

Spread of A Rumor

Nick Fedewa, Central Michigan University
Emily Krause, Central Michigan University
Alexandra Sisson, Central Michigan University

Advisor: James Angelos
Department of Mathematics
Central Michigan University
Mt. Pleasant, MI 48858

April 25, 2013

Abstract

Modelling the random spread of a rumor has a long history. In this article we consider a random process that is based on sampling without replacement leading to the use of the discrete hypergeometric distribution. First considered is the model with only spreaders and ignorants followed by more general models where there are spreaders, ignorants, and stiflers. In this case a multivariate hypergeometric model is applied. It is shown that, as in the traditional case, not all ignorants hear the rumor.

1 Introduction

The classical model for the spread of a rumor consists of n individuals one of whom, the spreader, knows a rumor. Those individuals who do not know the rumor are called ignorants. On day one the spreader chooses one individual at random, which could be the spreader himself, to whom to tell the rumor. As noted it may not spread the first day. If an ignorant is chosen, the spreader tells the rumor and on the next day there are two spreaders. Each spreader chooses one person at random (which could be himself, the other spreader, or an ignorant) to whom to tell the rumor. This continues until all persons have heard the rumor. In [7] it is shown that if S_n is the minimum time for all ignorants to hear the rumor then

$$S_n = \log_2 n + \log n + O(1) \text{ in probability as } n \rightarrow \infty. \quad (1)$$

We show that a similar result holds, but is approximately twice as fast in the case of sampling without replacement. When sampling without replacement care is taken to avoid choosing

2 SPREADERS AND IGNORANTS

any element of the population more than once. In this way independence of the sample values is lost.

The more general model has three classes of individuals: spreaders (S), ignorants (I), and stiflers (R). Stiflers are individuals who have heard the rumor, but do not spread the rumor. In the case where sampling without replacement is used, a multivariate hypergeometric distribution is used. Two models are considered depending on the interaction of the three groups. In one model all of the ignorants eventually hear the rumor and in the other model a certain percentage of the ignorants fail to hear the rumor, depending on the initial number of spreaders.

2 Spreaders and Ignorants

Assume the population is size n and suppose on a given day there are k spreaders and hence $n - k$ ignorants. We then take a sample of size k without replacement. Interpreting the $n - k$ ignorants as “successes” then if X is the random variable counting the number of individuals to be told the rumor on the given day. This is the same as counting the number of successes in a sample of size k . Therefore

$$P(X = x) = \frac{\binom{n-k}{x} \binom{k}{k-x}}{\binom{n}{k}}, \quad 0 \leq x \leq \min(n-k, k).$$

This is the hypergeometric distribution with population size n , $n - k$ successes, and sample size k , which is denoted by $HG(n, n - k, k)$. We have then

$$E(X) = \frac{k(n-k)}{n} = k - \frac{k^2}{n}.$$

Note that a doubling of the number that have heard the rumor takes place when $E(X) > k$ or $k < \sqrt{\frac{n}{2}}$. For n large, when initially there is one spreader, we expect to have 2^j spreaders after j days when j is small. The following graphs depict simulations of this process of sampling without replacement.

2 SPREADERS AND IGNORANTS

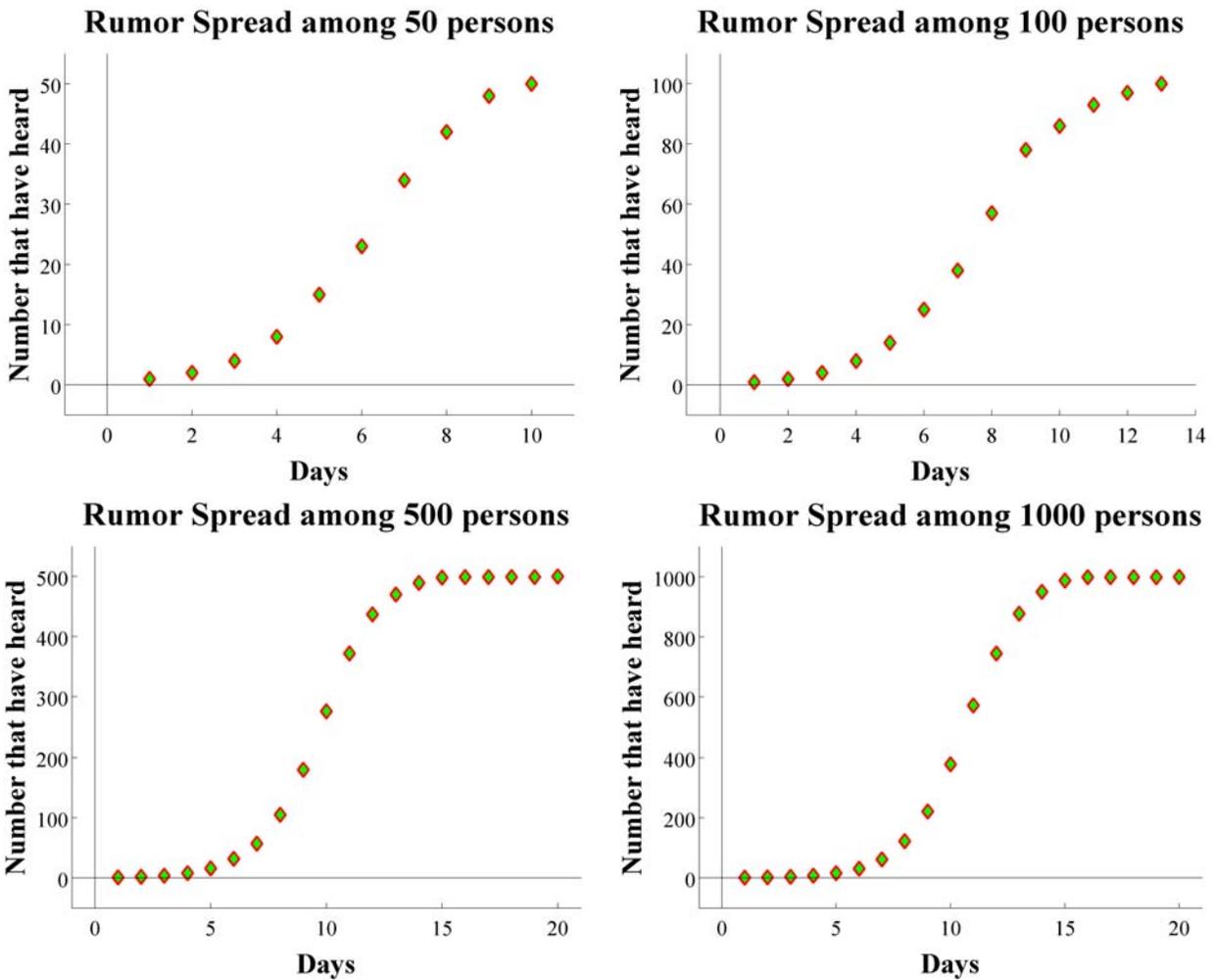


Figure 1: Four Population Sizes

Typical data are shown in table 1. Different simulations produce nearly the same data, but for length of time for all to hear the rumor.

2 SPREADERS AND IGNORANTS

Table 1
Simulation

	$n = 100$	$n = 500$	$n = 1000$	$n = 5000$
Day	k	k	k	k
1	1	1	1	1
2	2	2	2	2
3	4	4	4	4
4	8	8	8	8
5	15	15	16	16
6	27	29	31	32
7	44	54	61	64
8	69	102	121	127
9	89	179	227	251
10	100	297	401	490
11		420	654	926
12		483	888	1676
13		499	990	2790
14		500	1000	4022
15				4809
16				4994
17				5000

Table 2
Difference Equation

	$n = 100$	$n = 500$	$n = 1000$	$n = 5000$
Day	r_k	r_k	r_k	r_k
1	1.0	1.0	1.0	1.0
2	2.0	2.0	2.0	2.0
3	3.9	3.99	4.00	4.0
4	7.73	7.94	8.0	8.0
5	14.9	15.76	15.9	16.0
6	27.5	31.03	31.5	31.9
7	47.4	60.13	62.0	63.6
8	72.4	113.0	120.2	126.6
9	92.4	200.5	226.0	249.6
10	99.4	320.6	400.9	486.7
11	100.0	435.6	641.0	926.0
12		491.7	871.1	1681.0
13		499.9	983.4	2796.0
14			999.7	4029.0
15				4811.0
16				4993.0
17				5000.0

Note that, as expected the number of spreaders doubles early on in the simulations. We are now interested in finding an expression for how long it takes for the entire population to know the rumor. Let r_j be the expected number that know the rumor on day j . Then from the expected value of the hypergeometric distribution we have

$$r_{j+1} = r_j + r_j - \frac{r_j^2}{n} = r_j + \frac{1}{n}r_j(n - r_j), \quad r_1 = 1, \quad k = 1, 2, \dots, m.$$

Here m is the number of days for r_j to reach n . This is a logistic difference equation and produces similar data (Table 2) as in Table 1.

This logistic difference equation is an Euler's method approximation to the differential equation

$$\frac{dy}{dt} = y \left(1 - \frac{y}{n}\right), \quad y(0) = 1.$$

with step size $h = 1$. Since this is a separable equation the exact solution is found to be

$$y(t) = \frac{n}{1 + (n - 1)e^{-t}}, \quad t \geq 0.$$

We use this exact solution to estimate m . We want to find the time t_n at which $y(t)$ is closest to have the integer value n . So we let $0 < \varepsilon < 1/2$ and solve the equation $y(t_n) = n - \varepsilon$ for t_n , which yields the solution $t_n = \log((n - \varepsilon)(n - 1)/\varepsilon)$. Here and henceforth, \log denotes the natural logarithm. Using the Taylor series for $\log(1 - x) = -x + x^2/2 + O(x^3)$ where $x = \varepsilon/n$ or $x = 1/n$. Choose n large enough so that $|x| < 1$,

2 SPREADERS AND IGNORANTS

$$\begin{aligned}
 t_n &= \log((n - \varepsilon)(n - 1)/\varepsilon) = \log \left[n \left(1 - \frac{\varepsilon}{n} \right) \right] + \log \left[n \left(1 - \frac{1}{n} \right) \right] + \log(1/\varepsilon) \\
 &= 2 \log n + \left[-\frac{\varepsilon}{n} + O\left(\frac{\varepsilon^2}{n^2}\right) \right] + \left[-\frac{1}{n} + O\left(\frac{1}{n^2}\right) \right] + \log(1/\varepsilon) \\
 &= c \log_2 n + \log(1/\varepsilon) - \left(\frac{1 + \varepsilon}{n} \right) + O\left(\frac{1}{n^2}\right)
 \end{aligned}$$

where $c = 2 \log 2 \approx 1.386$. For n large the dominant term above is $c \log_2 n$ and so the time estimate for all of the ignorants to hear the rumor is

$$t_n \approx 1.386 \log_2 n, \quad \text{for } n \text{ large.}$$

Combining the dominant terms in (1) from [7], gives the time estimate in the case of sampling with replacement as approximately $2.443 \log n$. Thus, a rumor reaches the entire population approximately 1.75 times faster when using sampling without replacement. The following link gives an animation that demonstrates this increase in speed.

<http://people.cst.cmich.edu/angel1jr/Rumor/RumorMovie.htm>

The following graph compares simulation average time using 400 trials, average time, and the above estimate.

In [7], p. 213 at the bottom, it is noted that the random process modeling rumor spread is approximated by a difference equation. This difference equation is an Euler approximation for the differential equation

$$\frac{dy}{dt} = (n - y(t)) (1 - e^{-y(t)/n}), \quad y(0) = 1$$

with step size one. This equation cannot be solved explicitly. However, replacing the exponential term with its second degree Taylor approximation gives

$$\frac{dy}{dt} = (n - y(t)) \left(\frac{y(t)}{n} - \frac{y(t)^2}{2n^2} \right), \quad y(0) = 1.$$

This is a separable equation and the solution is given by

$$y(t) = n - \frac{n}{\sqrt{a_n e^t + 1}}, \quad \text{where } a_n = \frac{2n - 1}{(n - 1)^2}, \quad n > 1.$$

Note that $y(t) \rightarrow n$ as $t \rightarrow \infty$. As we did for the case without replacement, we let $0 < \varepsilon < 1/2$ and solve $y(s_n) = n - \varepsilon$. Using Taylor's series, this produces

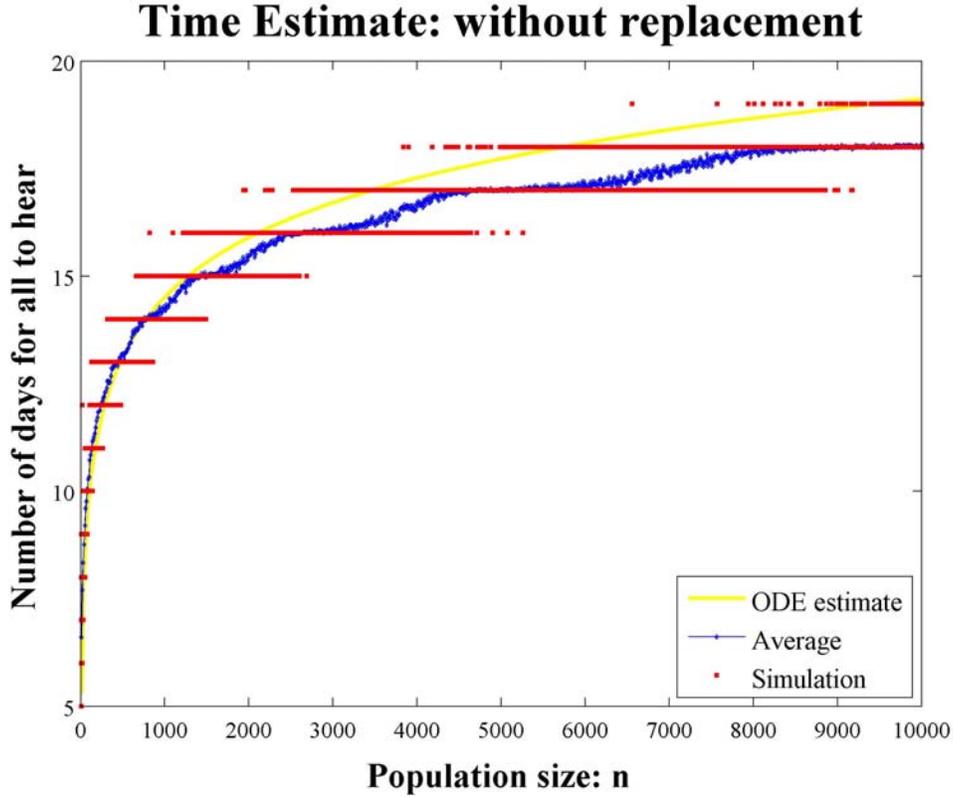


Figure 2: Time Estimate for entire population to hear the rumor: Without Replacement

$$\begin{aligned}
 s_n &= \log \left(\frac{(n^2 - \varepsilon^2)(n - 1)^2}{\varepsilon^2(2n - 1)} \right) \\
 &= 2 \log(n - 1) - \log(2n - 1) + \log(n^2 - \varepsilon^2) + \log(1/\varepsilon^2) \\
 &= 2 \log n + 2 \log \left(1 - \frac{1}{n} \right) - \log n - \log \left(2 - \frac{1}{n} \right) + 2 \log n + \log \left(1 - \frac{\varepsilon^2}{n^2} \right) + \log(1/\varepsilon^2) \\
 &= 2 \log n - \frac{2}{n} + O \left(\frac{1}{n^2} \right) - \log n - \log 2 + \frac{1}{2n} + O \left(\frac{1}{n^2} \right) + 2 \log n + O \left(\frac{1}{n^2} \right) + \log(1/\varepsilon^2) \\
 &= 3 \log n - \log 2 + \log(1/\varepsilon^2) - \frac{3}{2n} + O \left(\frac{1}{n^2} \right) \\
 &= \log_2 n + \left(2 - \frac{1}{\log 2} \right) \log n + \log \left(\frac{1}{2\varepsilon^2} \right) + O \left(\frac{1}{n} \right) \\
 &\approx \log_2 n + 1.56 \log n + \log \left(\frac{1}{2\varepsilon^2} \right) + O \left(\frac{1}{n} \right)
 \end{aligned}$$

3 ADDING STIFLERS

This estimate is slightly larger than that given in [7]. Figure 3 compares simulation average time using 400 trials, average time, the above estimate, and the bound, in green, according to the theory in [7].

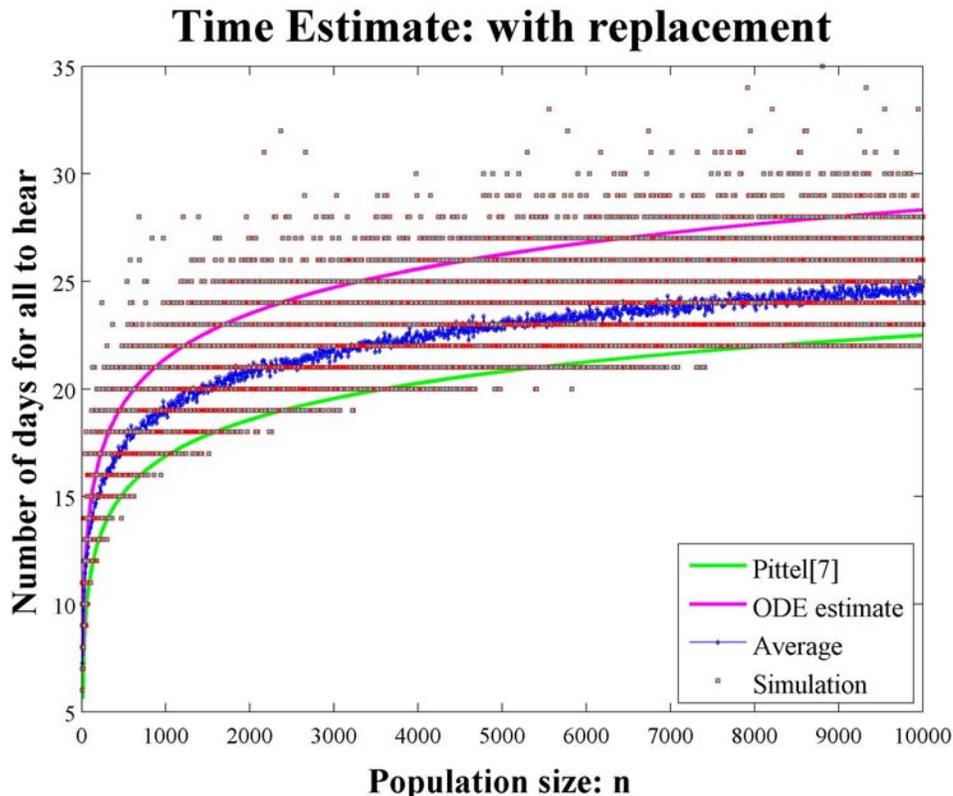


Figure 3: Time Estimate for entire population to hear the rumor: With Replacement

This result shouldn't be surprising since a spreader is not wasting any time talking to another spreader or two spreaders telling the same ignorant on a given day. When sampling without replacement each spreader is forced to talk to a unique person.

3 Adding Stiflers

In a model posed in [3] the population has three distinct groups, spreaders (S), ignorants (I), and stiflers (R). We pose that a person could become a stifter in three different ways, spreader-spreader interactions, spreader-stifler interactions, or spreader-ignorant interactions. We will consider each of these in order and in separate models.

The random vector (S, I, R) will be assumed to have a multivariate hypergeometric distribution and in order to perform simulations, a standard stepwise methodology is employed using inversion of the cumulative distribution functions $F_S(s)$ and $F_{I|S}(i)$.

3 ADDING STIFLERS

Step 1: Let $F_S(s)$ be the marginal distribution function of S and let u_1 be a value from a uniform distribution on $[0, 1)$. Then set $x_1 = F_S^{-1}(u_1)$. This is a sample from $HG(n, s, s)$, where s is the number of spreaders. Then $E(X_1) = \frac{s^2}{n}$

Step 2: Let $F_{I|S}(i)$ be the conditional distribution function of I given S and let u_2 be a value from a uniform distribution of $[0, 1)$. Then set $x_2 = F_{I|S}^{-1}(u_2)$. This is a sample from $HG(n - s, i, n - x_1)$. Then $E(X_2|X_1) = \frac{i(s - x_1)}{n - s}$.

Step 3: The value of x_3 , the number of stiflers, is $s - x_1 - x_2$. This value is from $HG(n - s - i, r, s - x_1 - x_2)$ and so $E(X_3|X_1, X_2) = \frac{r(s - x_1 - x_2)}{n - s - i}$.

3.1 Model 1

This model is characterized by the fact that stiflers are created by spreaders interacting with spreaders and spreaders are created from spreaders interacting with ignorants. In this case, when a spreader encounters another spreader they could assume that everyone has heard the rumor and so they stop spreading the rumor. Using the above sample procedure we create the following discrete dynamical system.

$$s_{t+1} = s_t - x_1 + x_2, \quad i_{t+1} = i_t - x_2, \quad r_{t+1} = r_t + x_1.$$

Using the expected values $E(X_1)$ and $E(X_2|X_1)$.

$$\begin{aligned} s_{t+1} &= s_t - E(X_1) + E(X_2|X_1) \\ &= s_t - \frac{s_t^2}{n} + \left(\frac{s_t - \frac{s_t^2}{n}}{n - s_t} \right) i_t \\ &= s_t - \frac{s_t^2}{n} + \left(\frac{s_t(n - s_t)}{n(n - s_t)} \right) i_t \\ &= s_t - \frac{s_t^2}{n} + \frac{s_t i_t}{n} \\ &= s_t + \frac{s_t}{n} (i_t - s_t). \end{aligned}$$

Therefore

$$s_{t+1} = s_t + \frac{1}{n} s_t (i_t - s_t).$$

This reflects that spreaders are created from s - i interactions and s - s interactions produce stiflers. This is an Euler approximation with step size one for the differential equation

$$\frac{dS}{dt} = \frac{S}{n} (I - S).$$

3 ADDING STIFLERS

In a similar fashion we have

$$i_{t+1} = i_t - \frac{s_t i_t}{n} \quad \text{which is an approximation of} \quad \frac{dI}{dt} = -\frac{1}{n}SI$$

and since $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$ we have

$$r_{t+1} = r_t + \frac{s_t^2}{n} \quad \text{which is an approximation of} \quad \frac{dR}{dt} = \frac{S^2}{n}.$$

Hence we will consider the following system of equations

$$\frac{dS}{dt} = \frac{S}{n}(I - S), \quad \frac{dI}{dt} = -\frac{1}{n}SI, \quad \frac{dR}{dt} = \frac{S^2}{n}. \quad (2)$$

The initial values are usually given as $S(0) = 1$ spreader, $I(0) = n - 2$ ignorants, and $R(0) = 1$ stifter.

We are interested in looking at the proportion each subpopulation represents. This is we are interested in finding $s = S/n$, $i = I/n$, and $r = R/n$. This is reflected in the following equations.

$$\frac{dS}{dt} \frac{1}{n} = \frac{ds}{dt}, \quad \frac{dI}{dt} \frac{1}{n} = \frac{di}{dt}, \quad \frac{dR}{dt} \frac{1}{n} = \frac{dr}{dt},$$

which give

$$\frac{ds}{dt} = s(i - s), \quad \frac{di}{dt} = -si, \quad \frac{dr}{dt} = s^2. \quad (3)$$

Here the initial values are $i(0) = \alpha$, $s(0) = \beta$, and $r(0) = \gamma$, where $0 < \alpha < 1$, $0 < \beta < 1$, $0 \leq \gamma < 1$, and $\alpha + \beta + \gamma = 1$.

It will be shown that the number of ignorants and spreaders approach zero. This indicates that eventually, since it is a closed population, the spreaders realize there is no one left to tell and hence stop spreading the rumor. Figure 4 shows the differential equations (2) compared with the simulations as described above. The simulations were stopped when the values of s , i and r repeated at least 5 times. Similar graphs result if $R(0) = 0$, i.e., no stiflers.

3.1.1 Asymptotic Behavior

We know that $0 \leq i < 1$ by definition, and we see from (3) that $\frac{di}{dt} < 0$ for all t . Hence the $i_\infty = \lim_{t \rightarrow \infty} i(t)$ exists and $0 \leq i_\infty < 1$.

Since $\frac{dr}{dt} > 0$ and $0 \leq r < 1$, r is increasing and bounded. Therefore $\lim_{t \rightarrow \infty} r = r_\infty$ exists. This implies that $s^2 = \frac{dr}{dt} \rightarrow 0$ as $t \rightarrow \infty$. Thus $s_\infty = \lim_{t \rightarrow \infty} s = 0$.

We claim that $i_\infty = 0$. Suppose that $i_\infty > 0$. Since $s(t) \rightarrow 0$ we have for t large enough, $\frac{ds}{dt} = s(i - s) > 0$. Hence, s is eventually increasing. We have then that if s is eventually increasing, positive, and bounded, then $s(t) \not\rightarrow 0$. This contradicts $s_\infty = 0$, and hence it cannot be the case that $i_\infty > 0$. This implies that $\lim_{t \rightarrow \infty} r(t) = 1$, i.e., the entire population becomes stiflers.

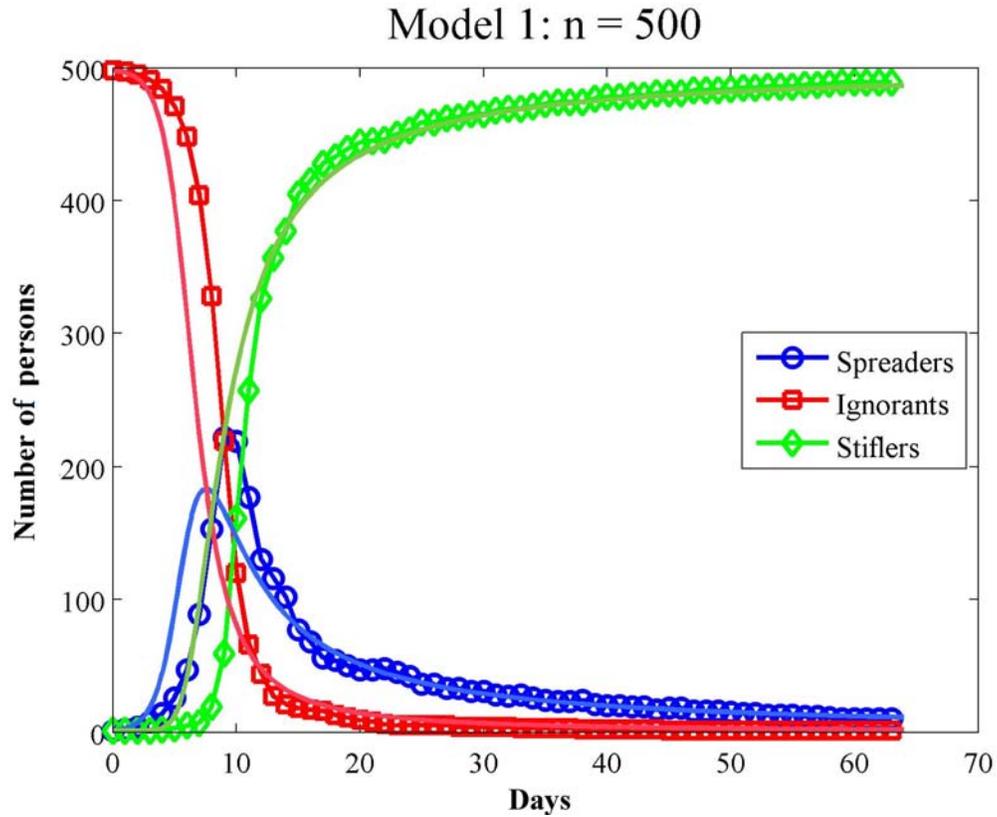


Figure 4: Model 1: Discrete system compared with differential equations (solid lines)

Since $\frac{di}{dt} < 0$, strictly decreasing, we can take i as an independent variable and we may consider the equation

$$\frac{ds}{di} = \frac{s-i}{i}, \quad s(\alpha) = \beta.$$

This equation is the first-order linear equation $\frac{ds}{di} - \frac{s}{i} = -1$ and has the solution

$$s(i) = \frac{\beta}{\alpha}i + i \log\left(\frac{\alpha}{i}\right). \tag{4}$$

We note that $s(t)$ reaches a maximum and its value can be found by noting from (3) that $\frac{ds}{dt} = 0$ when $s = 0$ or $s = i$. Taking $s = i$ in (4) we have

$$s = \frac{\beta}{\alpha}s + s \log\left(\frac{\alpha}{s}\right), \quad \text{and solving for } s \text{ gives } s = \alpha e^{(\beta-\alpha)/\alpha}. \tag{5}$$

In equation (5) take $\alpha = (n-2)/n$, $\beta = 1/n$, and replacing s with S/n , the maximum value of S is

3 ADDING STIFLERS

$$(n - 2) \exp\left(\frac{3 - n}{n - 2}\right).$$

For n large this is approximately $(n - 2)e^{-1}$ or about 37% of the population.

3.2 Model 2

For the second model the spreader-spreader interactions that create a stifter are removed and replaced with spreader-stifter interactions to create a stifter. We assume here that when a stifter meets a spreader, the spreader may be convinced that the truth of the rumor is brought into question and so stop spreading it or simply convinced to stop. Using the same method with the expected values for each interaction and removing the aforementioned interactions, we obtain the following new set of equations.

$$s_{t+1} = s_t - x_3 + x_2, \quad i_{t+1} = i_t - x_2, \quad r_{t+1} = r_t + x_3;$$

From which we obtain the system of difference equations

$$s_{t+1} = s_t - \frac{s_t}{n}(i_t - r_t), \quad i_{t+1} = i_t - \frac{s_t i_t}{n}, \quad r_{t+1} = r_t + \frac{s_t r_t}{n},$$

and the differential equations

$$\frac{dS}{dt} = \frac{S}{n}(I - R), \quad \frac{dI}{dt} = -\frac{SI}{n}, \quad \frac{dR}{dt} = \frac{SR}{n},$$

and for the proportions

$$\frac{ds}{dt} = s(i - r), \quad \frac{di}{dt} = -si, \quad \frac{dr}{dt} = sr. \quad (6)$$

Comparing the differential equations with simulations is shown in Figure 5. Here the initial values $S(0) = 1$, $I(0) = 479$, and $R(0) = 20$.

3.2.1 Asymptotic Behavior

A similar argument as in Model 1 shows that $\lim_{t \rightarrow \infty} s(t) = 0$. As before we consider the equation

$$\frac{ds}{di} = \frac{1 - s - 2i}{i}, \quad s(\alpha) = \beta.$$

This equation is a first-order linear equation, $\frac{ds}{di} + \frac{s}{i} = \frac{1}{i} - 2$, and so the solution is

$$s(i) = 1 - i + \frac{\alpha(\alpha + \beta - 1)}{i}. \quad (7)$$

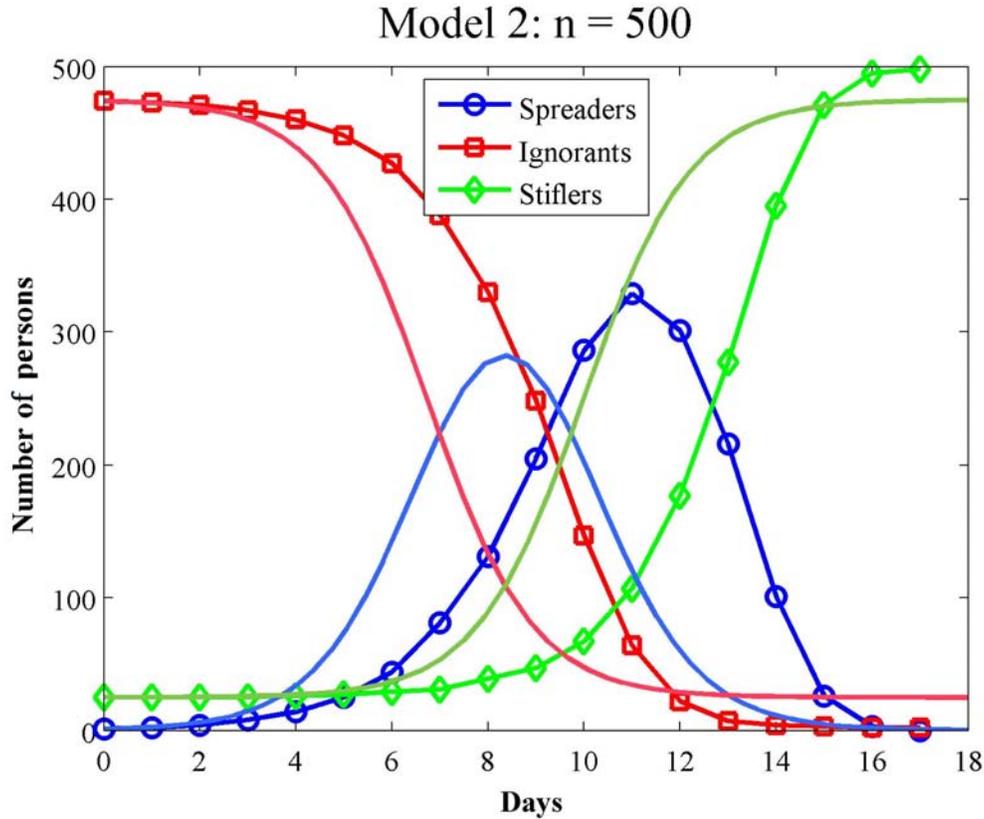


Figure 5: Model 2: Discrete system compared with differential equations (solid lines)

Since $s(t) \rightarrow 0$ we have from (7) setting $s = 0$

$$0 = 1 - i + \frac{\alpha(\alpha + \beta - 1)}{i}.$$

The solutions are

$$i = \frac{1}{2} \pm \frac{1}{2} \sqrt{1 - 4\alpha(1 - \alpha - \beta)} = \frac{1}{2} \left(1 \pm \sqrt{1 - 4\alpha\gamma} \right).$$

Let $l = \frac{1}{2} (1 - \sqrt{1 - 4\alpha\gamma})$ and $u = \frac{1}{2} (1 + \sqrt{1 - 4\alpha\gamma})$. We require that $1 - 4\alpha\gamma \geq 0$ or $4\alpha\gamma \leq 1$ or $\alpha\gamma \leq \frac{1}{4}$. Recall $\gamma = 1 - \alpha - \beta$ and $\alpha + \gamma \leq 1$ so $\alpha\gamma \leq \frac{1}{4}$ with equality when $\alpha = \gamma = \frac{1}{2}$. However, this would imply that $\beta = 0$, i.e., no spreaders which would mean that the rumor would never start so it is always assumed that α and β are positive. Hence $\alpha + \gamma < 1$. This implies that the roots satisfy $0 < l \leq \alpha \leq u < 1$ and $0 < l \leq \gamma \leq u < 1$.

We show that $i(t) \rightarrow l$ and $r(t) \rightarrow u$ as $t \rightarrow \infty$.

As it happens the system (6) can be solved explicitly. Using Maple[6] the following solutions

4 MODEL 3

are obtained. Let $A = \sqrt{a + b^2}$ and $B = e^{-Ac}$. Then

$$s(t) = \frac{4A^2Be^{-At}}{4A^2 - (Be^{-At} + 2b)^2} \quad (8)$$

$$i(t) = \frac{s^2s + (s')^2 - ss''}{2s^3} = \frac{4a(Be^{-At} - A) - B^2e^{-2At}(A + b) + 4ab}{2[4A^2 - (Be^{-At} + 2b)^2]} \quad (9)$$

$$r(t) = \frac{-s^2s + (s')^2 - ss''}{2s^3} = \frac{4a(Be^{-At} + A) - B^2e^{-2At}(A - b) + 4ab}{2[4A^2 - (Be^{-At} + 2b)^2]} \quad (10)$$

The constants a , b , and c depend on the initial values α , β , and γ .

$$a = -4\alpha\gamma, \quad b = \alpha + \beta + \gamma = 1, \quad c = \frac{1}{\left[4 \left(\frac{a(A+\alpha+\beta-\gamma)}{A-\alpha+\beta+\gamma}\right)\right]^{1/A}}.$$

Since $A > 0$ and $B > 0$ then $s(t) \rightarrow 0$ as $t \rightarrow \infty$. Also

$$\lim_{t \rightarrow \infty} i(t) = \frac{-4Aa + 4ab}{2(4A^2 - 4B^2)} = \frac{a(b - A)}{2(A^2 - b^2)} = \frac{1}{2} \left(1 - \sqrt{1 - 4\alpha\gamma}\right) = l$$

and

$$\lim_{t \rightarrow \infty} r(t) = \frac{4aA + 4ab}{2(4A^2 - 4b^2)} = \frac{a(b + A)}{2(A^2 - b^2)} = \frac{1}{2} \left(1 + \sqrt{1 + 4\alpha\gamma}\right) = u.$$

This shows that for $\alpha > 0$ and $\gamma > 0$, not all of the ignorants hear the rumor and not all of the population becomes a stifter.

This is similar to the situation in [2], [3], and [5] with the exception that the final number of ignorants depends on the initial values.

4 Model 3

This third model stiflers are created from a spreader-spreader interaction in which one of the spreaders becomes a stifter. For a spreader-ignorant interaction, the ignorant becomes a spreader. Finally, in a spreader-stifter interaction, the spreader becomes a stifter. When a spreader meets a stifter, it could be the case that the stifter imparts a sense of guilt on the spreader or relates that the veracity of the rumor is in serious question. In this way a spreader changes their mind and decides to no longer tell the rumor. This gives the following set of equations

$$s_{t+1} = s_t - x_1 - x_3 + x_2, \quad i_{t+1} = i_t - x_2, \quad r_{t+1} = r_t + x_1 + x_3.$$

Using the expected values as in the other models the following discrete system is produced.

$$s_{t+1} = s_t + \frac{s_t}{n}(i_t - s_t - r_t), \quad i_{t+1} = i_t - \frac{s_t i_t}{n}, \quad r_{t+1} = r_t + \frac{s_t}{n}(s_t + r_t),$$

4 MODEL 3

and the two systems of differential equations

$$\frac{dS}{dt} = \frac{S}{n}(I - S - R), \quad \frac{dI}{dt} = -\frac{SI}{n}, \quad \frac{dR}{dt} = \frac{S}{n}(S + R),$$

$$\frac{ds}{st} = s(i - s - r), \quad \frac{di}{dt} = -si, \quad \frac{dr}{dt} = s(s + r).$$

Using $s + i + r = 1$ and eliminating r we have

$$\frac{ds}{dt} = s(2i - 1), \quad \frac{di}{dt} = -si, \quad \frac{dr}{dt} = s(1 - i).$$

This is the classical model analyzed in [3], [5], and [2]. The following graph compares the differential equations with simulations. Here the initial values are $S(0) = 1$, $I(0) = 498$, and $R(0) = 1$. As is known, (e.g. [2] or [5]), approximately 20% of the ignorants do not hear the rumor.

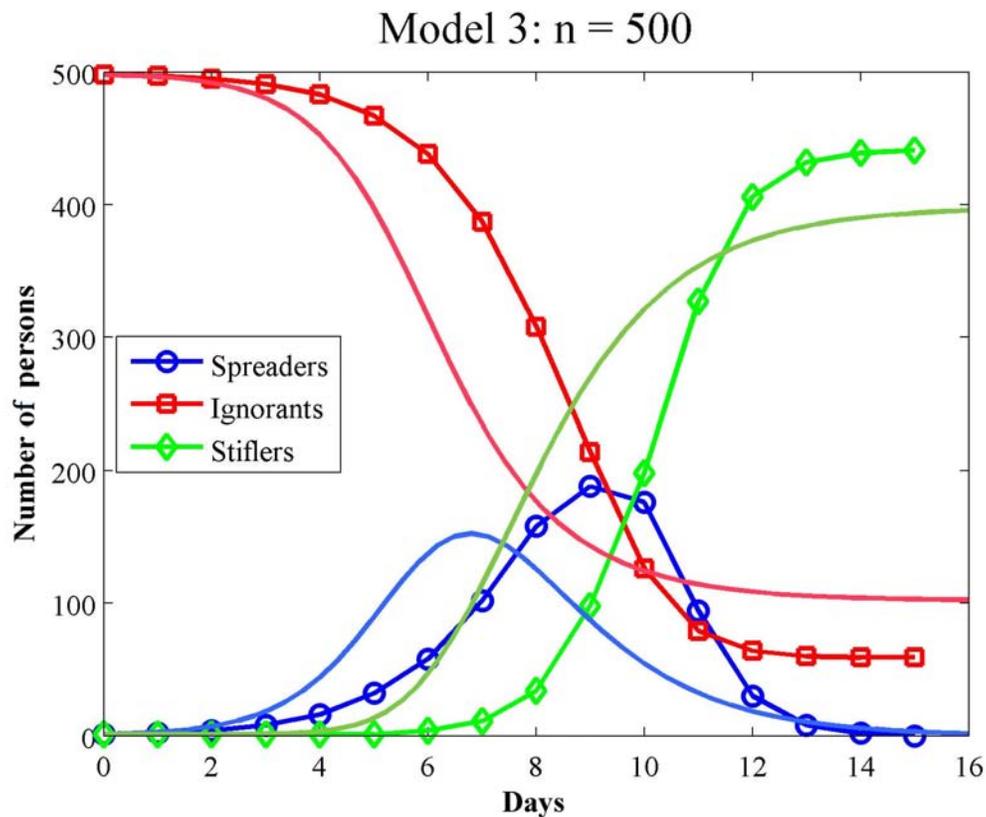


Figure 6: Model 3: Discrete system compared with differential equations (solid lines)

5 Conclusion

Using sampling without replacement various models of the spread of a rumor are analyzed. The analysis of these models require the use of both the single variable and multivariate hypergeometric distributions. The spreader-ignorant model produces a system where all the ignorants hear the rumor 1.75 times faster than the classical model. Adding stiflers and changing the ways stiflers are produced gives three models one of which is the classical model. In the first model the entire population becomes stiflers and in the second model not all ignorants hear the rumor, but unlike the classical model, the number of ignorants not hearing the rumor depends on the initial number of ignorants and stiflers.

Future work could entail including parameters that reflect imperfect information being passed, somewhat more in line with reality since once started, a rumor takes on a life of its own with different versions that evolve. This could also include adding terms that take into account that new information is added as the rumor spreads. For example, in the case that imperfect information is passed, model I takes the form

$$\frac{ds}{dt} = \mu si - \nu s^2 - \lambda sr, \quad \frac{di}{dt} = -\mu si, \quad \frac{dr}{dt} = \nu s^2 + \lambda sr$$

where $\lambda, \mu, \nu \in (0, 1]$. One question that comes to mind is whether or not parameter values of λ, μ, ν exist that produce chaotic behavior in any of the models.

6 Acknowledgements

The authors' work was done during the academic years and summers of 2008 to 2010 and was partially supported by NSF-LURE grant #06-36528.

REFERENCES

References

- [1] www.agner.org/random, A C++ class library of uniform and non-uniform random number generators.
- [2] Selma Belen and C. E. M. Pearce, Rumours with general initial conditions, The ANZIAM Journal, 45, 282-400, 2004
- [3] D. J. Daley and D. G. Kendal, Stochastic Rumours, J. Inst. Maths Applics, **1**, 42–55, 29 December 1964
- [4] William Feller, *An Introduction to Probability Theory and Its Applications*, Volume 1, Third Edition, Wiley, 1968.
- [5] D. P. Maki and M. Thompson, *Mathematical models and applications*, Prentice-Hall, Englewood Cliffs, 1973.
- [6] Maple 15, Maplesoft.
- [7] Boris Pittel, On Spreading A Rumor, SIAM J. Appl. Math., **47**, 213–223, February 1987.