} \qquad (23)$$

Where $y = \mu$ and $x = \sigma$. Thus,equation (23) minimizes the harmful consequence of risk. Hence, we choose the outcome(μ, σ) = (5.5, 1.573) pair with $j = 7$.

5.0 Interpretation

Table 1.2 shows (a) that since the confidence interval $(-12.3677, -0.13392)$ includes zero, the hypothesis that this parameter -6.2508 (intercept) is zero cannot be rejected at 0.0461 significance level, it implies the intercept is essentially zero. (b) The confidence interval $(3.06134, 11.4771)$ does not include zero, then the slope 7.2692 is significantly different from zero and cannot be rejected at 0.0035 significant level, the slope is essentially not zero.

The coefficient of variation R^2 shows the predictive strength of the model (22) which is 63% .

6.0 Conclusion

A probability generating function (pgf) based on $(i+1)$ hypodermic needles of the probabilities of infective was obtained and used to obtain a transition matrix of a markov chain. Consequently, this is used to determine the coefficients of the pgf of i hypodermic needles when this pgf is expressed in power series. If these coefficients are known it implies the pgf of i hypodermic needles is known, its expectation and standard deviation can be deduced for each number of hypodermic needles in use and the number of IVDUs. A sensitivity analysis is carried out on the model. Here we deduce that at least seven ($j = 7$) hypodermic needles should be in circulation to reduce the risk of contracting HIV whenthere are $n = 10$ IVDUs in the system, since 61 percent variation in the model is due to the standard deviation (risk).

References

[1] Bruce Winterhalder, Flora Lu, and Bram Tucker(1999) “Risk – Sensitive Adaptive Tactics. “Journal of Archaeological Research Vol. 7 No. 4., 301-348.

[2] Gani J.(1991) “ Generating Function Methods in a Random Allocation Problem of Epidemics”. Bull. Inst. Comb. Appl. 3, 43-50.

[3] Gani, J. (1993) “Random Allocation Methods in an Epidemic Model in Stochastic Processes”

[4] Gani J (2002a) “IVDUs.”Needle Sharing Infections among Heterogeneous Monatsh Math. 135, 25 – 36.

[5] Gani, J.(2002b) “Nonhomogeneous Susceptible and Infectives Among Needle Sharing IVDUs.”World Scientific, Singapore, 91 – 99.

[6] Gani, J.(2003)” Random Allocation and Urn Models. Available Online at : Http//www.maths. Anu.Edu.au.Accessed on 27th July 27, 2013.

[7] Gani, J. and Yakowitz, S. (1993)“Modeling the Spread of HIV among Intravenous Drug Users. “MA J. Math’s. Appl. Med and Biology 10, 51 – 65.

[8] Johnson, N.L. Kotz, S. and Kemp, A.W. (1993) “Univariate Discrete Distributions”. 2nd Edition New York: Wiley Inc.

[9] Rutherford, R.D. (1954) “On a Contagious Distribution.”Ann.Math.Statist. 25, 703-713.