

## 10.2: Branching Processes

### Historical Background

In this section we apply the theory of generating functions to the study of an important chance process called a

Until recently it was thought that the theory of branching processes originated with the following problem posed by Francis Galton in the in 1873.<sup>1</sup>

*Problem 4001: A large nation, of whom we will only concern ourselves with the adult males,  $N$  in number, and who each bear separate surnames, colonise a district. Their law of population is such that, in each generation,  $a_0$  per cent of the adult males have no male children who reach adult life;  $a_1$  have one such male child;  $a_2$  have two; and so on up to  $a_5$  who have five.*

*Find (1) what proportion of the surnames will have become extinct after  $r$  generations; and (2) how many instances there will be of the same surname being held by  $m$  persons.*

The first attempt at a solution was given by Reverend H. W. Watson. Because of a mistake in algebra, he incorrectly concluded that a family name would always die out with probability 1. However, the methods that he employed to solve the problems were, and still are, the basis for obtaining the correct solution.

Heyde and Seneta discovered an earlier communication by Bienaymé (1845) that anticipated Galton and Watson by 28 years. Bienaymé showed, in fact, that he was aware of the correct solution to Galton's problem. Heyde and Seneta in their book <sup>2</sup> give the following translation from Bienaymé's paper:

*If ...the mean of the number of male children who replace the number of males of the preceding generation were less than unity, it would be easily realized that families are dying out due to the disappearance of the members of which they are composed. However, the analysis shows further that when this mean is equal to unity families tend to disappear, although less rapidly ....*

*The analysis also shows clearly that if the mean ratio is greater than unity, the probability of the extinction of families with the passing of time no longer reduces to certainty. It only approaches a finite limit, which is fairly simple to calculate and which has the singular characteristic of being given by one of the roots of the equation (in which the number of generations is made infinite) which is not relevant to the question when the mean ratio is less than unity.<sup>3</sup>*

Although Bienaymé does not give his reasoning for these results, he did indicate that he intended to publish a special paper on the problem. The paper was never written, or at least has never been found. In his communication Bienaymé indicated that he was motivated by the same problem that occurred to Galton. The opening paragraph of his paper as translated by Heyde and Seneta says,

*A great deal of consideration has been given to the possible multiplication of the numbers of mankind; and recently various very curious observations have been published on the fate which allegedly hangs over the aristocracy and middle classes; the families of famous men, etc. This fate, it is alleged, will inevitably bring about the disappearance of the so-called <sup>4</sup>*

A much more extensive discussion of the history of branching processes may be found in two papers by David G. Kendall.<sup>5</sup>

Branching processes have served not only as crude models for population growth but also as models for certain physical processes such as chemical and nuclear chain reactions.

## Problem of Extinction

We turn now to the first problem posed by Galton (i.e., the problem of finding the probability of extinction for a branching process). We start in the 0th generation with 1 male parent. In the first generation we shall have 0, 1, 2, 3, ... male offspring with probabilities  $p_0, p_1, p_2, p_3, \dots$ . If in the first generation there are  $k$  offspring, then in the second generation there will be  $X_1 + X_2 + \dots + X_k$  offspring, where  $X_1, X_2, \dots, X_k$  are independent random variables, each with the common distribution  $p_0, p_1, p_2, \dots$ . This description enables us to construct a tree, and a tree measure, for any number of generations.

### ✓ Example 10.2.1

Assume that  $p_0 = 1/2, p_1 = 1/4$ , and  $p_2 = 1/4$ . Then the tree measure for the first two generations is shown in Figure ✓ 10.2.1.

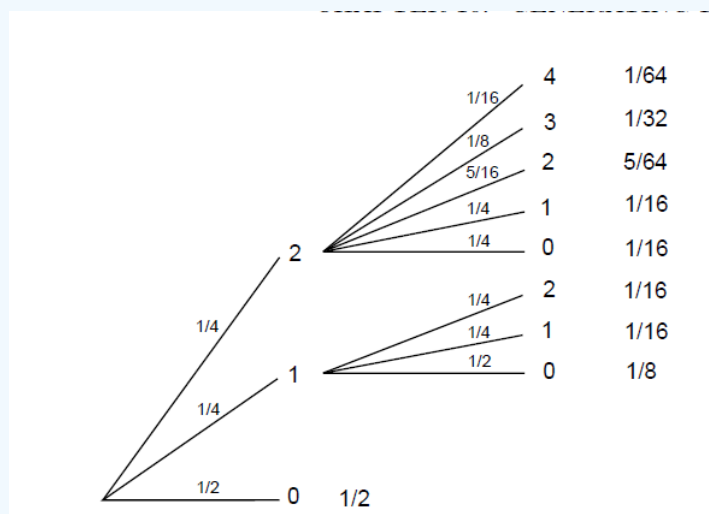


Figure 10.2.1: Tree diagram for Example 10.3.2

### Solution

Note that we use the theory of sums of independent random variables to assign branch probabilities. For example, if there are two offspring in the first generation, the probability that there will be two in the second generation is

$$\begin{aligned} P(X_1 + X_2 = 2) &= p_0 p_2 + p_1 p_1 + p_2 p_0 \\ &= \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{2} = \frac{5}{16}. \end{aligned}$$

We now study the probability that our process dies out (i.e., that at some generation there are no offspring).

Let  $d_m$  be the probability that the process dies out by the  $m$ th generation. Of course,  $d_0 = 0$ . In our example,  $d_1 = 1/2$  and  $d_2 = 1/2 + 1/8 + 1/16 = 11/16$  (see Figure [fig ✓ 10.2.1]). Note that we must add the probabilities for all paths that lead to 0 by the  $m$ th generation. It is clear from the definition that

$$0 = d_0 \leq d_1 \leq d_2 \leq \dots \leq 1. \quad (10.2.1)$$

Hence,  $d_m$  converges to a limit  $d, 0 \leq d \leq 1$ , and  $d$  is the probability that the process will ultimately die out. It is this value that we wish to determine. We begin by expressing the value  $d_m$  in terms of all possible outcomes on the first generation. If there are  $j$  offspring in the first generation, then to die out by the  $m$ th generation, each of these lines must die out in  $m - 1$  generations. Since they proceed independently, this probability is  $(d_{m-1})^j$ . Therefore

$$d_m = p_0 + p_1 d_{m-1} + p_2 (d_{m-1})^2 + p_3 (d_{m-1})^3 + \dots \quad (10.2.2)$$

Let  $h(z)$  be the ordinary generating function for the  $p_i$ :

$$h(z) = p_0 + p_1 z + p_2 z^2 + \dots \quad (10.2.3)$$

Using this generating function, we can rewrite Equation ✓ 10.2.1 in the form

$$d_m = h(d_{m-1}) . \quad (10.2.4)$$

Since  $d_m \rightarrow d$ , by Equation ✓ 10.2.2 we see that the value  $d$  that we are looking for satisfies the equation

$$d = h(d) . \quad (10.2.5)$$

One solution of this equation is always  $d = 1$ , since

$$1 = p_0 + p_1 + p_2 + \cdots . \quad (10.2.6)$$

This is where Watson made his mistake. He assumed that 1 was the only solution to Equation [eq 10.2.3]. To examine this question more carefully, we first note that solutions to Equation [eq 10.2.3] represent intersections of the graphs of

$$y = z \quad (10.2.7)$$

and

$$y = h(z) = p_0 + p_1 z + p_2 z^2 + \cdots . \quad (10.2.8)$$

Thus we need to study the graph of  $y = h(z)$ . We note that  $h(0) = p_0$ . Also,

$$h'(z) = p_1 + 2p_2 z + 3p_3 z^2 + \cdots , \quad (10.2.9)$$

and

$$h''(z) = 2p_2 + 3 \cdot 2p_3 z + 4 \cdot 3p_4 z^2 + \cdots . \quad (10.2.10)$$

From this we see that for  $z \geq 0$ ,  $h'(z) \geq 0$  and  $h''(z) \geq 0$ . Thus for nonnegative  $z$ ,  $h(z)$  is an increasing function and is concave upward. Therefore the graph of  $y = h(z)$  can intersect the line  $y = z$  in at most two points. Since we know it must intersect the line  $y = z$  at  $(1, 1)$ , we know that there are just three possibilities, as shown in Figure ✓ 10.2.2].

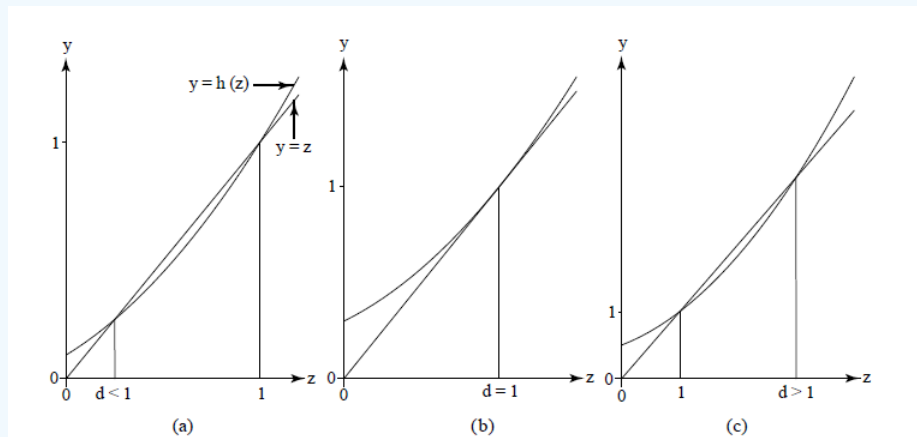


Figure 10.2.2: Graphs of  $y = z$  and  $y = h(z)$

In case (a) the equation  $d = h(d)$  has roots  $\{d, 1\}$  with  $0 \leq d < 1$ . In the second case (b) it has only the one root  $d = 1$ . In case (c) it has two roots  $\{1, d\}$  where  $1 < d$ . Since we are looking for a solution  $0 \leq d \leq 1$ , we see in cases (b) and (c) that our only solution is 1. In these cases we can conclude that the process will die out with probability 1. However in case (a) we are in doubt. We must study this case more carefully.

From Equation 10.2.4 we see that

$$h'(1) = p_1 + 2p_2 + 3p_3 + \cdots = m , \quad (10.2.11)$$

where  $m$  is the expected number of offspring produced by a single parent. In case (a) we have  $h'(1) > 1$ , in (b)  $h'(1) = 1$ , and in (c)  $h'(1) < 1$ . Thus our three cases correspond to  $m > 1$ ,  $m = 1$ , and  $m < 1$ . We assume now that  $m > 1$ . Recall that  $d_0 = 0$ ,  $d_1 = h(d_0) = p_0$ ,  $d_2 = h(d_1)$ , ..., and  $d_n = h(d_{n-1})$ . We can construct these values geometrically, as shown in Figure 10.2.3

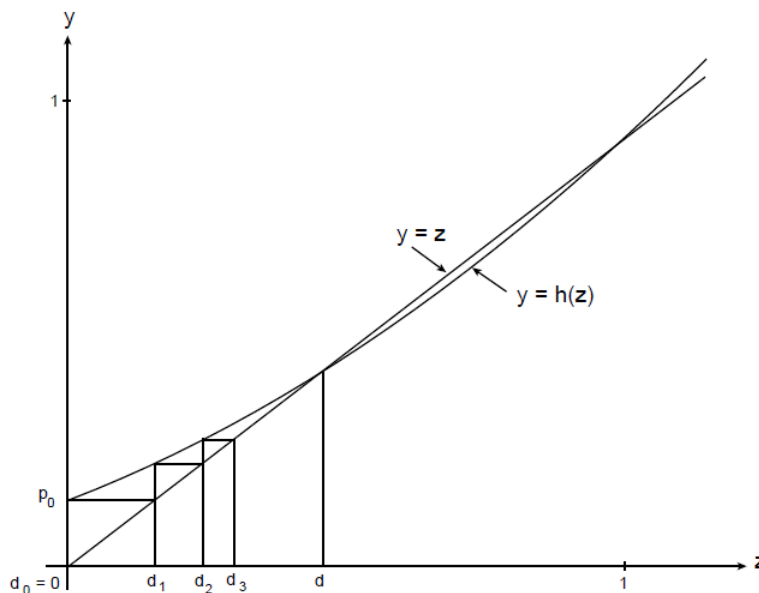


Figure 10.2.3: Geometric determination of  $d$ .

We can see geometrically, as indicated for  $d_0$ ,  $d_1$ ,  $d_2$ , and  $d_3$  in Figure 10.2.3 that the points  $(d_i, h(d_i))$  will always lie above the line  $y = z$ . Hence, they must converge to the first intersection of the curves  $y = z$  and  $y = h(z)$  (i.e., to the root  $d < 1$ ). This leads us to the following theorem.

#### Theorem 10.2.1

Consider a branching process with generating function  $h(z)$  for the number of offspring of a given parent. Let  $d$  be the smallest root of the equation  $z = h(z)$ . If the mean number  $m$  of offspring produced by a single parent is  $\leq 1$ , then  $d = 1$  and the process dies out with probability 1. If  $m > 1$  then  $d < 1$  and the process dies out with probability  $d$ .

We shall often want to know the probability that a branching process dies out by a particular generation, as well as the limit of these probabilities. Let  $d_n$  be the probability of dying out by the  $n$ th generation. Then we know that  $d_1 = p_0$ . We know further that  $d_n = h(d_{n-1})$  where  $h(z)$  is the generating function for the number of offspring produced by a single parent. This makes it easy to compute these probabilities.

The program **Branch** calculates the values of  $d_n$ . We have run this program for 12 generations for the case that a parent can produce at most two offspring and the probabilities for the number produced are  $p_0 = .2$ ,  $p_1 = .5$ , and  $p_2 = .3$ . The results are given in Table 10.2.1.

Table 10.2.1: Probability of dying out.

Generation	Probaboility of dying out
1	.2
2	.312
3	.385203
4	.437116
5	.475879
6	.505878

7	.529713
8	.549035
9	.564949
10	.578225
11	.589416
12	.598931

We see that the probability of dying out by 12 generations is about .6. We shall see in the next example that the probability of eventually dying out is  $2/3$ , so that even 12 generations is not enough to give an accurate estimate for this probability.

We now assume that at most two offspring can be produced. Then

$$h(z) = p_0 + p_1 z + p_2 z^2 . \quad (10.2.12)$$

In this simple case the condition  $z = h(z)$  yields the equation

$$d = p_0 + p_1 d + p_2 d^2 , \quad (10.2.13)$$

which is satisfied by  $d = 1$  and  $d = p_0/p_2$ . Thus, in addition to the root  $d = 1$  we have the second root  $d = p_0/p_2$ . The mean number  $m$  of offspring produced by a single parent is

$$m = p_1 + 2p_2 = 1 - p_0 - p_2 + 2p_2 = 1 - p_0 + p_2 . \quad (10.2.14)$$

Thus, if  $p_0 > p_2$ ,  $m < 1$  and the second root is  $> 1$ . If  $p_0 = p_2$ , we have a double root  $d = 1$ . If  $p_0 < p_2$ ,  $m > 1$  and the second root  $d$  is less than 1 and represents the probability that the process will die out.

#### ✓ Example 10.2.2

Keyfitz<sup>6</sup> compiled and analyzed data on the continuation of the female family line among Japanese women. His estimates at the basic probability distribution for the number of female children born to Japanese women of ages 45–49 in 1960 are given in Table 10.2.2

Table 10.2.2 Distribution of number of female children.

$p_1$	= .2584
$p_2$	= .2360
$p_3$	= .1593
$p_4$	= .0828
$p_5$	= .0357
$p_6$	= .0133
$p_7$	= .0042
$p_8$	= .0011
$p_9$	= .0002
$p_{10}$	= .0000

The expected number of girls in a family is then 1.837 so the probability  $d$  of extinction is less than 1. If we run the program **Branch**, we can estimate that  $d$  is in fact only about .324.

### Distribution of Offspring

So far we have considered only the first of the two problems raised by Galton, namely the probability of extinction. We now consider the second problem, that is, the distribution of the number  $Z_n$  of offspring in the  $n$ th generation. The exact form of the

distribution is not known except in very special cases. We shall see, however, that we can describe the limiting behavior of  $Z_n$  as  $n \rightarrow \infty$ .

We first show that the generating function  $h_n(z)$  of the distribution of  $Z_n$  can be obtained from  $h(z)$  for any branching process.

We recall that the value of the generating function at the value  $z$  for any random variable  $X$  can be written as

$$h(z) = E(z^X) = p_0 + p_1 z + p_2 z^2 + \cdots \quad (10.2.15)$$

That is,  $h(z)$  is the expected value of an experiment which has outcome  $z^j$  with probability  $p_j$ .

Let  $S_n = X_1 + X_2 + \cdots + X_n$  where each  $X_j$  has the same integer-valued distribution  $(p_j)$  with generating function  $k(z) = p_0 + p_1 z + p_2 z^2 + \cdots$ . Let  $k_n(z)$  be the generating function of  $S_n$ . Then using one of the properties of ordinary generating functions discussed in Section [sec 10.1], we have

$$k_n(z) = (k(z))^n, \quad (10.2.16)$$

since the  $X_j$ 's are independent and all have the same distribution.

Consider now the branching process  $Z_n$ . Let  $h_n(z)$  be the generating function of  $Z_n$ . Then

$$\begin{aligned} h_{n+1}(z) &= E(z^{Z_{n+1}}) \\ &= \sum_k E(z^{Z_{n+1}} | Z_n = k) P(Z_n = k). \end{aligned}$$

If  $Z_n = k$ , then  $Z_{n+1} = X_1 + X_2 + \cdots + X_k$  where  $X_1, X_2, \dots, X_k$  are independent random variables with common generating function  $h(z)$ . Thus

$$E(z^{Z_{n+1}} | Z_n = k) = E(z^{X_1 + X_2 + \cdots + X_k}) = (h(z))^k, \quad (10.2.17)$$

and

$$h_{n+1}(z) = \sum_k (h(z))^k P(Z_n = k). \quad (10.2.18)$$

But

$$h_n(z) = \sum_k P(Z_n = k) z^k. \quad (10.2.19)$$

Thus,

$$h_{n+1}(z) = h_n(h(z)). \quad (10.2.20)$$

If we differentiate Equation [eq 10.2.5] and use the chain rule we have

$$h'_{n+1}(z) = h'_n(h(z)) h'(z). \quad (10.2.21)$$

Putting  $z = 1$  and using the fact that  $h(1) = 1$ ,  $h'(1) = m$ , and  $h'_n(1) = m_n$  = the mean number of offspring in the  $n$ 'th generation, we have

$$m_{n+1} = m_n \cdot m. \quad (10.2.22)$$

Thus,  $m_2 = m \cdot m = m^2$ ,  $m_3 = m^2 \cdot m = m^3$ , and in general

$$m_n = m^n. \quad (10.2.23)$$

Thus, for a branching process with  $m > 1$ , the mean number of offspring grows exponentially at a rate  $m$ .

### ✓ Example 10.2.3

For the branching process of Example ✓ 10.2.1 we have

$$\begin{aligned}
 h(z) &= 1/2 + (1/4)z + (1/4)z^2, \\
 h_2(z) &= h(h(z)) = 1/2 + (1/4)[1/2 + (1/4)z + (1/4)z^2] \\
 &= 1/2 + (1/4)[1/2 + (1/4)z + (1/4)z^2]^2 \\
 &= 11/16 + (1/8)z + (9/64)z^2 + (1/32)z^3 + (1/64)z^4.
 \end{aligned}$$

The probabilities for the number of offspring in the second generation agree with those obtained directly from the tree measure (see Figure 10.2.1).

It is clear that even in the simple case of at most two offspring, we cannot easily carry out the calculation of  $h_n(z)$  by this method. However, there is one special case in which this can be done. Example 10.2.2

#### ✓ Example 10.2.4

Assume that the probabilities  $p_1, p_2, \dots$  form a geometric series:  $p_k = bc^{k-1}$ ,  $k = 1, 2, \dots$ , with  $0 < b \leq 1 - c$  and  $0 < c < 1$ . Then we have

$$\begin{aligned}
 p_0 &= 1 - p_1 - p_2 - \dots \\
 &= 1 - b - bc - bc^2 - \dots \\
 &= 1 - \frac{b}{1 - c}.
 \end{aligned}$$

The generating function  $h(z)$  for this distribution is

$$\begin{aligned}
 h(z) &= p_0 + p_1z + p_2z^2 + \dots \\
 &= 1 - \frac{b}{1 - c} + bz + bc^2z^2 + bc^2z^3 + \dots \\
 &= 1 - \frac{b}{1 - c} + \frac{bz}{1 - cz}.
 \end{aligned}$$

From this we find

$$h'(z) = \frac{bcz}{(1 - cz)^2} + \frac{b}{1 - cz} = \frac{b}{(1 - cz)^2} \quad (10.2.24)$$

and

$$m = h'(1) = \frac{b}{(1 - c)^2}. \quad (10.2.25)$$

We know that if  $m \leq 1$  the process will surely die out and  $d = 1$ . To find the probability  $d$  when  $m > 1$  we must find a root  $d < 1$  of the equation

$$z = h(z), \quad (10.2.26)$$

or

$$z = 1 - \frac{b}{1 - c} + \frac{bz}{1 - cz}. \quad (10.2.27)$$

This leads us to a quadratic equation. We know that  $z = 1$  is one solution. The other is found to be

$$d = \frac{1 - b - c}{c(1 - c)}. \quad (10.2.28)$$

It is easy to verify that  $d < 1$  just when  $m > 1$ .

It is possible in this case to find the distribution of  $Z_n$ . This is done by first finding the generating function  $h_n(z)$ .<sup>7</sup> The result for  $m \neq 1$  is:

$$h_n(z) = 1 - m^n \left[ \frac{1-d}{m^n-d} \right] + \frac{m^n \left[ \frac{1-d}{m^n-d} \right]^2}{1 - \left[ \frac{m^n-1}{m^n-d} \right] z} . \quad (10.2.29)$$

The coefficients of the powers of  $z$  give the distribution for  $Z_n$ :

$$P(Z_n = 0) = 1 - m^n \frac{1-d}{m^n-d} = \frac{d(m^n-1)}{m^n-d} \quad (10.2.30)$$

and

$$P(Z_n = j) = m^n \left( \frac{1-d}{m^n-d} \right)^2 \cdot \left( \frac{m^n-1}{m^n-d} \right)^{j-1} , \quad (10.2.31)$$

for  $j \geq 1$ .

### ✓ Example 10.2.5

Let us re-examine the Keyfitz data to see if a distribution of the type considered in Example ✓ 10.2.4 could reasonably be used as a model for this population. We would have to estimate from the data the parameters  $b$  and  $c$  for the formula  $p_k = bc^{k-1}$ . Recall that

$$m = \frac{b}{(1-c)^2} \quad (10.2.32)$$

and the probability  $d$  that the process dies out is

$$d = \frac{1-b-c}{c(1-c)} . \quad (10.2.33)$$

Solving Equation ✓ 10.2.7 and ✓ 10.2.8 for  $b$  and  $c$  gives

$$c = \frac{m-1}{m-d} \quad (10.2.34)$$

and

$$b = m \left( \frac{1-d}{m-d} \right)^2 . \quad (10.2.35)$$

We shall use the value 1.837 for  $m$  and .324 for  $d$  that we found in the Keyfitz example. Using these values, we obtain  $b = .3666$  and  $c = .5533$ . Note that  $(1-c)^2 < b < 1-c$ , as required. In Table [table 10.3] we give for comparison the probabilities  $p_0$  through  $p_8$  as calculated by the geometric distribution versus the empirical values.

Table 10.2.3: Comparison of observed and expected frequencies.

$p_j$	Data	Geometric
		Model
0	.2092	.1816
1	.2584	.3666
2	.2360	.2028
3	.1593	.1122



4	.0828	.0621
5	.0357	.0344
6	.0133	.0190
7	.0042	.0105
8	.0011	.0058
9	.0002	.0032
10	.0000	.0018

The geometric model tends to favor the larger numbers of offspring but is similar enough to show that this modified geometric distribution might be appropriate to use for studies of this kind.

Recall that if  $S_n = X_1 + X_2 + \cdots + X_n$  is the sum of independent random variables with the same distribution then the Law of Large Numbers states that  $S_n/n$  converges to a constant, namely  $E(X_1)$ . It is natural to ask if there is a similar limiting theorem for branching processes.

Consider a branching process with  $Z_n$  representing the number of offspring after  $n$  generations. Then we have seen that the expected value of  $Z_n$  is  $m^n$ . Thus we can scale the random variable  $Z_n$  to have expected value 1 by considering the random variable

$$W_n = \frac{Z_n}{m^n}. \quad (10.2.36)$$

In the theory of branching processes it is proved that this random variable  $W_n$  will tend to a limit as  $n$  tends to infinity. However, unlike the case of the Law of Large Numbers where this limit is a constant, for a branching process the limiting value of the random variables  $W_n$  is itself a random variable.

Although we cannot prove this theorem here we can illustrate it by simulation. This requires a little care. When a branching process survives, the number of offspring is apt to get very large. If in a given generation there are 1000 offspring, the offspring of the next generation are the result of 1000 chance events, and it will take a while to simulate these 1000 experiments. However, since the final result is the sum of 1000 independent experiments we can use the Central Limit Theorem to replace these 1000 experiments by a single experiment with normal density having the appropriate mean and variance. The program **BranchingSimulation** carries out this process.

We have run this program for the Keyfitz example, carrying out 10 simulations and graphing the results in Figure 10.2.4

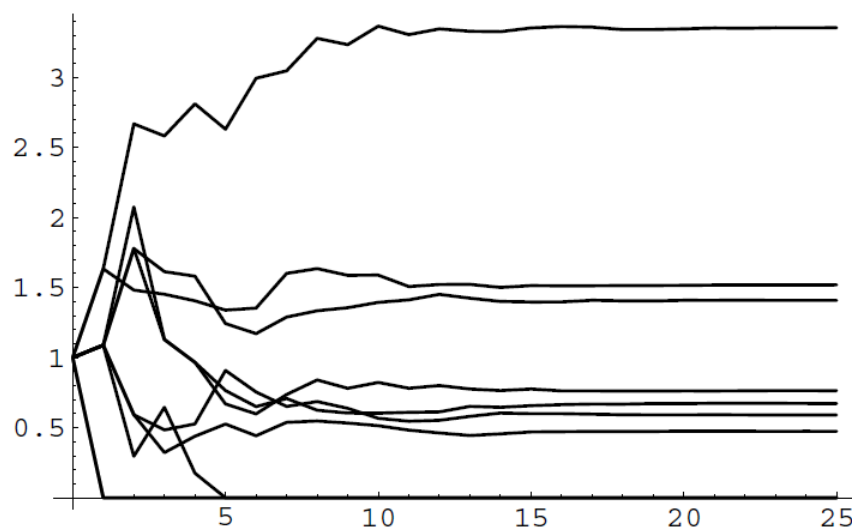


Figure 10.2.4: Simulation of  $Z_n / m^n$  for the Keyfitz example. }

The expected number of female offspring per female is 1.837, so that we are graphing the outcome for the random variables  $W_n = Z_n / (1.837)^n$ . For three of the simulations the process died out, which is consistent with the value  $d = .3$  that we found for this example. For the other seven simulations the value of  $W_n$  tends to a limiting value which is different for each simulation.

### ✓ Example 10.2.6

We now examine the random variable  $Z_n$  more closely for the case  $m < 1$  (see Example ✓ 10.2.4). Fix a value  $t > 0$ ; let  $[tm^n]$  be the integer part of  $tm^n$ . Then

$$\begin{aligned} P(Z_n = [tm^n]) &= m^n \left( \frac{1-d}{m^n-d} \right)^2 \left( \frac{m^n-1}{m^n-d} \right)^{[tm^n]-1} \\ &= \frac{1}{m^n} \left( \frac{1-d}{1-d/m^n} \right)^2 \left( \frac{1-1/m^n}{1-d/m^n} \right)^{tm^n+a}, \end{aligned}$$

where  $|a| \leq 2$ . Thus, as  $n \rightarrow \infty$ ,

$$m^n P(Z_n = [tm^n]) \rightarrow (1-d)^2 \frac{e^{-t}}{e^{-td}} = (1-d)^2 e^{-t(1-d)}. \quad (10.2.37)$$

For  $t = 0$ ,

$$P(Z_n = 0) \rightarrow d. \quad (10.2.38)$$

We can compare this result with the Central Limit Theorem for sums  $S_n$  of integer-valued independent random variables (see Theorem 9.3.5), which states that if  $t$  is an integer and  $u = (t - n\mu) / \sqrt{\sigma^2 n}$ , then as  $n \rightarrow \infty$ ,

$$\sqrt{\sigma^2 n} P(S_n = u \sqrt{\sigma^2 n} + \mu n) \rightarrow \frac{1}{\sqrt{2\pi}} e^{-u^2/2}. \quad (10.2.39)$$

We see that the form of these statements are quite similar. It is possible to prove a limit theorem for a general class of branching processes that states that under suitable hypotheses, as  $n \rightarrow \infty$ ,

$$m^n P(Z_n = [tm^n]) \rightarrow k(t), \quad (10.2.40)$$

for  $t > 0$ , and

$$P(Z_n = 0) \rightarrow d. \quad (10.2.41)$$

However, unlike the Central Limit Theorem for sums of independent random variables, the function  $k(t)$  will depend upon the basic distribution that determines the process. Its form is known for only a very few examples similar to the one we have considered here.

## Chain Letter Problem

### ✓ Example 10.2.1

An interesting example of a branching process was suggested by Free Huizinga.<sup>8</sup> In 1978, a chain letter called the "Circle of Gold," believed to have started in California, found its way across the country to the theater district of New York. The chain required a participant to buy a letter containing a list of 12 names for 100 dollars. The buyer gives 50 dollars to the person from whom the letter was purchased and then sends 50 dollars to the person whose name is at the top of the list. The buyer then crosses off the name at the top of the list and adds her own name at the bottom in each letter before it is sold again.

#### Solution

Let us first assume that the buyer may sell the letter only to a single person. If you buy the letter you will want to compute your expected winnings. (We are ignoring here the fact that the passing on of chain letters through the mail is a federal offense with certain obvious resulting penalties.) Assume that each person involved has a probability  $p$  of selling the letter. Then you will receive 50 dollars with probability  $p$  and another 50 dollars if the letter is sold to 12 people, since then your name would have

risen to the top of the list. This occurs with probability  $p^{12}$ , and so your expected winnings are  $-100 + 50p + 50p^{12}$ . Thus the chain in this situation is a highly unfavorable game.

It would be more reasonable to allow each person involved to make a copy of the list and try to sell the letter to at least 2 other people. Then you would have a chance of recovering your 100 dollars on these sales, and if any of the letters is sold 12 times you will receive a bonus of 50 dollars for each of these cases. We can consider this as a branching process with 12 generations. The members of the first generation are the letters you sell. The second generation consists of the letters sold by members of the first generation, and so forth.

Let us assume that the probabilities that each individual sells letters to 0, 1, or 2 others are  $p_0$ ,  $p_1$ , and  $p_2$ , respectively. Let  $Z_1$ ,  $Z_2$ , ...,  $Z_{12}$  be the number of letters in the first 12 generations of this branching process. Then your expected winnings are

$$50(E(Z_1) + E(Z_{12})) = 50m + 50m^{12}, \quad (10.2.42)$$

where  $m = p_1 + 2p_2$  is the expected number of letters you sold. Thus to be favorable we just have

$$50m + 50m^{12} > 100, \quad (10.2.43)$$

or

$$m + m^{12} > 2. \quad (10.2.44)$$

But this will be true if and only if  $m > 1$ . We have seen that this will occur in the quadratic case if and only if  $p_2 > p_0$ . Let us assume for example that  $p_0 = .2$ ,  $p_1 = .5$ , and  $p_2 = .3$ . Then  $m = 1.1$  and the chain would be a favorable game. Your expected profit would be

$$50(1.1 + 1.1^{12}) - 100 \approx 112. \quad (10.2.45)$$

The probability that you receive at least one payment from the 12th generation is  $1 - d_{12}$ . We find from our program **Branch** that  $d_{12} = .599$ . Thus,  $1 - d_{12} = .401$  is the probability that you receive some bonus. The maximum that you could receive from the chain would be  $50(2 + 2^{12}) = 204,900$  if everyone were to successfully sell two letters. Of course you can not always expect to be so lucky. (What is the probability of this happening?)

To simulate this game, we need only simulate a branching process for 12 generations. Using a slightly modified version of our program **BranchingSimulation** we carried out twenty such simulations, giving the results shown in Table ✓ 10.2.4

Note that we were quite lucky on a few runs, but we came out ahead only a little less than half the time. The process died out by the twelfth generation in 12 out of the 20 experiments, in good agreement with the probability  $d_{12} = .599$  that we calculated using the program **Branch**.

Table 10.2.4 Simulation of chain letter (finite distribution case).

									—	—	—	Profit
1	0	0	0	0	0	0	0	0	0	0	0	-50
1	1	2	3	2	3	2	1	2	3	3	6	250
0	0	0	0	0	0	0	0	0	0	0	0	-100
2	4	4	2	3	4	4	3	2	2	1	1	50
1	2	3	5	4	3	3	3	5	8	6	6	250
0	0	0	0	0	0	0	0	0	0	0	0	-100
2	3	2	2	2	1	2	3	3	3	4	6	300
1	2	1	1	1	1	2	1	0	0	0	0	-50
0	0	0	0	0	0	0	0	0	0	0	0	-100
1	0	0	0	0	0	0	0	0	0	0	0	-50

									—	—	—	Profit
2	3	2	3	3	3	5	9	12	12	13	15	750
1	1	1	0	0	0	0	0	0	0	0	0	-50
1	2	2	3	3	0	0	0	0	0	0	0	-50
1	1	1	1	2	2	3	4	4	6	4	5	200
1	1	0	0	0	0	0	0	0	0	0	0	-50
1	0	0	0	0	0	0	0	0	0	0	0	-50
1	0	0	0	0	0	0	0	0	0	0	0	-50
1	1	2	3	3	4	2	3	3	3	3	2	50
1	2	4	6	6	9	10	13	16	17	15	18	850
1	0	0	0	0	0	0	0	0	0	0	0	-50

Let us modify the assumptions about our chain letter to let the buyer sell the letter to as many people as she can instead of to a maximum of two. We shall assume, in fact, that a person has a large number  $N$  of acquaintances and a small probability  $p$  of persuading any one of them to buy the letter. Then the distribution for the number of letters that she sells will be a binomial distribution with mean  $m = Np$ . Since  $N$  is large and  $p$  is small, we can assume that the probability  $p_j$  that an individual sells the letter to  $j$  people is given by the Poisson distribution

$$p_j = \frac{e^{-m} m^j}{j!} . \quad (10.2.46)$$

The generating function for the Poisson distribution is

$$\begin{aligned} h(z) &= \sum_{j=0}^{\infty} \frac{e^{-m} m^j z^j}{j!} \\ &= e^{-m} \sum_{j=0}^{\infty} \frac{m^j z^j}{j!} \\ &= e^{-m} e^{mz} = e^{m(z-1)} . \end{aligned}$$

The expected number of letters that an individual passes on is  $m$ , and again to be favorable we must have  $m > 1$ . Let us assume again that  $m = 1.1$ . Then we can find again the probability  $1 - d_{12}$  of a bonus from **Branch**. The result is .232. Although the expected winnings are the same, the variance is larger in this case, and the buyer has a better chance for a reasonably large profit. We again carried out 20 simulations using the Poisson distribution with mean 1.1. The results are shown in Table Table 10.2.5

Table 10.2.5 Simulation of chain letter (Poisson case).

									—	—	—	Profit
1	2	6	7	7	8	11	9	7	6	6	5	200
1	0	0	0	0	0	0	0	0	0	0	0	-50
1	0	0	0	0	0	0	0	0	0	0	0	-50
1	1	1	0	0	0	0	0	0	0	0	0	-50
0	0	0	0	0	0	0	0	0	0	0	0	-100
1	1	1	1	1	1	2	4	9	7	9	7	300
2	3	3	4	2	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	-50

												Profit
2	1	0	0	0	0	0	0	0	0	0	0	0
3	3	4	7	11	17	14	11	11	10	16	25	1300
0	0	0	0	0	0	0	0	0	0	0	0	-100
1	2	2	1	1	3	1	0	0	0	0	0	-50
0	0	0	0	0	0	0	0	0	0	0	0	-100
2	3	1	0	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0	0	0	0	50
1	0	0	0	0	0	0	0	0	0	0	0	-50
3	4	4	7	10	11	9	11	12	14	13	10	550
1	3	3	4	9	5	7	9	8	8	6	3	100
1	0	4	6	6	9	10	13	0	0	0	0	-50
1	0	0	0	0	0	0	0	0	0	0	0	-50

We note that, as before, we came out ahead less than half the time, but we also had one large profit. In only 6 of the 20 cases did we receive any profit. This is again in reasonable agreement with our calculation of a probability .232 for this happening.

## Exercises

### Exercise 10.2.1

Let  $Z_1, Z_2, \dots, Z_N$  describe a branching process in which each parent has  $j$  offspring with probability  $p_j$ . Find the probability  $d$  that the process eventually dies out if

- $p_0 = 1/2, p_1 = 1/4$ , and  $p_2 = 1/4$ .
- $p_0 = 1/3, p_1 = 1/3$ , and  $p_2 = 1/3$ .
- $p_0 = 1/3, p_1 = 0$ , and  $p_2 = 2/3$ .
- $p_j = 1/2^{j+1}$ , for  $j = 0, 1, 2, \dots$
- $p_j = (1/3)(2/3)^j$ , for  $j = 0, 1, 2, \dots$
- $p_j = e^{-2} 2^j / j!$ , for  $j = 0, 1, 2, \dots$  (estimate  $d$  numerically).

### Exercise 10.2.2

Let  $Z_1, Z_2, \dots, Z_N$  describe a branching process in which each parent has  $j$  offspring with probability  $p_j$ . Find the probability  $d$  that the process dies out if

- $p_0 = 1/2, p_1 = p_2 = 0$ , and  $p_3 = 1/2$ .
- $p_0 = p_1 = p_2 = p_3 = 1/4$ .
- $p_0 = t, p_1 = 1 - 2t, p_2 = 0$ , and  $p_3 = t$ , where  $t \leq 1/2$ .

### Exercise 10.2.3

In the chain letter problem (see Example [exam 10.2.5]) find your expected profit if

- $p_0 = 1/2, p_1 = 0$ , and  $p_2 = 1/2$ .
- $p_0 = 1/6, p_1 = 1/2$ , and  $p_2 = 1/3$ .

Show that if  $p_0 > 1/2$ , you cannot expect to make a profit.

### Exercise 10.2.4

Let  $S_N = X_1 + X_2 + \dots + X_N$ , where the  $X_i$ 's are independent random variables with common distribution having generating function  $f(z)$ . Assume that  $N$  is an integer valued random variable independent of all of the  $X_j$  and having generating function

$g(z)$ . Show that the generating function for  $S_N$  is  $h(z) = g(f(z))$ . : Use the fact that

$$h(z) = E(z^{S_N}) = \sum_k E(z^{S_N} | N = k) P(N = k). \quad (10.2.47)$$

### Exercise 10.2.5

We have seen that if the generating function for the offspring of a single parent is  $f(z)$ , then the generating function for the number of offspring after two generations is given by  $h(z) = f(f(z))$ . Explain how this follows from the result of Exercise [exer 10.2.4].

### Exercise 10.2.6

Consider a queueing process such that in each minute either 1 or 0 customers arrive with probabilities  $p$  or  $q = 1 - p$ , respectively. (The number  $p$  is called the .) When a customer starts service she finishes in the next minute with probability  $r$ . The number  $r$  is called the .) Thus when a customer begins being served she will finish being served in  $j$  minutes with probability  $(1 - r)^{j-1}r$ , for  $j = 1, 2, 3, \dots$

### Exercise 10.2.7

Consider a by considering the offspring of a customer to be the customers who arrive while she is being served. Using Exercise [exer 10.2.4], show that the generating function for our customer branching process is  $h(z) = g(f(z))$ .

### Exercise 10.2.8

If we start the branching process with the arrival of the first customer, then the length of time until the branching process dies out will be the for the server. Find a condition in terms of the arrival rate and service rate that will assure that the server will ultimately have a time when he is not busy.

### Exercise 10.2.9

Let  $N$  be the expected total number of offspring in a branching process. Let  $m$  be the mean number of offspring of a single parent. Show that

$$N = 1 + \left( \sum p_k \cdot k \right) N = 1 + mN \quad (10.2.48)$$

and hence that  $N$  is finite if and only if  $m < 1$  and in that case  $N = 1/(1 - m)$ .

### Exercise 10.2.10

Consider a branching process such that the number of offspring of a parent is  $j$  with probability  $1/2^{j+1}$  for  $j = 0, 1, 2, \dots$

- a. Using the results of Example [exam 10.2.4] show that the probability that there are  $j$  offspring in the  $n$ th generation is

$$p_j^{(n)} = \begin{cases} \frac{1}{n(n+1)} \left( \frac{n}{n+1} \right)^j, & \text{if } j \geq 1, \\ \frac{n}{n+1}, & \text{if } j = 0. \end{cases} \quad (10.2.49)$$

- b. Show that the probability that the process dies out exactly at the  $n$ th generation is  $1/n(n+1)$ .  
c. Show that the expected lifetime is infinite even though  $d = 1$ .

---

This page titled [10.2: Branching Processes](#) is shared under a [GNU Free Documentation License 1.3](#) license and was authored, remixed, and/or curated by [Charles M. Grinstead & J. Laurie Snell](#) ([American Mathematical Society](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.