International Journal of Applied Mathematical Sciences.
ISSN 0973-0176 Vol.2 No.2 (2005), pp. 220-229

http://www.gbspublisher.com/ijams.htm

Allocating Infected Needles Randomly among Susceptibles

J. Gani
Mathematical Sciences Institute
Australian National University
Canberra ACT 0200, Australia
E-mail: gani@maths.anu.edu.au

R. J. Swift
Department of Mathematics and Statistics
California State Polytechnic University, Pomona
Pomona, CA 91768
E-mail: rjswift@csupomona.edu

Abstract

This paper considers the random allocation of i infected needles among a fixed number n of susceptibles. The expected numbers of susceptibles hit by $k = 1, 2, \ldots, i$ needles are derived, and the numbers of 1-hit and 2 or more-hits infected susceptibles are compared. Finally, we discuss the case where infection is transmitted by each needle independently with a probability $p < 1$.
Keywords: Random allocation, infectives, susceptibles, probability generating function (pgf)

AMS Subject Classification: 60G99, 92D30

1. Introduction

A model which has often been used for the spread of infection among susceptibles through hypodermic needles is the random allocation model. In this model, it is assumed that i needles, infected with HIV, hepatitis or some other disease, are distributed randomly among n susceptibles. This is identical with the random allocation of i balls (needles) among n cells (susceptibles), as illustrated in Figure 1.

Let $S_0(i,n), S_1(i,n), \ldots, S_{j-1}(i,n)$ be the numbers of susceptibles subject to $0, 1, \ldots, j-1$ needle hits respectively, with $S_j(i,n)$ the number subject to $j \leq i$ or more needle hits. Then if

$$p_{s_0 s_1 \ldots s_j}(i,n) = P\{S_0(i,n) = s_0, S_1(i,n) = s_1, \ldots, S_j(i,n) = s_j\},$$

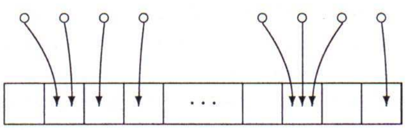

Figure 1: Random allocation of i balls in n cells

given that $S_0(0,n) = n$, and $S_k(0,n) = 0$ for all $k = 1,2,\ldots,j$, we can see that on increasing the number of needles from i to $i+1$, the following relation will hold

$$\begin{aligned} p_{s_0 s_1 \ldots s_j}(i,n) &= p_{s_0+1 s_1-1 \ldots s_j}(i-1,n)\left(\frac{s_0+1}{n}\right) \\ &+ \ldots \\ &+ p_{s_0 s_1 \ldots s_{j-1}+1, s_j-1}(i-1,n)\left(\frac{s_{j-1}+1}{n}\right) \\ &+ p_{s_0 s_1 \ldots s_j}(i-1,n)\left(1-\frac{s_0+\ldots+s_{j-1}}{n}\right). \end{aligned} \quad (1)$$

Defining the probability generating function (pgf)

$$F_{i,n}(u_0,u_1,\ldots,u_j) = \sum p_{s_0 s_1 \ldots s_j}(i,n) u_0^{s_0} \ldots u_j^{s_j},$$

where the sum is taken over all values of s_k, $k = 0,1,\ldots,j$, such that $s_1 + 2s_2 + \ldots + js_j = i$, and $s_0 + s_1 + \ldots + s_j = n$, we deduce from (1) the difference-differential equation

$$\begin{aligned} F_{i,n}(u_0,u_1,\ldots,u_j) &= \frac{1}{n}\left((u_1-u_0)\frac{\partial F_{i-1,n}}{\partial u_0} + \ldots + (u_k-u_{k-1})\frac{\partial F_{i-1,n}}{\partial u_{k-1}}\right. \\ &+ \left.\ldots + (u_j-u_{j-1})\frac{\partial F_{i-1,n}}{\partial u_{j-1}}\right) + F_{i-1,n}. \end{aligned} \quad (2)$$

Further details of such pgfs may be found in Gani [1], [2] and [3]. In this paper, we are specifically interested in the expected numbers

$$M_k(i,n) = E[S_k(i,n)], \ k = 0,1,\ldots,j,$$

and their asymptotic values as $n \to \infty$.

If we consider a single susceptible, and ask what the probability is that it receives $k \leq i$ needle hits, it is clear that

$$p_k = P(k \text{ needle hits on a susceptible }) = \binom{i}{k}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{i-k}.$$

Hence the expected number of susceptibles with $k < j$ needle hits will be

$$M_k(i,n) = n \binom{i}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{i-k},$$

while for $k \geq j$, we have

$$M_k(i,n) = n \sum_{k \geq j} p_k.$$

We now use an alternative method based on the pgf $F_{i,n}$ defined above.

2. Expectations

If we take the derivative of $F_{i,n}$ with respect to u_0 , we find that

$$\begin{aligned} \frac{\partial F_{i,n}}{\partial u_0} &= \frac{1}{n}\left(-\frac{\partial F_{i-1,n}}{\partial u_0} + (u_1 - u_0)\frac{\partial^2 F_{i-1,n}}{\partial u_0^2} + \ldots + \right. && (3) \\ &+ \left. (u_{k+1} - u_k)\frac{\partial^2 F_{i-1,n}}{\partial u_k \partial u_0} + \ldots + (u_j - u_{j-1})\frac{\partial^2 F_{i-1,n}}{\partial u_{j-1} \partial u_0}\right) \\ &+ \frac{\partial F_{i-1,n}}{\partial u_0}. && (4) \end{aligned}$$

Setting $u_0 = u_1 = \ldots = u_j = 1$, we see that

$$M_0(i,n) = (1 - \frac{1}{n})M_0(i-1,n), \tag{5}$$

or, since $M_0(0,n) = n$, then

$$M_0(i,n) = (1 - \frac{1}{n})^i n.$$

Similarly, for any u_k, $k = 1, 2, \ldots, j-1$, we find that

$$\begin{aligned} \frac{\partial F_{i,n}}{\partial u_k} &= \frac{1}{n}\left((u_1 - u_0)\frac{\partial^2 F_{i-1,n}}{\partial u_0 \partial u_k} + \ldots + (u_k - u_{k-1})\frac{\partial^2 F_{i-1,n}}{\partial u_{k-1} \partial u_k} + \frac{\partial F_{i-1,n}}{\partial u_{k-1}}\right. \\ &+ \left. (u_{k+1} - u_k)\frac{\partial^2 F_{i-1,n}}{\partial u_k^2} - \frac{\partial F_{i-1,n}}{\partial u_k} + \ldots + (u_j - u_{j-1})\frac{\partial^2 F_{i-1,n}}{\partial u_{j-1} \partial u_k}\right) \\ &+ \frac{\partial F_{i-1,n}}{\partial u_k}. && (6) \end{aligned}$$

Hence, it follows that

$$M_k(i,n) = (1 - \frac{1}{n})M_k(i-1,n) + \frac{1}{n}M_{k-1}(i-1,n). \tag{7}$$

For u_j , the derivative is

$$\begin{aligned}\frac{\partial F_{i,n}}{\partial u_j} &= \frac{1}{n}\left((u_1-u_0)\frac{\partial^2 F_{i-1,n}}{\partial u_0 \partial u_j}+\ldots+(u_j-u_{j-1})\frac{\partial^2 F_{i-1,n}}{\partial u_{j-1}\partial u_j}+\frac{\partial F_{i-1,n}}{\partial u_{j-1}}\right)\\ &+ \frac{\partial F_{i-1,n}}{\partial u_k},\end{aligned} \tag{8}$$

so that

$$M_j(i,n) = M_j(i-1,n)+\frac{1}{n}M_{j-1}(i-1,n). \tag{9}$$

We may express these results in matrix form as

$$\begin{bmatrix} M_j(i,n)\\ M_{j-1}(i,n)\\ \vdots\\ M_1(i,n)\\ M_0(i,n)\end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{n} & 0 & \ldots & 0 & 0\\ 0 & 1-\frac{1}{n} & \frac{1}{n} & 0 & \ldots & 0\\ 0 & 0 & 1-\frac{1}{n} & \frac{1}{n} & \ldots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & 0 & \ldots & 1-\frac{1}{n}\end{bmatrix}\begin{bmatrix} M_j(i-1,n)\\ M_{j-1}(i-1,n)\\ \vdots\\ M_1(i-1,n)\\ M_0(i-1,n)\end{bmatrix} \tag{10}$$

or

$$M(i,n) = A(n)M(i-1,n),$$

where $A(n)$ is a $(j+1)\times(j+1)$ matrix. If, for simplicity, we take $j=i$, we then have that $A(n)$ is an $(i+1)\times(i+1)$ matrix, and

$$M(i,n) = A(n)^i M(0,n), \tag{11}$$

where $M'(0,n) = [0,0,0,\ldots,0,n]$.

A matrix of the general form

$$\begin{bmatrix} 1 & b & 0 & \ldots & 0 & 0\\ 0 & a & b & 0\ldots & 0 & 0\\ 0 & 0 & a & b & \ldots & 0\\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots\\ 0 & 0 & 0 & 0 & \ldots & a\end{bmatrix}$$

where $b = 1/n$, and $a = (1-1/n)$ in $A(n)$, is such that its i-th power is

$$\begin{bmatrix} 1 & b(1+a+\ldots+a^{i-1}) & b^2(1+2a+\ldots+(i-1)a^{i-2}) & \ldots & b^{i-1}(1+(i-1)a) & b^i\\ 0 & a^i & ia^{i-1}b & \ldots & \frac{i(i-1)}{2!}a^2b^{i-1} & iab^{i-1}\\ 0 & 0 & a^i & \ldots & \frac{i!(i-3)!}{3!}a^3b^{i-3} & \frac{i(i-1)}{2!}a^2b^{i-2}\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \ldots & a^i\end{bmatrix}$$

with the last column consisting of the terms of the binomial $(b+a)^i$. It follows that

$$M'(i,n) = [n(1/n)^i, ni(1/n)^{i-1}(1-1/n), ..., ni(1/n)(1-1/n)^{i-1}, n(1-1/n)^i].$$

We note that the sum of these terms is n.

If n is very large, we can write the asymptotic expected values as

$$M_k(i,n) \sim \frac{n}{n^{i-k}} \frac{i! e^{-k/n}}{k!(i-k)!}$$

for $k = 0, 1, \ldots, i$. Suppose that two or more hits by an infected needle are necessary to transmit the infection, then the expected number of infected individuals will be

$$\begin{aligned} n[1-(1-1/n)^i - i(1/n)(1-1/n)^{i-1} &= n[1-(1-1/n)^{i-1}(1+(i-1)/n] \\ &\sim n[1-(1+(i-1)/n)e^{-(i-1)/n}]. \end{aligned}$$

As a simple example, let us take $n = 100$, and $i = 25$, then the expected number of susceptibles infected by two or more needle hits is

$$100[1-0.99^{24}(1.24)] = 2.5759,$$

while the asymptotic approximation is

$$100[1-(1.24)e^{-0.24}] = 2.4581.$$

The ratio of individuals infected by two or more needle hits to those suffering only a single hit is

$$\frac{n[1-(1-\frac{1}{n})^{i-1}(1+\frac{i-1}{n})]}{i(1-\frac{1}{n})^{i-1}}.$$

In the case where $n = 100$ and $i = 25$, this is

$$\frac{2.5729}{19.6420} = 0.131143,$$

while the asymptotic approximation gives the slightly lower value

$$\frac{2.4581}{19.6657} = 0.124994.$$

3. Infections due to 2 or More Needle hits

The expected number of individuals subject to a single needle hit rises to a maximum, and then decreases as the number i of needles increases, while the number of those subject to two or more-hits continues to rise. Table 1 lists the one-hit and two or more-hits individuals for increasing values of i when $n = 100$.

i	Single-hit individuals $i(1-\frac{1}{n})^{i-1}$	Two or more-hits individuals $n[1-(1-\frac{1}{n})^{i-1}(1+\frac{i-1}{n})]$
25	19.64	2.58
50	30.56	8.94
75	35.65	17.29
99	36.97	26.05
100	36.97	26.42
125	35.95	35.58
150	33.55	44.30
168	31.36	50.16
175	30.45	52.33
200	27.07	59.54
300	14.86	80.24
400	7.25	90.95
500	3.32	96.02
600	1.46	98.30
650	0.96	98.90
662	0.86	99.01
675	0.77	99.12

Table 1: One-hit and two or more-hits individuals, $n = 100$ and increasing i.

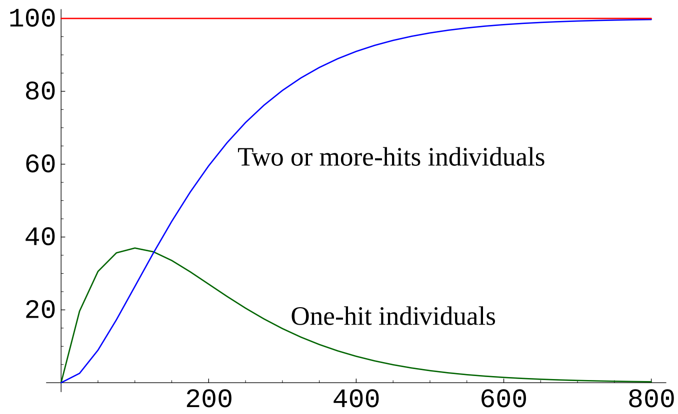

Figure 2: Graphs of one-hit and two or more-hits individuals for $n = 100$

We see that the expected number of one-hit individuals reaches a peak of 36.97 for $i = 99$ and 100, and then declines to less than 1 for $i = 650$. The expected number of two or more-hits individuals rises steadily, reaching over 50% of the population when $i = 168$, and over 99% when $i = 662$. The graphs in Figure 2 illustrate the situation.

The approximate asymptotic results for $i = 168$ and 662 give

$$\begin{aligned} n[1-(1+(i-1)/n)e^{-(i-1)/n}] &= 49.74 \text{ for } i = 168 \\ &= 98.98 \text{ for } i = 662 \end{aligned}$$

for two or more-hits individuals, slightly lower than the exact values of 50.16 and 99.01 respectively, and

$$\begin{aligned} ie^{-(i-1)/n} &= 31.63 \text{ for } i = 168 \\ &= 0.89 \text{ for } i = 662 \end{aligned}$$

for one-hit individuals, slightly larger than the exact values 31.36 and 0.86 respectively.

4. Needles with Random Infectivity

Suppose that infection is transmitted by each infected needle independently with the probability $0 < p < 1$; then $q = 1 - p$ is the probability of non-infection, and for a susceptible subjected to $1 < k < i$ needle hits, and we assume that at least $1 \leq j \leq k$ hits are required for the individual to be infected, then the probability of being infected will be

$$\sum_{x=j}^{k} \binom{k}{x} p^x q^{k-x} = 1 - \sum_{x=0}^{j-1} \binom{k}{x} p^x q^{k-x},$$

so since the expected number of individuals subjected to k needle hits is

$$n\left[\frac{i!(1-\frac{1}{n})^{i-k}}{k!(i-k)!n^k}\right], \tag{12}$$

then the number infected among individuals subject to k needle hits is

$$n\left[\frac{i!(1-\frac{1}{n})^{i-k}}{k!(i-k)!n^k}\right]\left(1-\sum_{x=0}^{j-1}\binom{k}{x}p^x q^{k-x}\right). \tag{13}$$

Summing over all values of k, the total number of infected individuals will be

$$\sum_k\left[n\left[\frac{i!(1-\frac{1}{n})^{i-k}}{k!(i-k)!n^k}\right]\left(1-\sum_{x=0}^{j-1}\binom{k}{x}p^x q^{k-x}\right)\right]. \tag{14}$$

Table 2 compares the expected numbers of infected individuals for $n = 100$, when infection is due to two or more needle hits with needles having random infectivity $p = 0.5$ requiring at least $j = 1$ or 2 hits to become infected.

i	Two or more-hits individuals $n[1-(1-\frac{1}{n})^{i-1}(1+\frac{i-1}{n})]$	Random Infectivity with at least j hits $j \geq 1$	$j \geq 2$
25	2.58	11.78	0.6948
50	8.944	22.169	2.613
75	17.291	31.336	5.457
100	26.424	39.423	8.982
125	35.581	46.558	12.988
150	44.302	52.852	17.314
175	52.327	58.405	21.827
200	59.535	63.304	26.424
300	80.235	77.771	44.259
400	90.952	86.534	59.467
500	96.025	91.843	71.347
600	98.302	95.059	80.160
650	98.900	96.154	83.592
700	99.290	97.007	86.477

Table 2: Expected numbers of infected individuals under three scenarios for $n = 100, p = 0.5$

We note that under random infectivity with $p = 0.5$, the expected number of infected individuals for $j \geq 1$ is larger than the two or more-hits individuals where $p = 1$ and $j \geq 2$ until $i = 251$, when it begins to lag, as can be seen in Figure 3.

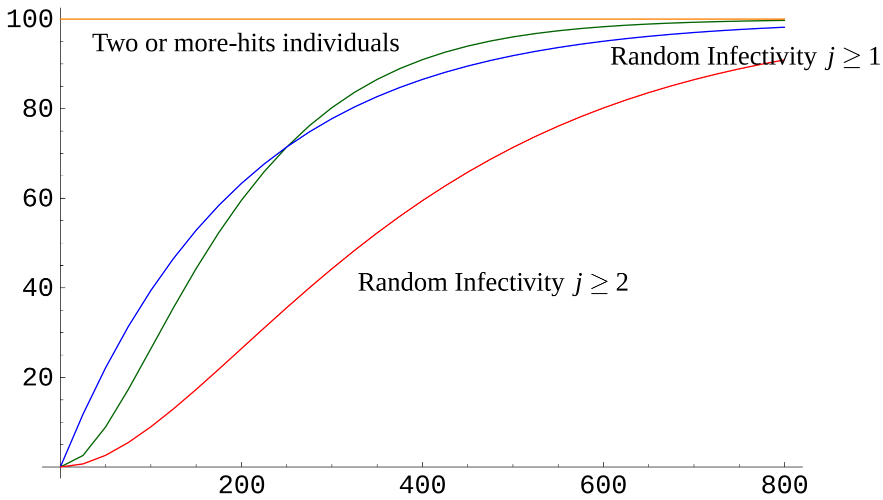

Figure 3: Graphs of expected numbers of infected individuals under three scenarios for $n = 100$

The turning point occurs generally when

$$1-(1-p/n)^i < 1-(1-1/n)^{i-1}(1+(i-1)/n),$$

or

$$(1-p/n)^i > (1-1/n)^{i-1}(1+(i-1)/n).$$

When $n = 100$ and $p = 0.5$, the turning point is at $i = 251$, as can readily be verified, since after taking logarithms for $i = 251$,

$$-1.2581 = 251\ln 0.995 > 250\ln 0.99 + \ln 3.50 = -1.2598$$

while for $i = 250$,

$$-1.2531 = 250\ln 0.995 < 249\ln 0.99 + \ln 3.49 = -1.2526.$$

Acknowledgements

The second author expresses his sincere thanks to The Centre for Mathematics and its Applications and the Mathematical Sciences Institute at The Australian National University for their kind generosity during his stay as a Visiting Fellow. He also wishes to thank the Department of Mathematics and Statistics and the College of Science at California State Polytechnic University, Pomona for providing travel funds. Without the assistance of these institutions, this work would not have been possible.

References

[1] Gani, J. (1991) Generating function methods in a random allocation problem of epidemics. Bull. Inst. Combin. Applns 3, 43-50.

[2] Gani, J. (1992) Random allocation methods in an epidemic model. In"Stochastic Processes: A Festschrift in Honour of Gopinath Kallianpur", Edited by S. Cambanis, J.K. Ghosh, R. Karandikar and P.K. Sen. Springer-Verlag, New York, 97-106.

[3] Gani, J. (2004) Random allocation and urn models. J.Appl.Prob.41A, 313-320.