Exploding dice
A special case of summing a random number of random variables decided by a branching process

Mathias Berggren
Exploding dice

A special case of summing a random number of random variables decided by a branching process

Mathias Berggren
0. Abstract

This thesis looks at "exploding dice", a type of dice scheme commonly used in (table top) role-playing games. Unlike ordinary dice rolls, this dice scheme means that if one rolls some subset of the dice's numbers an extra number of dice are rolled. Here we present results about branching processes, integer partitions, and full K-ary trees to find an expression for the density for the final sum of the exploding dice, as well as their mean and variance. We also present a MATLAB script for calculating these probabilities. Finally some notes are given on how these results relate to the more general problem of summing a random number of random variables.

Acknowledgements

I would like to thank my supervisor Prof. Rolf Larsson for his great help with the theory needed for writing this thesis, in particular regarding how one uses probability generating functions when working with branching processes.

In addition I would like to thank Prof. Svante Jansson for presenting me with a method of finding the number of ordered partitions of an integer; as well as Prof. em. Allan Gut for help with finding the generating function of the total number of individuals in a branching process that dies out.
1. The problem

This thesis is concerned with answering the following problem: Suppose you roll a number of m-sided dice and want to count their sum. However if a die rolls strictly above a value $m_\alpha$ an extra $K$ number of dice are rolled, also to be included in the sum. What then is the probability distribution for the total sum?

Called "exploding dice", such a randomization-scheme is often used within role-playing- and other games (particularly table-top games) when you want to simulate an outcome which usually takes values within some interval, but has a chance of taking much higher values. For example, within the Swedish table-top game Eon (Eon IV, 2014) each player-character has a set of skill-values equal to a number of six-sided dice (plus some constant integer between 0 and 3). When rolling these dice, the sixes aren't counted, but instead result in two more dice being rolled. So when using ones skills, the dice are rolled and the final sum is compared to a difficulty level, and if ones sum exceeds the difficulty (or the sum of an opposing roll) the attempt is successful, with higher values possibly offering extra benefits.

Within this thesis we will first find the probability distribution for one Eon-die and derive its expected value and variance (so we can use the normal approximation for large number of dice). This will then be generalized to any m-sided die (any discrete uniform distribution between 1 and m) with any new number of dice thrown, either when counting or not counting the branching dice, and when branching occurs over some cut-off value of the possible values. Finally we will make some notes on how the model can be further generalized.

First of all however, we need to put the problem into its proper mathematical background. To achieve this we will begin with a description of the theory needed to solve our problem, that is: the theory behind branching processes. For the problem we will also present some combinatorics needed for finding an expression for our distribution. However, first we will give a short presentation of generating functions, since these play a vital part in the theory needed for our problem.
2. Background

2.1 Generating functions

We will exclusively look at ordinary generating functions, from now on simply called generating functions. These are defined in the following way: given a sequence $a_n$, the (ordinary) generating function of that sequence is given by:

$$g_{a_n}(t) = \sum_{n=0}^{\infty} a_n t^n$$

An important special case for our purposes is when $a_n$ denotes the probability mass function of a discrete random variable $X \geq 0$. In this case we will use the following notation for the probability generating function:

$$g_X(t) = \sum_{x=0}^{\infty} p_X(x) t^x = E_X(t^X)$$

where we will leave out the subscript when it is obvious which random variable (R.V.) we mean. We see that upon setting $t=0$ and $t=1$ we get:

$$g_{a_n}(0) = a_0, \quad g_X(0) = p_X(0)$$

$$g_{a_n}(1) = \sum_{n=0}^{\infty} a_n, \quad g_X(1) = \sum_{x=0}^{\infty} p_X(x) = 1$$

Differentiating the probability generating function (with respect to $t$) we get:

$$g_X'(t) = \sum_{x=1}^{\infty} x \cdot p(x) t^{x-1}$$

And more generally:

$$g_X^{(k)}(t) = \sum_{x=k}^{\infty} x(x-1)(x-2) \ldots (x-k+1) \cdot p(x) t^{x-k}$$

So setting $t=1$ we get:

$$g_X^{(k)}(1) = \sum_{x=k}^{\infty} x(x-1)(x-2) \ldots (x-k+1) \cdot p(x) = E(X(X-1) \ldots (X-k+1))$$

which is a useful relation for finding moments of positive discrete random variables. Furthermore setting $t=0$ we get:
\[ g_X^{(k)}(0) = k(k - 1)(k - 2) \ldots 2 \cdot 1 \cdot p(k) = k! \cdot p(k) \]

So the probability generating function single handedly defines the probability distribution of the R.V., meaning that if two R.V.:s have the same probability generating function, they must have the same distribution (and vice versa).

An important question for when we turn to branching processes is how to find the generating function of a sum of a random number of i.i.d. random variables. As such we will now prove the following theorem:

**Theorem 1:** For \( S = X_1 + X_2 + \cdots + X_N \), where the \( X_i \):s are i.i.d. R.V.:s and \( N \) is an R.V. independent of the \( X \):s, we have \( g_S(t) = g_N(g_X(t)) \).

**Proof** (by Grimmett & Stirzaker, 2004):

\[
g_S(t) = E(t^S) = E(E(t^S|N)) = \sum_n E(t^S|N = n)P(N = n) = \\
= \sum_n E(t^{X_1+\cdots+X_n}|N = n)P(N = n) = \{\text{by independence}\} = \\
= \sum_n E(t^{X_1} \ldots t^{X_n})P(N = n) = \sum_n g_X(t)^n P(N = n) = g_N(g_X(t))
\]

\( \square \)

With this relation we can now turn to branching processes.

### 2.2 Branching processes

A branching process is a process wherein each member of the \( n \):th generation gives birth to a number of members (possibly zero) of the \( n+1 \):th generation, the number of which are described by a random variable (Grimmett & Stirzaker, 2004). As such our problem is one of a branching process, since each die has a probability of branching, e.g. the Eon-die has probability \( 1/6 \) of generating two new dice in the next generation, and probability \( 5/6 \) of not generating any new dice.

Often one assumes the following about a branching process, both of which hold for our exploding dice:

(i) The number of "offspring" of all members of the branching process form a collection of independent random variables.

(ii) All members have the same probability mass function (and thus generating function) for their number of offspring.

Together with information about the number of "founding" (or starting) members of the process these assumptions specify the random evolution of the process. Throughout this chapter we will assume one founding member, as we otherwise simply get a number of independent branching processes of this type equal to the number of starting members.
2.2.1 Size of generation n

Usually when dealing with a branching process, one is interested in how the size of generations vary over time. In our problem however, we are only interested in the total number of dice, but the theory developed here will nevertheless be necessary to arrive at a solution for our problem.

We will denote the number of members in generation $n$ by $Z_n$ (the founding generation being denoted as generation 0, and so by our assumption $Z_0=1$). Then, if $g_n(t)$ is the generating function of $Z_n$, we have:

**Theorem 2:** $g_{m+n}(t) = g_m(g_n(t))$, so $g_n(t) = g\left(\ldots(g(t)\ldots)\right)$ is the $n$-fold iteration of $g(t) = g_1(t)$.

**Proof** (by Grimmett & Stirzaker, 2004): Each member of the $(m+n)$:th generation has a unique ancestor in the $m$:th generation, and so if we denote the number of members of the $(m+n)$:th generation which stem from the $i$:th member of the $m$:th generation by $X_i$, we have:

$$Z_{m+n} = X_1 + X_2 + \ldots + X_{Z_m}$$

I.e. a sum of a random number $Z_m$ of random variables. The $X$:es are independent by assumption (i) and identically distributed by assumption (ii), with the same distribution as the number $Z_n$ of the $n$:th generation offspring of the founding member in generation 0 (since each member generates a "new" branching process with it as its founding member). Now, by Theorem 1 we have that:

$$g_{m+n}(t) = g_m\left(g_{X_1}(t)\right) = \{\text{by similarity in distribution}\} = g_m(g_n(t))$$

And by iterating backwards we get:

$$g_n(t) = g_1(g_{n-1}(t)) = g_1\left(g_1(g_{n-2}(t))\right) = \ldots = g_1\left(\ldots(g_1(t)\ldots)\right)$$

and the proof of Theorem 2 is complete. □

2.2.2 Mean and variance of generation $n$

For simplicity we will just denote $g_1(t)$ by $g(t)$, since we see that $g(t)$ tells us all about our $Z_n$ and their distribution. For the mean and variance we have:

**Corollary 2.1:** Let $\mu = E(Z_1)$, and $\sigma^2 = \text{Var}(Z_1)$, then:

$$E(Z_n) = \mu^n, \quad \text{Var}(Z_n) = \begin{cases} n\sigma^2, & \text{if } \mu = 1 \\ \frac{\sigma^2(\mu^n-1)}{\mu-1}, & \text{if } \mu \neq 1 \end{cases}$$

**Proof** (by Grimmett & Stirzaker, 2004): We differentiate $g_n(t) = g_{n-1}(g(t))$ once to get $g'_n(t) = g'(t)g'_{n-1}(g(t))$, and set $t=1$ to obtain $E(Z_n) = \mu E(Z_{n-1})$ (since $g(1)=1$, and $g^{(k)}(1)=E[X(X-1)\ldots(X-k+1)]$ by property of the generating function). So iterating backwards
we get \( E(Z_n) = \mu E(Z_{n-1}) = \mu^2 E(Z_{n-2}) = \cdots = \mu^n \). Differentiating again and setting \( t = 1 \) we get:

\[
g''_n(1) = E(Z_n(Z_n - 1)) = g''(1)g'_n(g(1)) + (g'(s))^2g''_n(g(1)) =
\]

\[= E(Z_1(Z_1 - 1))\mu^{n-1} + \mu^2 E(Z_{n-1}(Z_{n-1} - 1)) \leftrightarrow \]

\[\leftrightarrow Var(Z_n) + \mu^{2n} - \mu^n = (\sigma^2 + \mu^2 - \mu)\mu^{n-1} + \mu^2 Var(Z_{n-1}) + \mu^{2(n-1)} - \mu^{n-1} \leftrightarrow \]

\[\leftrightarrow Var(Z_n) = \sigma^2 \mu^{n-1} + \mu^{n+1} - \mu^n + \mu^2 Var(Z_{n-1}) + \mu^{2n} - \mu^{n+1} - \mu^{2n} + \mu^n = \]

\[= \sigma^2 \mu^{n-1} + \mu^2 Var(Z_{n-1}) = \sigma^2 \mu^{n-1} + \sigma^2 \mu^n + \mu^4 Var(Z_{n-2}) = \sigma^2 \sum_{i=1}^{n} \mu^{(n-2)+i} = \]

\[= \sigma^2 \mu^{n-2} \sum_{i=1}^{n} \mu^i = \{ \text{if } \mu \neq 1 \} = \sigma^2 \mu^{n-2} \frac{\mu - \mu^{n+1}}{1 - \mu} = \frac{\sigma^2(\mu^n - 1)\mu^{n-1}}{\mu - 1} \]

By the formula for geometric series. Furthermore it is easy to see that we get \( n\sigma^2 \) if \( \mu = 1 \). \( \square \)

2.2.3 Extinction of a branching process

A question of importance for any branching process is whether it will eventually die out, e.g. in our dice problem we’d prefer not to have to roll new dice forever. From the formula for the mean of generation \( n \) we might expect that the branching process dies out if \( \mu < 1 \). To prove this we formulate the probability of extinction as:

\[\lim_{n \to \infty} P(Z_n = 0)\]

And letting \( d_n \) denote the probability that the branching process is extinct at generation \( n \) (i.e. that \( Z_n = 0 \)), we must have:

\[0 \leq d_1 \leq d_2 \leq \cdots \leq 1\]

Since if the process is extinct by generation \( n \), then surely it is so for generation \( n+1 \). But this means that \( d_n \) must converge to some limit \( d \) (\( \leq 1 \)) as \( n \) goes to infinity. Now, since:

\[d_n = P(Z_n = 0) = g_n(0)\]

And:

\[g_n(s) = g(g_{n-1}(s))\]

We must have:

\[d_n = g(d_{n-1})\]
When \( n \) goes to infinity \( d_n \) and \( d_{n-1} \) must converge to a limit \( d \), so:

\[
d = g(d) = \sum_{k=0}^{\infty} p(k)d^k
\]

As can be seen, by property of the generating function, \( g(0) = p(0) \) and \( g(1) = 1 \). Furthermore, since the \( d \):s and \( p \):s are all positive we must have:

\[
g'(d) > 0, \quad g''(d) > 0
\]

(Since the differentiated terms all become positive as well.) This means that \( g(d) \) is an increasing convex function for \( 0 \leq d \leq 1 \). Now our problem of finding \( d = g(d) \) is the same as finding the intersection between \( y = d \) and \( y = g(d) \). We find that \( d = 1 \) is always an intersection (since \( g(1) = 1 \)), and since \( y = g(d) \) is an increasing convex function it can only intersect \( y = d \) at most one other time between \( 0 \leq d \leq 1 \) if \( g'(1) > 1 \), and no other times otherwise. But since \( g'(1) = \mu \), this means that if \( \mu \leq 1 \) then \( d = 1 \) is the only solution, and so the branching process dies out with probability one, as we wanted to show. If \( \mu > 1 \) however, then the proper choice of \( d \) is the intersection less than 1, since \( d_n \) is increasing towards its bound, meaning that the bound must be the lower one. This in turn means that the branching process has positive probability of going on forever.

2.2.4 Generating function of the total population

For the mean of the total population of the branching process we can simply use the formula for the mean for generation \( n \) to get:

\[
E(\text{Total population}) = \sum_{n=0}^{\infty} \mu^n = \{ \text{if } \mu < 1 \} = \frac{1}{1 - \mu}
\]

By the formula for geometric series, with infinite expectation otherwise. However, for the variance things are not so simple, and once again we need to use generating functions to find an answer.

First, observing that since the mean of the total population does not exist if \( \mu \geq 1 \), the variance cannot exist either (since \( \text{Var}(X) = E(X^2) - E(X)^2 \)), and so we limit ourselves to the case when \( \mu < 1 \) and the branching process is sure to die out.

Denoting the total population of the branching process up to generation \( n \) by \( T_n \), and letting \( G_n(t) \) be the generating function for \( T_n \), so that:

\[
T_n = \sum_{k=0}^{n} Z_k = 1 + \sum_{k=1}^{n} Z_k
\]

We have:
Theorem 3: $G_n(t) = t \ast g(G_{n-1}(t))$

Proof (by Gut, 2013): Every individual in the first generation generates a new branching process. All these new processes are independent. Denote the total progeny of these branching processes up to (their) generation $n-1$ by $Y_1, Y_2, \ldots$ As such we have that our $Y$'s have the same distribution as $T_{n-1}$. Now, we have that:

$$T_n = 1 + \sum_{k=1}^{Z_1} Y_k$$

Since we have only one founding member. If we denote the rest of the sum (the $Y$'s) as $U$ we have that $U$ is a sum of a random number of variables (as we had before). Now, we note that for the generating function we have:

$$g_{1+U}(t) = E(t^{1+U}) = t \ast E(t^U) = t \ast g_U(t)$$

So for any non-negative integer valued R.V. $U$, using Theorem 1 we obtain:

$$G_n(t) = t \ast G_U(t) = t \ast G_{Z_1}(G_Y(t)) = t \ast g(G_{T_{n-1}}(t)) = t \ast g(G_{n-1}(t))$$

As wanted. $\square$

This generating function can be used to find a formula for the mean and variance of a branching process that dies out. We have:

**Corollary 3.1:** If $\mu < 1$ and $\sigma^2 < \infty$ we have that the mean $\mu_T$ and variance $\sigma_T^2$ of the total population of a branching process are:

$$\mu_T = \frac{1}{1-\mu}, \quad \sigma_T^2 = \frac{\sigma^2}{(1-\mu)^3}$$

Proof (by Gut 2013): Differentiating the generating function $G_n(t)$ once we get:

$$G'_n(t) = g(G_{n-1}(t)) + t \ast G'_{n-1}(t) \ast g'(G_{n-1}(t))$$

Now, since $G_n(t)$ denotes the generating function for the total population up to generation $n$, to find the generating function of the total population we let $n$ go to infinity. Since the process dies out, this means that $G_n(t)$ and $G_{n-1}(t)$ will converge to a limit $G(t)$. So setting $t=1$ and denoting the mean of the total population by $\mu_T$ we have:

$$G'(1) = g(G(1)) + 1 \ast G'(1) \ast g'(G(1)) \iff \mu_T = g(1) + \mu_T \ast g'(1) =$$

$$= 1 + \mu_T \ast \mu \iff \mu_T = \frac{1}{1-\mu}$$

As above. Assuming $\sigma^2 < \infty$ and differentiating again we get:
So letting $n$ go to infinity and setting $t=1$:

$$G'(1) = 2G'(1) * g'(G(1)) + 1 * g'(G(1)) + 1 * (g'(G(1))^2 * g''(G(1))$$

Next we denote the variance of the total population by $\sigma^2_t$, and since $G''(1) = \sigma^2_t + \mu^2_t - \mu_t$ we get:

$$\sigma^2_t = \frac{\mu - \mu^2 + \sigma^2}{(1 - \mu)^3} - \frac{1}{(1 - \mu)^2} + \frac{1}{1 - \mu} = \frac{\mu - \mu^2 + \sigma^2 - 1 + \mu + (1 - \mu)^2}{(1 - \mu)^3} = \frac{\mu - \mu^2 + \sigma^2 - 1 + \mu + 1 - 2\mu + \mu^2}{(1 - \mu)^3} = \frac{\sigma^2}{(1 - \mu)^3}$$

And we are done. □

2.3 Some combinatorics

Having all the theory about branching processes needed for our problem, we can see that it unfortunately gives us little insight into how to construct a probability function for the process (since it is often very hard to formulate it for a branching process). So in order to formulate a probability function for our problem we are going to need some additional theory of combinatorics, which we turn to next.

2.3.1 Ordered integer partitions

Since our problem results in summing a number of discrete uniform variables, a question of importance is in how many ways we can sum to an integer $n$ using $k$ integers between 1 and $m$, where the order of the integers is important. These numbers will be denoted by $p_m(k, n)$, where $p$ here stands for partition. There appears to be no simple analytic function to represent these numbers, however using generating functions the following is true:

**Theorem 4**: Letting $m$ and $k$ be fixed, we have: $\sum_{n=1}^{\infty} p_m(k, n)x^n = (\sum_{j=1}^{m} x^j)^k$

**Proof**: The right hand side can be interpreted as follows: the $m$ x:es within the parenthesis correspond to the $m$ different integers we can use when summing to $n$. The $k$ parentheses then corresponds to adding $k$ integers together. Multiplying out the x:es result in their exponents (the j:s) being added together to form a sum corresponding to the number $n$ they add up to, and then summing the x:es with equal n:s will give us the $p_m(k, n)$ values. □
Example: Say \( m=3 \) and \( k=2 \), then we have:

\[
\left( \sum_{j=1}^{3} x^j \right)^2 = (x^1 + x^2 + x^3)^2 =
\]

\[
= x^{1+1} + x^{1+2} + x^{1+3} + x^{2+1} + x^{2+2} + x^{2+3} + x^{3+1} + x^{3+2} + x^{3+3} =
\]

\[
= (x^{1+1}) + (x^{1+2} + x^{2+1}) + (x^{1+3} + x^{2+2} + x^{3+1}) + (x^{2+3} + x^{3+2}) + (x^{3+3}) =
\]

\[
= 1x^2 + 2x^3 + 3x^4 + 2x^5 + 1x^6
\]

The different exponents correspond to the different ways to sum to the possible \( n \):s (2-6):

\[
2 = 1 + 1, \quad 3 = 1 + 2 = 2 + 1, \quad 4 = 1 + 3 = 2 + 2 = 3 + 1,
\]

5 = 2 + 3 = 3 + 2, \quad 6 = 3 + 3

So we have one way to sum two integers between one and three to 2 or 6, two ways to sum to 3 or 5, and three ways to sum to 4 (and 0 ways for all other \( n \)).

We observe the following:

**Corollary 4.1:** \( n < k \rightarrow p_m(k, n) = 0 \), and \( n > mk \rightarrow p_m(k, n) = 0 \).

**Proof:** We sum \( k \) integers between one and \( m \), so if we sum \( k \) ones and this is higher than \( n \), there is no way to sum to \( n \) (all other sums exceed it), and similarly if we sum \( k \) integers \( m \) and this is lower than \( n \) there is no way to sum to \( n \) (it exceeds all other sums). \( \square \)

A MATLAB script for finding \( p_m(k, n) \) is given in Appendix A. Some more results about \( p_m(k, n) \) are given in Appendix C.

### 2.3.2 Full K-ary trees

A full K-ary tree is a rooted\(^1\) tree where each vertex has exactly 0 or \( K \) children. We can easily see that if a full K-ary tree has \( k \) internal vertices (the vertices with \( K \) children), it must have \( Kk + 1 \) vertices in total \((K - 1)k + 1\) of which are leaves (the vertices with 0 children). A question of interest for our branching process problem then is how many different full K-ary trees with \( k \) internal vertices exist.

**Theorem 5:** The number of full K-ary trees with \( k \) internal vertices is \( \frac{1}{(K-1)k+1} \binom{Kk}{k} \).

These are called Fuss-Catalan numbers.

**Proof for \( K=2 \):** We present a proof for \( K=2 \) only, for one example of a general proof we refer to Aval (2008) instead. In this case we have:

\(^1\) A rooted tree is a tree where one vertex has been designated the "root" from which the now directed tree stems. For our problem then the root corresponds to the founding member of the branching process.
Called the Catalan numbers, the k\:t of which we will refer to as \( C_k \), with the corresponding full 2-ary trees referred to as full binary trees. From above we have that a full binary tree of \( k \) internal vertices has \((2 - 1)k + 1 = k + 1\) leaves.

First we observe that \( C_0 = C_1 = 1 \), which is trivial as there is only one full binary tree of no or one internal vertex (i.e. the tree with just the root, and the tree with the root plus its two children), which agrees with our formula.

In order to prove the formula for general \( k > 0 \), first we observe that one of our internal vertices must be the root, which can be any of these \( k \) vertices. Let us say that if the root is the \( j \):th internal vertex then \( j \)-1 internal vertices make up the leftmost offspring tree, and the remaining \( k-j \) internal vertices make up the rightmost offspring tree (so we “count” the leftmost tree first, followed by the root, and then the rightmost tree), see Figure 1. Then there are \( C_{j-1} \) possible trees for the leftmost offspring tree, and \( C_{k-j} \) for the rightmost offspring tree. But this means that the \( k \):th Catalan number must satisfy:

\[
C_k = C_0C_{k-1} + C_1C_{k-2} + \cdots + C_{k-1}C_0 = \sum_{j=1}^{k} C_{j-1}C_{k-j}
\]

\textbf{Figure 1:} Counting the internal vertices down and up from left to right means this full binary tree of three internal vertices has the root as its third internal vertex. As such the leftmost subtree has 3-1=2 internal vertices, and the rightmost subtree has 3-3=0 internal vertices. There must be \( C_2C_0 \) such trees with the root as its third internal vertex. The full binary trees of three internal vertices with the root as their first internal vertex are the reflection of these \( (C_0C_2 \text{ in total}) \). Finally if the root is the second internal vertex both its children must be internal vertices, resulting in \( C_1C_1 \) such trees. Summing gives the total number of full binary trees of three internal vertices.

Now, the generating function for the Catalan numbers is:

\[
c(x) = \sum_{k=0}^{\infty} C_k x^k
\]

Squaring we get:

\[
c(x)^2 = (C_0C_0) + (C_0C_1 + C_1C_0)x + (C_0C_2 + C_1C_1 + C_2C_1)x^2 + \cdots
\]

Which we see corresponds to the equation for \( C_k \) above, so:

\[
c(x)^2 = C_1 + C_2x + C_3x^2 + \cdots
\]
And:

\[ c(x) = C_0 + c(x)^2 x = 1 + c(x)^2 x \]

I.e. a quadratic function of \( c(x) \), which gives us:

\[
c(x) = \frac{1}{2x} \pm \sqrt{\frac{1}{4x^2} - \frac{1}{x}} = \frac{1 \pm \sqrt{1 - 4x}}{2x} \rightarrow \{c(0) = C_0 = 1\} \rightarrow
\]

\[
\rightarrow c(x) = \frac{1 - \sqrt{1 - 4x}}{2x} \left( \frac{1 - \sqrt{1 - 4x}(1 + \sqrt{1 - 4x})}{2x(1 + \sqrt{1 - 4x})} \right) = \frac{2}{1 + \sqrt{1 - 4x}}
\]

As the other solution gives \( C_0 \neq 1 \), and must be a false solution. By Newton’s generalized binomial theorem we have:

\[
\sqrt{1 - 4x} = (1 - 4x)^{1/2} = 1 - \frac{1}{2!} 4x + \frac{1}{2!} \left( -\frac{1}{2} \right) (4x)^2 - \frac{1}{3!} \left( -\frac{1}{2} \right) \left( -\frac{3}{2} \right) (4x)^3 + \frac{1}{4!} \left( -\frac{1}{2} \right) \left( -\frac{3}{2} \right) \left( -\frac{5}{2} \right) (4x)^4 + \cdots =
\]

\[
= 1 - \frac{1}{2!} 2x - \frac{1}{2!} 4x^2 - \frac{3}{3!} 8x^3 - \frac{3 \cdot 5}{4!} 16x^4 - \frac{3 \cdot 5 \cdot 7}{5!} 32x^5 - \cdots
\]

So:

\[
c(x) = 1 + \frac{1}{2!} 2x + \frac{3}{3!} 4x^2 + \frac{5 \cdot 3}{4!} 8x^3 + \frac{7 \cdot 5 \cdot 3}{5!} 16x^4 + \frac{9 \cdot 7 \cdot 5 \cdot 3}{6!} 32x^5 + \cdots
\]

We observe that our numerators are factorials missing their even numbers. Now, since

\[
2^1 \cdot 1! = 2, 2^2 \cdot 2! = 4 \cdot 2, 2^3 \cdot 3! = 6 \cdot 4 \cdot 2, 2^4 \cdot 4! = 8 \cdot 6 \cdot 4 \cdot 2 \text{ and so on, we get:}
\]

\[
c(x) = 1 + \frac{1}{2!} \left( \frac{2!}{(1!)^2} \right) x + \frac{1}{3!} \left( \frac{4!}{(2!)^2} \right) x^2 + \frac{1}{4!} \left( \frac{6!}{(3!)^2} \right) x^3 + \frac{1}{5!} \left( \frac{8!}{(4!)^2} \right) x^4 + \frac{1}{6!} \left( \frac{10!}{(5!)^2} \right) x^5 + \cdots
\]

But this is the same as:

\[
c(x) = \sum_{k=0}^{\infty} \frac{1}{k+1} \binom{2k}{k} x^k \rightarrow c_k = \frac{1}{k+1} \binom{2k}{k}
\]

As wanted. □
3. Results

Now with some background in branching processes and combinatorics we can start solving our problem. First we will create a solution for an Eon die (six sided die, sixes aren’t counted and result in two new die-rolls), and then generalize our results. Throughout this chapter we will denote our sums $S$ of a random number of random variables, where the number of variables is decided by a branching process, by:

$$S = \sum_{i=1}^{D_\Omega-D} X_{\alpha,i} + \sum_{i=1}^{D} X_{\beta,i}$$

Which with a specification of the distribution of the $X$:es, $D$ and $D_\Omega$ gives us all information we need for finding a distribution of $S$. A short clarification of the variables in this summary is given below.

Table 1: A summary of the variables in our model of the random sums $S$.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>The random sum of primary interest.</td>
</tr>
<tr>
<td>$X$</td>
<td>R.V.:s denoting the individual contribution of the dice to the sum $S$.</td>
</tr>
<tr>
<td>$X_\alpha$</td>
<td>R.V.:s denoting the individual contribution of the dice that don't branch.</td>
</tr>
<tr>
<td>$X_\beta$</td>
<td>R.V.:s denoting the individual contribution of the dice that do branch.</td>
</tr>
<tr>
<td>$D_\Omega$</td>
<td>R.V. denoting the total number of dice.</td>
</tr>
<tr>
<td>$D$</td>
<td>R.V. denoting the total number of branching dice.</td>
</tr>
</tbody>
</table>

It is also useful to think of the distribution of $X_\beta$ and $X_\alpha$ as the conditional distributions of a R.V. $X$ (the normal die's distribution without branching) given whether the corresponding die branches or not, i.e.:

$$P(X_{\beta,i} = k) = P(X_i = k|\text{die } i \text{ branches})$$
$$P(X_{\alpha,i} = k) = P(X_i = k|\text{die } i \text{ does not branch})$$

which gives us:

$$P(X_i = k) = (1 - p)P(X_{\alpha,i} = k) + pP(X_{\beta,i} = k)$$

where $p$ is the probability of branching.
3.1 The single Eon die

For the single Eon die we have the following model for our sum $S$:

$$S = \sum_{i=1}^{D+1} X_{\alpha,i}$$

where the $X_{\alpha,i}$s are i.i.d. R.V.:s with discrete uniform distribution taking values between 1 and 5. We will from now on denote such a discrete uniform distribution that takes all integer values between $a$ and $b$ by $UD(a, b)$. Since we start with one die and get $D$ branchings this means that we will have $2D+1-D=D+1$ dice that don’t branch ($= D_{\alpha} - D$), and since we don’t count the dice that branch this is all we need for the sum.

Now we would like to find a distribution for $S$, so first we make the observation that if we end up with more than $n$ non-branching dice (dice not showing six after all branchings have taken place) then the probability that $S=n$ is zero (from Corollary 4.1). So:

$$P(S = n | D \geq n) = 0$$

since we then have to sum at least $n+1$ dice, each taking value at least 1. Similarly, from Corollary 4.1, since each die summed can take value at most 5, we have:

$$P(S = n | 5D + 1 < n) = P\left( S = n \bigg| D < \left\lfloor \frac{n}{5} - 1 \right\rfloor \right) = 0$$

(Where we round up because $D$ has to be integer valued.) Using these results, and the general result that:

$$P(S = n) = \sum_{k=-\infty}^{\infty} P(S = n \cap D = k) = \sum_{k=-\infty}^{\infty} P(S = n | D = k) P(D = k)$$

We observe that the conditional distribution of $S$ given $D$ gives us an expression for the probability that $S=n$ that requires summing a finite number of terms, i.e.:

$$P(S = n) = \sum_{k=\left\lceil \frac{n}{5} - 1 \right\rceil}^{n-1} P(S = n | D = k) P(D = k)$$

So the problem simplifies to finding $P(S=n | D=k)$ and $P(D=k)$, which we will turn to next.

3.1.1 $P(D=k)$

For the distribution of $D$ we observe that if we get exactly $k$ branchings, then $k$ dice will branch (probability $1/6$ each) and $k+1$ dice will not branch (probability $5/6$ each), so the probability distribution must satisfy:

$$P(D = k) = C(k) \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{k+1}$$
where \( C(k) \) counts the number of ways to get exactly \( k \) branchings. Now we observe that our resulting dice must be in the form of a full binary tree with \( k \) internal vertices and \( k+1 \) leaves, with each generation in their respective level of the tree, see Figure 2. The number of such trees with \( k \) internal vertices is given by the Catalan number, so:

\[
P(D = k) = \frac{(2k)!}{(k+1)!k!} \left( \frac{1}{6} \right)^k \left( \frac{5}{6} \right)^{k+1}
\]

\textbf{Figure 2:} An example of a throw with an Eon die resulting in \( D=3 \) and \( S=11 \).

3.1.2 \( P(S=n|D=k) \)

For this probability we observe that given that we have \( k \) branching dice and \( k+1 \) non-branching dice we are summing \( k+1 \) dice taking integer values between 1 and 5 with equal probability \( \frac{1}{5} \), so the probability has to conform to:

\[
P(S = n|D = k) = p_5(k+1, n) \left( \frac{1}{5} \right)^{k+1}
\]

Where \( p_5(k, n) \) is the number of ordered integer partitions of \( n \) with \( k \) natural numbers between 1 and 5, which we can acquire by using Theorem 5. As such we finish our expression of \( P(S=n|D=k) \) here\(^2\).

\(^2\) Note that the probability function of \( S|D=k \) is simply the probability function of the sum of \( k+1 \) i.i.d. R.V.:s with a \( UD(1,5) \) distribution.
3.1.3 $P(S=n)$

Having a formula for our respective distribution, we can now write a formula for $P(S=n \cap D=k)$:

$$P(S = n \cap D = k) = P(S = n|D = k)P(D = k) =$$

$$= p_5(k + 1, n) \left( \frac{1}{5} \right)^{k+1} \frac{(2k)!}{(k+1)!k!} \left( \frac{5}{6} \right)^k \left( \frac{1}{6} \right)^{k+1} = p_5(k + 1, n) \frac{(2k)!}{(k+1)!k!} \left( \frac{1}{6} \right)^{2k+1}$$

which is the probability of $2k + 1$ given numbers for $2k + 1$ dice times the number of full binary trees on $k$ internal vertices times the number of ways to sum $k+1$ natural numbers between 1 and 5 to $n$. So the probability we looked for is:

$$P(S = n) = \sum_{k=\lceil \frac{n-1}{2} \rceil}^{n-1} p_5(k + 1, n) \frac{(2k)!}{(k+1)!k!} \left( \frac{1}{6} \right)^{2k+1}$$

since all other values for $k$ result in $P(S = n|D = k)$ being zero (as mentioned above). A bar graph of the distribution of $S$ up to $n=30$ is given in Figure 3.

![Probability function for the single Eon die](image-url)

*Figure 3:* Bar graph of the probabilities for the single Eon die up to $n=30$. 

17
3.1.4 Mean and variance of the Eon die

In order to find the mean and variance of our distribution, we use the general results:

\[ E(S) = E_D(ES(S|D)), \quad Var(S) = E_D(Var_S(S|D)) + Var_D(ES(S|D)) \]

For the mean and variance of the conditional distribution \( S|D \), which has probability \( P(S=n|D=k) \), we note that this is the summation of \( k+1 \) i.i.d. R.V.'s with \( UD(1,5) \) distribution, that is \( k+1 X_\alpha \) (from the model for our die). We have:

\[
E(X_\alpha) = \frac{5 + 1}{2} = 3, \quad Var(X_\alpha) = \frac{(5 - 1 + 1)^2 - 1}{12} = 2
\]

by usual rules for discrete uniform distributions. So:

\[
E_S(S|D) = E_X \left( \sum_{i=1}^{D+1} X_{\alpha,i} \right) = 3(D + 1), \quad Var_S(S|D) = Var_X \left( \sum_{i=1}^{D+1} X_{\alpha,i} \right) = 2(D + 1)
\]

Now, instead of using \( E(D) \) and \( Var(D) \) we use the variable \( D_\Omega \) denoting the total number of dice when no new branchings occur (i.e. the final number of dice). Since we start with one die and each branching results in two new dice, we must have:

\[
D_\Omega = 2D + 1 \iff D = \frac{D_\Omega - 1}{2} \iff D + 1 = \frac{D_\Omega + 1}{2}
\]

So:

\[
E(D + 1) = E \left( \frac{D_\Omega + 1}{2} \right) = \frac{1}{2} (E(D_\Omega) + 1), \quad Var(D + 1) = Var \left( \frac{D_\Omega + 1}{2} \right) = \frac{1}{4} Var(D_\Omega)
\]

We observe that \( D_\Omega \) simply counts the total population of our branching process, which has the following mean and variance for generation one:

\[
\mu = E(Z_1) = \frac{5}{6} * 0 + \frac{1}{6} * 2 = \frac{1}{3}
\]

\[
\sigma^2 = Var(Z_1) = E(Z_1^2) - E(Z_1)^2 = \left( \frac{5}{6} * 0^2 + \frac{1}{6} * 2^2 \right) - \left( \frac{1}{3} \right)^2 = \frac{2}{3} - \frac{1}{9} = \frac{5}{9}
\]

And since \( \mu < 1 \) and \( \sigma^2 < \infty \) we can use Corollary 3.1 to get:

\[
E(D_\Omega) = \frac{1}{1 - \mu} = \frac{1}{1 - 1/3} = \frac{3}{2}, \quad Var(D_\Omega) = \frac{\sigma^2}{(1 - \mu)^3} = \frac{5/9}{(1 - 1/3)^3} = \frac{15}{8}
\]

Which gives us:

\[
E(S) = E(3(D + 1)) = \frac{3}{2}(E(D_\Omega) + 1) = \frac{3}{2} \left( \frac{3}{2} + 1 \right) = \frac{15}{4} = 3.75
\]
Generalizations

In this part we will generalize the results of the Eon die to any \( m \)-sided die, with any number \( K \geq 2 \) new dice when branching, either when counting or not counting the branching dice, and finally when branching occurs over some cut-off value \( m_\alpha \). We will generalize from the Eon die in this order. A MATLAB script for finding the probabilities in the most general case is given in Appendix A. The special case when \( K=1 \), perhaps the most common type of exploding die, is treated further in its own section (where it will be assumed that we count the branching dice). Another special section will be devoted to finding the distribution when we only count number of successes (when a die rolls over some number), not the actual sum of the dice. Finally we will give some short notes how to deal with the probability function for the sum when we have multiple starting dice (using convolutions and normal approximation).

3.2.1 Any \( m \)-sided die

Model:

\[
S = \sum_{i=1}^{D+1} X_{\alpha,i}
\]

i.e. the same as in the Eon-die case, with the exception that \( X_{\alpha} \) now have a \( UD(1, m-1) \) distribution (since the final side \( m \) branches and is not counted). We will assume \( m \geq 2 \) (to avoid the trivial case without any randomness). Now, instead of six sided dice we have dice with \( m \)-sides, meaning that each side has probability \( 1/m \), and instead of 5 possible values to sum to \( n \) we now have \( m-1 \) values (since we still assume the branching value is not counted). Furthermore (from Corollary 4.1) to be able to sum to \( n \), we must have:

\[(m - 1)(k + 1) \geq n \leftrightarrow k \geq \frac{n}{m - 1} - 1\]

So the probability function becomes:

\[
P(S = n) = \sum_{k=\left\lfloor \frac{n}{m-1} \right\rfloor}^{n-1} p_{m-1}(k + 1, n) \left( \frac{2k}{(k+1)! k!} \right)^{2k+1} \left( \frac{1}{m} \right)^{2k+1}
\]

The probabilities for the Eon die, together with a corresponding four-, and eight-sided die is shown in Figure 4 (up to \( n=30 \)):
3.2.2 K new dice when branching

Model:

\[ S = \sum_{i=1}^{(K-1)D+1} X_{a,i} \]

Since if we get \(D\) branchings, we get \(KD + 1\) dice in total, \(D\) of which branched and aren’t included in the sum, resulting in \((K - 1)D + 1\) non-branching dice summing to \(S\). We will assume \(K \geq 2\) (\(K = 1\) will be treated in a separate section). Now, if we get \(k\) branchings, we will have \(kk + 1\) dice in total, \(k\) of which branches and \((K - 1)k + 1\) of which are used when summing to \(S\). But now instead of forming a full bipartite tree, the dice form a full \(K\)-partite tree, the number of which were given by Theorem 5. Furthermore since we are using values between 1 and \(m-1\) to sum to \(n\) we must have:

\[(K - 1)k + 1 \leq n \iff k \leq \frac{n - 1}{K - 1}\]

\[(m - 1)((K - 1)k + 1) \geq n \iff k \geq \frac{n}{m - 1} - 1 = \frac{n - m + 1}{(m - 1)(K - 1)}\]

So our probability function becomes:
Figure 5 below shows the probabilities up to n=30 for the Eon die, and six sided dice that instead result in 3 or 4 new dice when a six is rolled:

$$P(S = n) = \sum_{k=\left\lceil \frac{n-1}{m(K-1)} \right\rceil}^{\left\lfloor \frac{n-m+1}{(m-1)(K-1)} \right\rfloor} p_{m-1}((K-1)k + 1, n) \frac{1}{(K - 1)k + 1} \binom{Kk}{k} \left( \frac{1}{m} \right)^{k+1}$$

**Figure 5:** Probabilities up to n=30 for six sided exploding die that result in, from up to down, two (the Eon die), three, and four new dice when a six is rolled (sixes aren't counted).

### 3.2.3 Counting the branching dice

**Model:**

$$S = \sum_{i=1}^{(K-1)D+1} X_{\alpha,i} + \sum_{i=1}^{D} m = \sum_{i=1}^{(K-1)D+1} X_{\alpha,i} + mD$$

Since we now also count the branching dice which all take value m (as this is the only side that results in a branching). Now we need:
\[(K - 1)k + 1 + mk \leq n \iff k \leq \frac{n - 1}{K + m - 1}\]

\[(m - 1)((K - 1)k + 1) + mk \geq n \iff k \geq \frac{n - m + 1}{(m - 1)K + 1}\]

Furthermore, since a part equal to \(mk\) of the sum \(n\) is already taken care of by the \(k\) branching dice, the remaining dice must sum to \(n - mk\) instead, so our probability function when counting the dice that branches becomes:

\[P(S = n) = \sum_{k=\left\lfloor \frac{n - m + 1}{(m - 1)K + 1} \right\rfloor}^{\left\lfloor \frac{n - 1}{K + m - 1} \right\rfloor} p_{m-1}( (K - 1)k + 1, n - mk ) \frac{1}{(K - 1)k + 1} \binom{Kk}{k} \left( \frac{1}{m} \right)^{kk+1}\]

Figure 6 below shows the probabilities up to \(n=30\) for the Eon die, and a corresponding die where the sixes are counted.

---

*Figure 6:* Probabilities up to \(n=30\) for, from up to down, the Eon die, and a corresponding die where the sixes are counted.
3.2.4 Branching occurs over some value \( m_{\alpha} \)

For this generalization, we will look at the probability distribution when a branching occurs not just for one value of the die, but for values strictly above a value \( m_{\alpha} \) (so that \( m_{\alpha} \) is the highest value that doesn't result in a branching). It will be assumed that \( 0 < m_{\alpha} < m \) so that at least one but not all sides result in a branching. We will make one generalization for when we don't count the dice that branches, and one for when we do count them.

Beginning with finding a common formula for \( D \) we observe that the probability of a branching is simply \( (m - m_{\alpha})/m \), and the probability of no branching is \( m_{\alpha}/m \), so the probability for exactly \( k \) branchings becomes:

\[
P(D = k) = \frac{1}{(K - 1)k + 1} \left( \frac{m_{\alpha}}{m} \right)^k \left( \frac{m - m_{\alpha}}{m} \right)^{(K - 1)k + 1}
\]

3.2.4.1 Branching dice not counted

Model (branching dice are not counted):

\[
S = \sum_{i=1}^{(K-1)D+1} X_{\alpha,i}
\]

where the non-branching dice now have a \( UD(1, m_{\alpha}) \) distribution. Now, when we don't count the dice that branches, we simply sum the remaining dice showing value at most \( m_{\alpha} \), so we get:

\[
P(S = n | D = k) = p_{m_{\alpha}}((K - 1)k + 1, n) \left( \frac{1}{m_{\alpha}} \right)^{(K - 1)k + 1}
\]

And we find our limits of summation by the inequalities:

\[
(K - 1)k + 1 \leq n \iff k \leq \frac{n-1}{K-1}, \quad m_{\alpha}((K - 1)k + 1) \geq n \iff k \geq \frac{n-m_{\alpha}}{m_{\alpha}(K-1)}
\]

So after simplifying the probability becomes:

\[
P(S = n) = \sum_{k=\left[\frac{n-m_{\alpha}}{m_{\alpha}(K-1)}\right]}^{\left[\frac{n-1}{K-1}\right]} p_{m_{\alpha}}((K - 1)k + 1, n) \left( \frac{m - m_{\alpha}}{m} \right)^k \left( \frac{m_{\alpha}}{m} \right)^{(K - 1)k + 1}
\]

Figure 7 below shows the Eon die together with corresponding dice where branching occurs over values two, three, and four.
Figure 7: Probabilities up to n=30 for six sided not counting the branching dice that result in a branching to to new dice when rolling over, from up to down, two (infinite expected value), three, four, and five (the Eon die).

3.2.4.2 Branching dice counted

Model (branching dice are counted):

\[ S = \sum_{i=1}^{(K-1)D+1} X_{\alpha,i} + \sum_{i=1}^{D} X_{\beta,i} \]

where the non-branching dice now have a \( UD(1, m_\alpha) \) and the branching dice have a \( UD(m_\alpha + 1, m) \) distribution. In this case where we are counting the branching dice we have to further break up \( P(S=n|D=k) \), since the values of the branching dice now is not entirely decided by them having branched. Denoting the sum of the branching dice by \( n_\beta \), and the variable for this by \( S_\beta \) we must have that:

\[ P(S = n|D = k) = P(S = n|S_\beta = n_\beta \land D = k)P(S_\beta = n_\beta|D = k) \]

Now, since the \( k \) branching dice take values \( m_\alpha + 1, m_\alpha + 2, \ldots, m - 1, m \) which we shall sum to \( n_\beta \), there have to be as many ways to do this as there are numbers of ways to use \( k \) values between 1 and \( m - m_\alpha \) to sum to \( n_\beta - m_\alpha k \), so assuming \( k \geq 1 \):

\[ P(S_\beta = n_\beta|D = k) = p_{m-m_\alpha}(k, n_\beta - m_\alpha k) \left( \frac{1}{m - m_\alpha} \right)^k \]
And the other probability becomes:

\[ P(S = n | S_\beta = n_\beta \land D = k) = p_{m_\alpha} \binom{(K - 1)k + 1, n - n_\beta}{(K - 1)k + 1} \left( \frac{1}{m_\alpha} \right)^{(K - 1)k + 1} \]

since a part equal to \( n_\beta \) of the sum \( n \) is taken care of by the branching dice. After simplifying we get:

\[ P(S = n \land S_\beta = n_\beta \land D = k) = \]

\[ = \frac{p_{m_\alpha} \binom{(K - 1)k + 1, n - n_\beta}{(K - 1)k + 1} p_{m - m_\alpha} \binom{k, n_\beta - m_\alpha k}{k} \left( \frac{1}{m} \right)^{k + 1}}{Kk} \]

when \( p_{m - m_\alpha} \binom{k, n_\beta - m_\alpha k}{Kk} \) plays a role (i.e. when \( k > 0 \)), with that term equal to 1 otherwise. To find \( P(S = n) \) we have to sum over the possible values, where we now have two sums to take care of, one pertaining to \( k \) and the other to \( n_\beta \). Since the \((K - 1)k + 1\) non-branching dice take values at least 1 and at most \( m_\alpha \), and the \( k \) branching dice take values at least \( m_\alpha + 1 \) and at most \( m \) we have to have:

\[(K - 1)k + 1 + (m_\alpha + 1)k \leq n \leftrightarrow k \leq \frac{n - 1}{K + m_\alpha} \]

\[m_\alpha ((K - 1)k + 1) + mk \geq n \leftrightarrow k \geq \frac{n - m_\alpha}{m_\alpha(K - 1) + m}\]

And given that we have \( k \) branchings, the sum of the branching dice have to conform to:

\[n_\beta \leq mk \land n_\beta \leq n - (K - 1)k - 1\]

\[n_\beta \geq (m_\alpha + 1)k \land n_\beta \geq n - m_\alpha((K - 1)k + 1)\]

Where the latter terms comes from the fact that the sum of the non-branching dice must be at least \((K - 1)k + 1\) and at most \( m_\alpha((K - 1)k + 1)\). So:

\[ P(S = n) = \]

\[= \left( \frac{1}{m}, if \ k = 0 \right) \]

\[\sum_{k=\left[\frac{n - n_\beta}{m + m_\alpha(K - 1)}\right]}^{\left[\frac{n - n_\beta}{m + m_\alpha(K - 1)}\right]} P(S = n \land S_\beta = n_\beta \land D = k), if k \neq 0 \]

Where \( P(S = n \land S_\beta = n_\beta \land D = k) \) is as above. If we add the property that \( p_m(0,0) = 1 \) for all \( m \) then the lower term suffices for all \( k \) \( (\geq 0) \).

Figure 8 below shows the probabilities for six sided dice that result in two new dice when branching, count the branching dice, and result in a branching over values two to five.
Figure 8: Probabilities up to n=30 for six sided that that branch to two new dice, count the branching dice, and result in a branching over values, from up to down, two, three, four, and five.

All our generalizations now complete, Appendix A presents a MATLAB script for finding the probabilities for any exploding dice of these types, with any number of starting dice (the procedure for which is shortly presented in section 3.4). Appendix B presents a MATLAB script that simulates rolling exploding dice of these types.

3.2.5 Mean and variance in the general case

For our most general case of branching die we have the following model:

$$S = \sum_{i=1}^{(K-1)D+1} X_{\alpha,i} + I_\beta \sum_{i=1}^{D} X_{\beta,i}$$

Where the $X$'s are independent and $X_{\alpha} \sim U(1, m_\alpha)$, $X_{\beta} \sim U(m_\alpha + 1, m)$, and where $I_\beta$ is one if we count the branching dice and zero if we don't count them. We will now find an expression for the mean and variance in our most general case of exploding dice which were found in section 3.2.4. We will assume that the process dies out, i.e. that the mean $\mu$ of generation one is lower than 1, i.e. that $K(m - m_\alpha)/m < 1$. Furthermore, the variance $\sigma^2$ of generation one becomes $K^2(m - m_\alpha)/m - K^2((m - m_\alpha)/m)^2 = K^2(m - m_\alpha)(m_\alpha/m)$ so it will be finite for any given $K < \infty$. As such we can use Corollary 3.1 to first find an expression for the mean and variance of the branching process, and then use these to find corresponding expressions for the sum.
Whether we do or do not count the branching dice will not affect the mean and variance of the total number of dice in the process. In the section above we gave an expression of the mean and variance of the first generation, and so the mean and variance of the total number of dice becomes (from Corollary 3.1):

\[
E(D_\alpha) = \frac{1}{1 - \mu} = \frac{1}{1 - K \frac{(m - m_\alpha)}{m}} = \frac{m}{m - K(m - m_\alpha)}
\]

\[
Var(D_\alpha) = \frac{\sigma^2}{(1 - \mu)^2} = \frac{K^2 \left( \frac{m - m_\alpha}{m} \right) \frac{m_\alpha}{m}}{\left( 1 - K \frac{(m - m_\alpha)}{m} \right)^2} = \frac{K^2 m(m - m_\alpha)m_\alpha}{(m - K(m - m_\alpha))^3}
\]

Furthermore since we have that:

\[
S|D = \sum_{i=1}^{(K-1)D+1} X_{\alpha,i} + I_\beta \sum_{i=1}^{D} X_{\beta,i}
\]

And by property of the discrete uniform distribution:

\[
E(X_{\alpha,i}) = \frac{m_\alpha + 1}{2}, \quad Var(X_{\alpha,i}) = \frac{(m_\alpha - 1 + 1)^2 - 1}{12} = \frac{m_\alpha^2 - 1}{12}
\]

\[
E(X_{\beta,i}) = \frac{m + m_\alpha + 1}{2}, \quad Var(X_{\beta,i}) = \frac{(m - (m_\alpha + 1) + 1)^2 - 1}{12} = \frac{(m - m_\alpha)^2 - 1}{12}
\]

So:

\[
E(S) = E(E(S|D)) = E\left( ((K - 1)D + 1)E(X_{\alpha,i}) + I_\beta DE(X_{\beta,i}) \right) =
\]

\[
= ((K - 1)E(D) + 1) \frac{m_\alpha + 1}{2} + I_\beta E(D) \frac{m + m_\alpha + 1}{2} =
\]

\[
= \frac{m_\alpha + 1 + ((K - 1)(m_\alpha + 1) + I_\beta (m + m_\alpha + 1)) E(D)}{2}
\]

\[
Var(S) = Var(E(S|D)) + E(Var(S|D)) =
\]

\[
= Var\left( ((K - 1)D + 1)E(X_{\alpha,i}) + I_\beta DE(X_{\beta,i}) \right) + E\left( ((K - 1)D + 1)Var(X_{\alpha,i}) + I_\beta DVar(X_{\beta,i}) \right) =
\]

\[
= \frac{(K - 1)^2(m_\alpha + 1)^2 + I_\beta (m + m_\alpha + 1)^2}{4} Var(D) - \frac{m_\alpha^2 - 1 + ((K - 1)(m_\alpha^2 - 1) + I_\beta ((m - m_\alpha)^2 - 1)) E(D)}{12} =
\]
Now we use that:

\[ D_\alpha = KD + 1 \iff D = \frac{D_\alpha - 1}{K} \]

So we get the mean:

\[
E(S) = \frac{m_\alpha + 1 + ((K-1)(m_\alpha + 1) + \beta(m + m_\alpha + 1))}{2} E\left(\frac{D_\alpha - 1}{K}\right)
\]

\[
= \frac{K(m_\alpha + 1) + ((K-1)(m_\alpha + 1) + \beta(m + m_\alpha + 1))}{2K} \left(\frac{m}{m - K(m - m_\alpha)} - 1\right)
\]

\[
= \frac{(m_\alpha + 1)(m - K(m - m_\alpha)) + ((K-1)(m_\alpha + 1) + \beta(m + m_\alpha + 1))(m - m_\alpha)}{2(m - K(m - m_\alpha))}
\]

\[
= \frac{m_\alpha(m_\alpha + 1) + \beta(m + m_\alpha + 1)(m - m_\alpha)}{2(m - K(m - m_\alpha))} = \frac{(1 - \beta)m_\alpha(m_\alpha + 1) + \beta m(m + 1)}{2(m - K(m - m_\alpha))}
\]

And variance:

\[
Var(S) = \frac{m_\alpha^2 - 1}{12} + \frac{(K-1)(m_\alpha^2 - 1)}{12} E\left(\frac{D_\alpha - 1}{K}\right) + \frac{(K-1)^2(m_\alpha + 1)^2}{4} Var\left(\frac{D_\alpha - 1}{K}\right)
\]

\[
+ \beta \left(\frac{((m - m_\alpha)^2 - 1)}{12} E\left(\frac{D_\alpha - 1}{K}\right) + \frac{(m + m_\alpha + 1)^2}{4} Var\left(\frac{D_\alpha - 1}{K}\right)\right)
\]

\[
= \frac{m_\alpha^2 - 1}{12K} + \frac{(K-1)(m_\alpha^2 - 1)}{12K} E(D_\alpha) + \frac{(K-1)^2(m_\alpha + 1)^2}{4K^2} Var(D_\alpha)
\]

\[
+ \beta \left(\frac{((m - m_\alpha)^2 - 1)}{12K} (E(D_\alpha) - 1) + \frac{(m + m_\alpha + 1)^2}{4K^2} Var(D_\alpha)\right)
\]
And upon inserting the mean and variance of the total number of dice and simplifying we get:

\[
\begin{align*}
\text{Var}(S) &= \frac{(1 - l_\beta)m_\alpha(m_\alpha^2 - 1) + l_\beta m(m^2 - 3m_\alpha(m - m_\alpha) - 1)}{12(m - K(m - m_\alpha))} \\
&\quad + \frac{m(m - m_\alpha)m_\alpha((K - 1)^2(m_\alpha + 1)^2 + l_\beta(m + m_\alpha + 1)^2)}{4(m - K(m - m_\alpha))^3}
\end{align*}
\]

3.3 The special case when K=1

We will begin with the case where just the highest value branches, since this is very common and gives us a nice distribution. In this section where \( K = 1 \) we will assume that we are counting the branching dice, since \( S \) simply get the discrete uniform distribution between 1 and \( m_\alpha \) otherwise (since in that case if we roll over this we roll a new die and count that one instead, and so on until we roll low enough). The probability of section 3.2.4 where we count the branching dice works in these cases as well (since we never had to assume \( K \neq 1 \) in these calculations), and so can be simplified to these formula under the assumptions about \( K \) and \( m_\alpha \) here. But since this is such a common type of exploding die, it warrants showing the special simplified distribution.

For \( m_\alpha = m - 1 \) our model summary becomes:

\[
S = \sum_{i=1}^{(1-1)^D+1} X_{\alpha,i} + \sum_{i=1}^D m = X_{\alpha,1} + mD
\]

where \( X_{\alpha} \sim UD(1, m - 1) \). Now, we observe that since we always get a new die when rolling \( m \), there is no way to sum to \( n = mk \) for any \( k \in \mathbb{N} \). Furthermore, for \( n = mk + j \) where \( 1 \leq j \leq m - 1 \), we must get \( k \) dice that rolls \( m \) and branches, followed by a dice that does not roll \( m \), but rolls \( j \) instead. So our distribution becomes:

\[
P(S = n = mk + j) = \left( \frac{1}{m} \right)^k \left( \frac{n}{m} \right)^j, \quad \text{for } 1 \leq j \leq m - 1
\]

(Where the last equalities work since we exclude the values where \( m|n \).) Furthermore having \( k \) branchings means that we must roll \( k \) values \( m \) followed by a value that is not \( m \), so \( D \) has a geometrical distribution, i.e.:

\[
P(D = k) = \left( \frac{1}{m} \right)^k \left( 1 - \frac{1}{m} \right)
\]
Now, if we (more generally) have a value $0 < m_\alpha < m$ over which a branching occurs we get the following model:

$$S = X_{\alpha,i} + \sum_{i=1}^{D} X_{\beta,i}$$

Where $X_{\alpha} \sim UD(1, m_\alpha)$ and $X_{\alpha} \sim UD(m_\alpha + 1, m)$. The probability that $D = k$ must still conform to the geometrical distribution:

$$P(D = k) = \left(1 - \frac{m_\alpha}{m}\right)^k \left(\frac{m_\alpha}{m}\right)^k = \left(\frac{m - m_\alpha}{m}\right)^k \left(\frac{m_\alpha}{m}\right)^k$$

However, the formula for $S$ is not as simple as before in this case, and once again we have to use $P(S = n | D = k)$. From section 3.2.4 we have:

$$P(S = n | D = k) =$$

$$= p_{m-m_\alpha}(k, n_\beta - m_\alpha k) \left(\frac{1}{m - m_\alpha}\right)^k p_{m_\alpha}((K-1)k + 1, n - n_\beta) \left(\frac{1}{m_\alpha}\right)^{(K-1)k+1} =$$

$$= \{K = 1\} = p_{m-m_\alpha}(k, n_\beta - m_\alpha k) p_{m_\alpha}(1, n - n_\beta) \left(\frac{1}{m - m_\alpha}\right)^k \left(\frac{1}{m_\alpha}\right)$$

So:

$$P(S = n \cap S_\beta = n_\beta \cap D = k) = p_{m-m_\alpha}(k, n_\beta - m_\alpha k) p_{m_\alpha}(1, n - n_\beta) \frac{1}{m^{k+1}}$$

We get our limits of summation by:

$$1 + (m_\alpha + 1)k \leq n \iff k \leq \frac{n - 1}{m_\alpha + 1}, \quad m_\alpha + mk \geq n \iff k \geq \frac{n - m_\alpha}{m}$$

$$n_\beta \geq (m_\alpha + 1)k \cap n_\beta \geq n - m_\alpha, \quad n_\beta \leq mk \cap n_\beta \leq n - 1$$

So the probability distribution for $S$ becomes:

$$P(S = n) =$$

$$= \sum_{k=\frac{n-m_\alpha}{m}}^{\left\lfloor \frac{n-1}{m_\alpha+1} \right\rfloor} \left(1, if \ k = 0\right) p_{m-m_\alpha}(k, n_\beta - m_\alpha k) p_{m_\alpha}(1, n - n_\beta), if \ k \neq 0$$

Or once again just the lower one if we add that $p_m(0,0) = 1$ for all $m$. 

30
3.3.1 Mean and variance when K=1

Since this is such a common special case of exploding die, we will find a special formula for the mean and variance. By property of the geometric distribution, we have:

\[ E(D) = \frac{(m - m_\alpha)/m}{m_\alpha/m} = \frac{m - m_\alpha}{m_\alpha}, \quad Var(D) = \frac{(m - m_\alpha)/m}{(m_\alpha/m)^2} = \frac{m(m - m_\alpha)}{m_\alpha^2} \]

And since we have that \( X_\alpha \) is \( UD(1, m_\alpha) \) distributed and \( X_\beta \) is \( UD(m_\alpha + 1, m) \) distributed:

\[ E(X_\alpha) = \frac{m_\alpha + 1}{2}, \quad Var(X_\alpha) = \frac{(m_\alpha + 1 + 1)^2 - 1}{12} = \frac{m_\alpha^2 - 1}{12} \]
\[ E(X_\beta) = \frac{m_\alpha + 1 + m}{2}, \quad Var(X_\beta) = \frac{(m - m_\alpha - 1 + 1)^2 - 1}{12} = \frac{(m - m_\alpha)^2 - 1}{12} \]

So finally since \( S = X_\alpha + \sum^D_{i=1} X_\beta,i \) we get:

\[ E(S) = E(E(S|D)) = E(X_\alpha) + E(D)E(X_\beta) = \frac{m_\alpha + 1}{2} + \frac{(m_\alpha + 1 + m)(m - m_\alpha)}{2m_\alpha} = \]
\[ = \frac{m(m + 1)}{2m_\alpha} \]
\[ Var(S) = Var(E(S|D)) + E(Var(S|D)) = \]
\[ = Var\left( E(X_\alpha) + DE(X_\beta) \right) + E\left( Var(X_\alpha) + DVar(X_\beta) \right) = \]
\[ = \left( \frac{m_\alpha + 1 + m}{2} \right)^2 Var(D) + \frac{m_\alpha^2 - 1}{12} + \frac{(m - m_\alpha)^2 - 1}{12} E(D) = \]
\[ = \frac{(m_\alpha + 1 + m)^2}{4} \cdot \frac{m(m - m_\alpha)}{m_\alpha^2} + \frac{m_\alpha^2 - 1}{12} + \frac{(m - m_\alpha)^2 - 1}{m_\alpha} \]
\[ = \frac{3(m_\alpha + 1 + m)^2(m(m - m_\alpha))}{12m_\alpha} + \frac{(m_\alpha^2 - 1)m_\alpha^2}{12m_\alpha} + \frac{(m - m_\alpha)^2 - 1)m_\alpha(m(m - m_\alpha))}{12m_\alpha} \]
\[ = \frac{m(m + 1)(4m_\alpha(m - 1) + 3m(m + 1) - 6m_\alpha^2)}{12m_\alpha^2} \]

When \( m_\alpha = m - 1 \) these formula simplify to:

\[ E(S) = \frac{m(m + 1)}{2(m - 1)} \]
\[ Var(S) = \frac{m(m + 1)(m^2 + 7m - 2)}{12(m - 1)^2} \]
3.4 Counting number of successes

Sometimes, rather than counting the exploding dice, there is merely some value over which a roll is denoted a "success", and one is only interested in counting the total number of successes. When there is also some chance of branching, our model summary becomes:

\[ S = \sum_{i=1}^{(K-1)D+1} X_{\alpha,i} + \sum_{i=1}^{D} 1 = \sum_{i=1}^{(K-1)D+1} X_{\alpha,i} + D \]

Where the \( X_{\alpha} \)'s takes value 1 with probability \( p \), and 0 with probability \( 1-p \) (a Bernoulli distribution). It is assumed that a branching is also a success. As before we get:

\[ P(D = k) = \frac{1}{(K-1)k + 1} \binom{Kk}{k} \left( \frac{m - m_\alpha}{m} \right)^k \left( \frac{m_\alpha}{m} \right)^{(K-1)k+1} \]

But since we now only count successes we must have that \( P(S = n|D = k) \) has a binomial distribution:

\[ P(S = n|D = k) = \binom{(K-1)k + 1}{n-k} p^{n-k} (1-p)^{(K-1)k+1-(n-k)} = \]

\[ = \binom{(K-1)k + 1}{n-k} p^{n-k} (1-p)^{Kk+1-n} \]

since a sum equal to \( k \) is already taken care of by the branching dice. This gives us:

\[ P(S = n \cap D = k) = \]

\[ = \frac{1}{(K-1)k + 1} \binom{Kk}{k} \left( \frac{m - m_\alpha}{m} \right)^k \left( \frac{m_\alpha}{m} \right)^{(K-1)k+1} \binom{(K-1)k + 1}{n-k} p^{n-k} (1-p)^{Kk+1-n} = \]

\[ = \frac{(Kk)!}{k! (n-k)! (Kk + 1 - n)!} \left( \frac{m - m_\alpha}{m} \right)^k \left( \frac{m_\alpha}{m} \right)^{(K-1)k+1} p^{n-k} (1-p)^{Kk+1-n} \]

The limits of summation become:

\[ k \leq n, \quad k + (K-1)k + 1 \geq n \iff k \geq \frac{n-1}{K} \]

So:

\[ P(S = n) = \]

\[ = \sum_{k=\frac{n-1}{K}}^{\binom{(K-1)k}{n-1}} \frac{(Kk)!}{k! (n-k)! (Kk + 1 - n)!} \left( \frac{m - m_\alpha}{m} \right)^k \left( \frac{m_\alpha}{m} \right)^{(K-1)k+1} p^{n-k} (1-p)^{Kk+1-n} \]
Note that we did not require the probability of success $p$ to be decided by the sides of a die. However, if it is decided by such we must have that $p = m_s/m_\alpha$, where $m_s$ is the number of sides that result in a success but not in a branching.

The mean and variance is found by (provided the mean of generation one is less than one and that the variance is finite):

$$E(X_\alpha) = p, \quad Var(X_\alpha) = p - p^2 = p(1 - p)$$

$$E(D_\alpha) = \frac{m}{m - K(m - m_\alpha)}, \quad Var(D_\alpha) = \frac{K^2 m (m - m_\alpha) m_\alpha}{(m - K(m - m_\alpha))^2}$$

$$E(D) = E\left(\frac{D_\alpha - 1}{K}\right) = \frac{m}{K (m - K(m - m_\alpha))} - \frac{1}{K} = \frac{m - m_\alpha}{m - K(m - m_\alpha)}$$

$$Var(D) = \frac{m(m - m_\alpha)m_\alpha}{(m - K(m - m_\alpha))^3}$$

$$E(S) = E(E(S|D)) = E\left(\left(\left(K - 1\right)D + 1\right)E(X_\alpha) + D\right) = (pK + 1 - p)E(D) + p = \frac{m - m_\alpha(1 - p)}{m - K(m - m_\alpha)}$$

$$Var(S) = +Var(E(S|D)) + E(Var(S|D)) = (pK + 1 - p)^2Var(D) + E\left(\left(\left(K - 1\right)D + 1\right)Var(X)\right) = \frac{(pK + 1 - p)^2 m(m - m_\alpha)m_\alpha}{(m - K(m - m_\alpha))^3} + (K - 1)Var(X)E(D) + Var(X) = \frac{(pK + 1 - p)^2 m(m - m_\alpha)m_\alpha}{(m - K(m - m_\alpha))^3} + \frac{m_\alpha p(1 - p)}{m - K(m - m_\alpha)}$$

3.5 Starting with multiple dice

The convolution of two probability distributions has the following form: If $Z = X + Y$, where $X$ and $Y$ are independent, then:

$$P(Z = z) = P(X + Y = z) = \sum_{k=-\infty}^{\infty} P(X = k)P(Y = z - k)$$

In our case, since we sum similar exploding dice, when $X$ and $Y$ each represents a single die they will have the same distribution and will have probability zero for $k \leq 0$ and $z - k \leq 0$ respectively (since their sum must be at least one). Furthermore when summing extra dice
beyond the two, the probability corresponding to A starting dice must have probability zero for \(k < A\). We will only give an outline for how to go about in the general case when having two dice, and give an example for the Eon dice, and the case when \(K = 1\) and \(m_A = m - 1\), although the process for all types of exploding dice and adding extra dice beyond the two is similar. When \(X\) and \(Y\) correspond to one exploding die, the limits of summation give us:

\[
P(Z = z) = \sum_{k=1}^{z-1} P(X = k)P(Y = z - k), \quad \text{for } z \geq 2, \quad P(Z) = 0 \text{ otherwise}
\]

Furthermore, when \(X\) corresponds to A starting dice and \(Y\) is an extra die, we have:

\[
P(Z = z) = \sum_{k=A}^{z-1} P(X = k)P(Y = z - k), \quad \text{for } z \geq A + 1, \quad P(Z) = 0 \text{ otherwise}
\]

**Example:** The probability distribution for two Eon dice becomes (for \(n \geq 2\)):

\[
P(S = n) = \sum_{j=1}^{n-1} \left( \sum_{k=\left\lceil \frac{j}{n} \right\rceil - 1}^{j-1} p_{5}(k+1,j) \frac{(2k)!}{(k+1)!k!} \left( \frac{1}{6} \right)^{2k+1} \right) \left( \sum_{k=\left\lceil \frac{n-j-1}{n} \right\rceil - 1}^{n-j-1} p_{5}(k+1,n-j) \frac{(2k)!}{(k+1)!k!} \left( \frac{1}{6} \right)^{2k+1} \right)
\]

where the \(n:s\) in the expressions for the single Eon dice are substituted for \(j\) and \(n-j\) respectively. As can be seen there is no way to break out any part of the expressions from the parentheses (perhaps because we have no closed form for expressing \(p_m(k,n)\)), and so the summation is best done by computer. The same appears to be true for most types of exploding dice.

**Example:** In the case when \(K = 1\) and \(m_A = m - 1\) we get the following expression for \(n \geq 2\):

\[
P(S = n) = \sum_{j=1}^{n-1} l_{m,n}(j) \left( \frac{1}{m} \right)^{\left\lfloor \frac{j}{m} \right\rfloor} \left( \frac{1}{m} \right)^{\left\lfloor \frac{n-j}{m} \right\rfloor} = \sum_{j=1}^{n-1} l_{m}(j,n) \left( \frac{1}{m} \right)^{\left\lfloor \frac{j}{m} \right\rfloor + \left\lfloor \frac{n-j}{m} \right\rfloor}
\]

Where \(l_{m}(j,n)\) equals zero if \(m \mid j + n - j\) and one otherwise. Since this means that the terms when either \(j/m = j/m\) or \(\left\lfloor (n-j)/m \right\rfloor = (n-j)/m\) are discarded, the remaining exponents simply add up to \(\left\lfloor (n-1)/m \right\rfloor + 1\) regardless of \(j\). This is because the exponential terms change value at the same \(j\) so that when one increases by one the other one decreases by one. Also when \(j=1\) the first term in the exponent is equal to 1 and the other is equal to \(\left\lfloor (n-1)/m \right\rfloor\), from where we get our equality. Furthermore we add \(n-1\) terms together, and if \(m\mid n\) then \(\left\lfloor (n-1)/m \right\rfloor\) terms of these will be zero (when both \(j/m = j/m\) and \(\left\lfloor (n-j)/m \right\rfloor = (n-j)/m\) and if \(m \nmid n\) then \(2\left\lfloor (n-1)/m \right\rfloor\) terms will be zero (when the same is true for one part of the exponent). So the probability becomes:
\[ P(S = n) = \begin{cases} 
\left( n - 1 - \left\lfloor \frac{n - 1}{m} \right\rfloor \right) \left( \frac{1}{m} \right)^{\left\lfloor \frac{n - 1}{m} \right\rfloor + 1}, & \text{if } m \mid n \\
\left( n - 1 - 2 \left\lfloor \frac{n - 1}{m} \right\rfloor \right) \left( \frac{1}{m} \right)^{\left\lfloor \frac{n - 1}{m} \right\rfloor + 1}, & \text{if } m \nmid n 
\end{cases} \]

For \( n \geq 2 \). Furthermore, the exponent in the upper probability can be substituted for \( n/m + 1 \) and the multiplying factor by \( n - n/m \) if preferred (since \( m \mid n \)).

As can be seen, the expressions for our exploding dice seem to get increasingly more difficult the more dice we start with, since there appears to be no satisfying way to simplify our expressions after using convolutions. So if one has a large number of starting dice the distributions become quite cumbersome. As such we now turn to the normal approximation. The MATLAB script in Appendix A can be used for dealing with any amount of starting dice however (barring round-off error for higher values).

### 3.6 Normal approximations

As could be seen from the figures in our generalizations of the exploding dice, changing values \( m, K, m_\alpha, \) and whether or not the branching dice where counted does not appear to cause the distribution for the single exploding dice to get much closer to normal.

However, by the classical central limit theorem we have that the distribution for the sum of multiple R.V.:s approaches a normal distribution as the number of R.V.:s increases; assuming they are i.i.d., and with finite mean and variance (which we found the requirements for in section 3.2.5). So the value when we found an expression for the mean and variance of the different types of single exploding dice lied in being able to use normal approximations for a larger number of starting dice. In this section we will use the script in Appendix A, which we got from the final generalized formula from section 3.2 together with the formula for convolutions of section 3.5, to look at how the distribution of the sum of exploding dice approaches the normal distribution as the starting number of dice increases. We will only look at examples when the mean and variance is finite for the single die. However, since all our distributions are significantly skewed (the majority of the density lies in the lower part, with small but existent probabilities for very large values), and since most games that use exploding dice rarely use more than a small number of starting dice, it is perhaps rare that the normal approximation will be good enough in normal usage. As such these approximations are primarily interesting on a theoretical level. We will illustrate the asymptotic distributions using three types of dice: a die with \( m = 10, m_\alpha = 9, K = 1, l_\beta = 1 \); one like the first but with \( m_\alpha = 6 \); and a final that is like the first but with \( K = 4 \). The other manipulations, for different \( m \) and not counting the branching dice, do not do much to change the asymptotic behavior.

Starting with the first type of dice, Figure 10 shows the distribution up to \( n=200 \). As can be seen the skew results in lower values around the mean being more common, and higher values around the mean being less common than predicted by the normal distribution. Even so, at 5 starting dice the distributions have begun to look roughly normal. However depending on how correct an approximation one wants, the approximation might not be good enough until at roughly 20 starting dice. Nevertheless, the normal approximation seems to get acceptable quite quickly (compared to some other more skewed distributions).
For our second example, where $m_{\alpha} = 6$, seen in Figure 11, things look a little less promising, as now the normal approximation with 5 starting dice seems less acceptable. However, with 20 starting dice the approximation once again appears to work. The reason behind the somewhat slower asymptotic normality is of course due to the greater skewness of this distribution, with higher values now more common.
For our final example with $K = 4$, presented in Figure 12, however, things are much worse, as not even starting with 20 dice results in a distribution even closely resembling the normal distribution. Once again this has to be attributed to the skewness, which here is even greater since every branching means four more dice to add to the sum (compared to just one before). We see that for this example, the normal approximation cannot be justified unless for quite a large number of starting dice (many more than shown here).

![Figure 12](image)

*Figure 12*: Probabilities up to $n=200$ of, from up to down, 1, 5, and 20 starting exploding dice with $m = 10$, $m_\alpha = 9$, $K = 4$, $l_\beta = 1$ (blue bars), with corresponding normal distribution (red line).

Some general guidelines for the distribution of our exploding dice to approach the normal distribution more quickly appears to be to have small $K$, as well as to some extent bigger $m_\alpha$ compared to $m$ (so that $m - m_\alpha$, the number of sides that result in a branching, is small). Indeed, the perhaps most common types of exploding dice have $K = 1$ and $m_\alpha = m - 1$ (and $l_\beta = 1$), and so for these the normal approximation seems useful even at a medium amount of starting dice.
4. Discussion

The aim of this thesis was to do a mathematical analysis of exploding dice, commonly used in (primarily table top) role playing games. In particular we wanted to find an analytical way to write their probability distribution, as well as finding their mean and variance. We found that, assuming the die has finitely many sides, the only requirements for the mean and variance to be finite was that \( K(m - m_u)/m < 1 \) (which is the requirement for the dice throw to "die out" and not go on forever), and under these assumptions we found the mean and variance. Furthermore, we found the probability distribution for the most general case of (non-trivial) exploding dice (without using this assumption), with special cases for when \( K = 1 \) and when we only counted successes. When we had found our expression for the probability distribution for single starting dice we finally used convolutions and the normal approximation to deal with multiple starting dice.

As such this thesis succeeded in its primary goal of finding a way to derive the probabilities. However, our expressions are perhaps not strictly analytical, due to the inclusion of the \( p_m(k, n) \) function. Nevertheless, since we found a simple (if for higher values somewhat time consuming) way to find these values, our expressions are perfectly usable for finding the exact probabilities (in particular with the aid of a computer).

This being done, we will end this thesis with some general discussion about counting the sum of a random number of R.V.'s, primarily where the number is decided by a branching process, and when this might be a useful model of more general real life processes.

4.1 Summing a random number of random variables

When we summed a random number of random variables, where the number is decided by a branching process, in this thesis, the reason we could find a probability distribution was twofold. Firstly, the number of offspring of each individual was either 0 or \( K \), which meant that we could use the Fuss-Catalan numbers to find a distribution for \( P(D = k) \), which completely decided the total number of individuals of the branching process (i.e. the total number of individuals when we started with one individual is \( KD + 1 \)). This allowed for \( P(S = n) \) to be expressed by this probability and \( P(S = n | D = k) \). The ability to find a closed formula for this second probability was the second factor why our procedure worked. As such it is perhaps not always so easy to find an expression for the sum of a random number of random variables decided by a branching process. However, even if this is not so, the use of generating functions in section 2.2 about branching processes meant that we could find expressions for the moments of a branching process. Since:

\[
S = \sum_{i=1}^{N} X_i
\]

This means that we can find expressions for the moments of the sum for the total process, or for the process in or up to some generation \( n \). For example, if we denote \( S_n \) to be the sum of the \( n \):th generation we must have:

\[
E(S_n) = E(E(S_n | Z_n)) = E \left( \sum_{i=1}^{z_n} E(X_i) \right)
\]

38
And if all the $X$:es have the same distribution this becomes:

$$E(S_n) = E(Z_n)E(X)$$

With extra terms if we have different distribution for the $X$:es (such as in our dice case where the distribution between dice that did branch and dice that did not branch differ).

Another useful property that was hinted upon in our dice examples was the fact that we could use normal approximations for a sum dependent upon a branching process. If our sum depends on a large number of starting individuals that form a branching process, then we can find a normal approximation of the sum up to some generation $n$. This is because this will always have finite mean and variance provided that the $X$:es and the function that decides the offspring distribution of each member (i.e. $Z_1$) has this, while that of the total population might not (if $E(Z_1) = \mu \geq 1$).

To widen our view and look at when a sum of a random number of variables decided by a branching process might be a useful model for natural phenomena, we must first look at when branching processes are used. That is, when we are looking at a population of individuals, who each are assumed to propagate with independent and identical probability. Our sum then must correspond to an effect this population has on something in the natural world. An example could be if we are interested in the amount of a solution a strand of bacteria will consume up to some time (corresponding to some maximum generation). The amount each bacteria will consume during its lifetime would decide the $X$:es distribution, and a branching probability would decide the number of offspring of the bacteria. Due to the above, if we start with a larger number of bacteria we could expect the total solution consumed by the strand up to a generation $n$ to be roughly normally distributed. So if we have an idea about the offspring distribution and the amount of solution consumed by each bacteria, we could use the results of section 2.2 to find the mean and variance of the population size up to generation $n$, and then use this to find the mean and variance of the consumed amount up to this generation. Whether the model holds could then be tested. Another example that might be modeled in this way is how quickly a newly introduced predator might extinguish some indigenous species.

When modeling this way it might be reasonable to assume that the individuals who produce offspring might "affect more" (i.e. contribute more to the sum $S$) than individuals who don't produce offspring. This means that the $X$:es will have different distribution whether their individual produce offspring or not, requiring us to condition upon whether an individual produces offspring (such as in our dice example).

Furthermore, if the $X$:es have a distribution that is closed under addition (such as the poisson or the normal distribution), then the probability $P(S = s|N = n)$ might give us some insight into how the process develops, for example if:

$$S = \sum_{i=1}^{N} X_i$$

and the $X$:es are i.i.d. with $Po(\lambda)$ distribution we will have:

$$P(S = s|N = n) = \frac{(n\lambda)^s}{s!}e^{-n\lambda}$$
Similarly, if the X:es have i.i.d. \( N(\mu, \sigma^2) \) distribution we get:

\[
f(s|N = n) = \frac{1}{\sqrt{2n\sigma^2\pi}} e^{-\frac{(s-n\mu)^2}{2n\sigma^2}}
\]

(Note that the distribution of the X:es - and so the sum S - is not required to be discrete.) The same sort of formula holds if we only look at or up to some generation. If we can also find a function for \( P(N = n) \) the problem of finding the distribution of \( S \) is easy. Although in the case of branching processes this need not be so, the conditional distributions might nevertheless be valuable in understanding the effect the population \( N \) has on the factor \( S \). Especially together with the formula from section 2.2 about how the mean number of individuals change over generations.

This thesis looked at a special case of this more general problem, and due to the important factor that the offspring was always either 0 or \( K \) we were able to find a formula for the distribution of \( S \). The results in this thesis are surely welcome for anyone who has played table top role playing games with some type of exploding dice, and who have wanted to find the probability that a roll will succeed or fail. However, the broader subject of summing a random number of random variables, where the number is decided by a branching process, is of course also interesting more generally. In this final section we have tried hinting on some ways in which such a model might arise, and how the results of this thesis relate to this more general problem.
5. References


Appendix A

MATLAB script for finding the probability distribution of any number of any type of exploding die in the generalizations.

```
function density = ExpDice(N,m,ma,K,Ib,A)
% Finds the probabilities for exploding dice.
% N = The highest probability term counted.
% m = The number of sides of the dice.
% ma = The highest number not resulting in a branching. Default: m-1
% K = The number of new dice when a branching occurs. Default: 1
% Ib = 1 if the branching dice are counted, 0 otherwise. Default: 1
% A = The number of starting dice. Default: 1
% N and m need to be specified.
% Output: a matrix with the probabilities for 1 to N in the columns, with
% the different starting dice from 1 to A in the rows.
if nargin<2
    error('N and m need to be specified.');
elseif nargin==2
    ma=m-1; K=1; Ib=1; A=1;
elseif nargin==3
    K=1; Ib=1; A=1;
elseif nargin==4
    Ib=1; A=1;
elseif nargin==5
    A=1;
end
density=zeros(A,N);
Pma= Intpart(ma,N);
Pma= Pma';
Pmma= Intpart(m-a,N);
Pmma= Pmma';
for n=1:N
    if Ib==1
        for k= ceil((n-ma)/(m+ma*(K-1))): floor((n-1)/(K+ma))
            for nb=max((ma+1)*k,n-ma*((K-1)*k+1)): min(m*k,n-(K-1)*k-1)
                if k==0
                    density(1,n)=density(1,n)+1/m;
                else
                    density(1,n)=density(1,n)+Pma((K-1)*k+1,n-nb)*Pmma(k,nb-ma*k)*nchoosek(K*k,k)/((K-1)*k+1)*(1/m)^(K*k+1);
                end
            end
        end
    elseif Ib==0
        for k= ceil((n-ma)/(ma*(K-1))): floor((n-1)/(K-1))
            density(l,n)=density(l,n)+Pma((K-1)*k+1,n-nb)*Pmma(k,nb-ma*k)*nchoosek(K*k,k)/((K-1)*k+1)*(1/m)^(K*k+1);
        end
    end
end
for i=2:A
    for n=i:N
        for k=i-1:n-1
            density(i,n)=density(i,n)+density(i-1,k)*density(1,n-k);
        end
    end
end
```
The ordered integer partitions in the beginning utilizes the following separate MATLAB script:

```matlab
function pm = Intpart(m,n)
% Creates a matrix of the number of ways of summing k integers between 1
% and m to values between 1 and n (where the order of the integers is
% important). The columns correspond to the different k and the rows
% correspond to the different n.
% The code uses the fact that these numbers can be found by polynomial
% multiplication of k parentheses of m x:es of degree 1 to m inside.
% When the parentheses are multiplied the degrees add up in such a way
% that the number of created x:es of degree n is equal to this number
% (since we get all possible ways to sum to n this way).

x=ones(1,m);
x=[x 0]; % Creates a vector representing the polynomial in MATLAB.

pm=zeros(n,n); % The output matrix.
for i=1:min(m,n)
    pm(i,1)=1;
end

for k=2:n
    c=x;
    c=conv(c,x); % Polynomial multiplication with k parentheses.
    C=zeros(1,length(c));
    for i=1:length(c)
        C(i)=c(length(c)-i+1);
    end
    C=C(2:length(C)); % Switches the order of the vector (for simplicity).

    for i=1:min(length(C),n)
        pm(i,k)=C(i);
    end % Puts each number in their correct place.
end
```
Appendix B

function density = ExpDiceEst(N,m,ma,K,Ib,A)
% Simulates rolling exploding dice.
% N = The number of rolls.
% m = The number of sides of the dice.
% ma = The highest value of the die that doesn't result in a branching.
% Default: ma=m-1
% K = The number of new dice that a branching results in. Default: K=1
% Ib = 1 if the branching dice are counted, 0 otherwise. Default: Ib=1
% A = The number of starting dice. Default: A=1
% N and m need to be specified.
% Output: a vector where entry j denotes the number of rolls that resulted
% in the sum j. If N=1 the function only returns the value of the roll.
if nargin<2
    error('N and m need to be specified.');
elseif nargin==2
    ma=m-1; K=1; Ib=1; A=1;
elseif nargin==3
    K=1; Ib=1; A=1;
elseif nargin==4
    Ib=1; A=1;
elseif nargin==5
    A=1;
end
density=zeros(1,m*K*A*3);
for i=1:N
    D=A;
    S=0;
    while D>0
        s=randi(m);
        if s>ma
            D=D+K;
            if Ib==1
                S=S+s;
            end
        else
            S=S+s;
        end
        D=D-1;
    end
    if S>length(density)
        dplus=zeros(1,S-length(density));
        density=[density dplus];
    end
    density(S)=density(S)+1;
end
if N==1
    density=S;
end
44
Appendix C

We always assume that \( m \geq 1 \).

**Theorem C.1:** Denote by \( p_{(a,b)}(k, n) \) the number of ordered integer partitions of \( n \) using \( k \) positive integers between \( a \) and \( b \) so that \( 1 \leq a \leq b \). Then:

\[
\sum_{n=1}^{\infty} p_{(a,b)}(k, n)x^n = \left(\sum_{j=a}^{b} x^j\right)^k
\]

**Proof:** This is just a more general case of Theorem 4, and follows immediately from that proof. \( \square \)

**Corollary C.1.1:** \( p_{(a,b)}(k, n) = p_{b-a+1}(k, n - (a - 1)k) \)

**Proof:** This fact was used in section 3.2.4.2, we have:

\[
\sum_{n=1}^{\infty} p_{(a,b)}(k, n)x^n = \left(\sum_{j=a}^{b} x^j\right)^k = \left(x^{a-1} \sum_{j=1}^{b-a+1} x^j\right)^k = x^{(a-1)k} \left(\sum_{j=1}^{b-a+1} x^j\right)^k
\]

\[
= x^{(a-1)k} \sum_{n=1}^{\infty} p_{b-a+1}(k, n - (a - 1)k)x^n
\]

So since the \( x \):es before the sum does not change the constant in front of the \( x \):es to the power of \( n \) we have our equality. \( \square \)

Due to the equality above, we see that we will only need to work with \( p_m(k, n) \) in order to get proofs about these more general ordered integer partitions using values between \( a \) and \( b \).

**Theorem C.2:** \( p_n(k, n) = \binom{n - 1}{k - 1} \)

**Proof:** Say we have \( n \) number of dots in a row so that there are \( n-1 \) empty spaces between them. Out of these empty spaces, choose \( k-1 \) of them to draw lines separating the dots so that they are divided into \( k \) groups. Then each group has at least one dot and at most \( n \) dots. Let the number of dots in a group correspond to an integer value equal to the number of dots. The numbers between 1 and \( n \) from the \( k \) groups will then add up to \( n \). There are \( n-1 \) over \( k-1 \) ways of choosing the places for the lines, from which we get our equality. \( \square \)

**Corollary C.2.1:** If \( k + m > n \) then \( p_m(k, n) = \binom{n - 1}{k - 1} \).

**Proof:** Write the inequality as \( k - 1 + m \geq n \) (which is the same thing as \( k, m, \) and \( n \) are all integers). The left hand side corresponds to summing \( k - 1 \) ones and one \( m \). So if this is at least as high as \( n \) then there is no way to sum to \( n \) using any value over \( m \). As such we must have that \( p_m(k, n) = p_n(k, n) \), from which we get our equality. \( \square \)
Theorem C.3: Define \( p_m(0,0) = 1 \), otherwise unless \( k \cap n > 0 \) define \( p_m(k,n) = 0 \), then for \( k \cap n > 0 \):

\[
p_m(k,n) = \sum_{j=1}^{m} p_m(k-1,n-j)
\]

Proof:

\[
\sum_{n=1}^{\infty} p_m(k,n)x^n = \left( \sum_{j=1}^{m} x^j \right)^k = \left( \sum_{j=1}^{m} x^j \right)^{k-1} \left( \sum_{j=1}^{m} x^j \right) = \\
(p_m(k,k-1)x^{k-1} + p_m(k,k)x^k + \cdots + p_m(k,m(k-1)x^{m(k-1)}) \left( \sum_{j=1}^{m} x^j \right)
\]

So in order to find \( p_m(k,n) \) we multiply the values corresponding to \( p_m(k-1,n-m) \) to \( p_m(k-1,n-1) \) in the left parenthesis with the corresponding x in the right parenthesis so the power becomes n (these are always defined by the definitions above, and the definition of \( p_m(0,0) = 1 \) ensures that \( p_m(1,1) = 1 \) which starts our recursion). Since all values in the right parenthesis have 1 as their constant in front of the x:es this multiplication doesn't change the values from the left parenthesis. So when we finally sum the x:es of the same power we just sum these \( p_m(k-1,n-m) \) to \( p_m(k-1,n-1) \) to get the value for \( p_m(k,n) \) in front of the x that is to the power of n in this way. □

Corollary C.3.1: Under the same definitions as for Theorem C.3 we have:

\[
p_m(k,n) = p_m(k-1,n-1) + p_m(k,n-1) - p_m(k-1,n-m-1)
\]

Proof: From Theorem C.3:

\[
p_m(k,n) = \sum_{j=1}^{m} p_m(k-1,n-j) , \quad p_m(k,n-1) = \sum_{j=1}^{m} p_m(k-1,n-j-1)
\]

So:

\[
p_m(k,n) = \sum_{j=1}^{m} p_m(k-1,n-j) + p_m(k,n-1) - \sum_{j=1}^{m} p_m(k-1,n-j-1)
\]

We see that the values from the left sum and the right sum all are the same except for the first value for the left sum and the last value from the right sum. So all these other values cancel out, from which we get our equality. □
**Theorem C.4**: For larger $k$-values we have:

$$p_m(k, n) \approx m^k \left( \Phi \left( \frac{n + 0.5 - k \frac{m+1}{2}}{\sqrt{k(m^2 - 1)/12}} \right) - \Phi \left( \frac{n - 0.5 - k \frac{m+1}{2}}{\sqrt{k(m^2 - 1)/12}} \right) \right)$$

**Proof**: In section 3.1.2 we found that the probability distribution of a sum of $k$ numbers of independent $UD(1, m)$ distributed random variables could be written as:

$$p(n) = p_m(k, n) \left( \frac{1}{m} \right)^k = \frac{p_m(k, n)}{m^k}$$

Which by the classical central limit theorem will be roughly normally distributed for larger $k$. But this means that for large $k$:

$$\frac{p_m(k, n)}{m^k} \approx \Phi \left( \frac{n + 0.5 - k \frac{m+1}{2}}{\sqrt{k(m^2 - 1)/12}} \right) - \Phi \left( \frac{n - 1 + 0.5 - k \frac{m+1}{2}}{\sqrt{k(m^2 - 1)/12}} \right) \Leftrightarrow$$

$$\Leftrightarrow p_m(k, n) \approx m^k \left( \Phi \left( \frac{n + 0.5 - k \frac{m+1}{2}}{\sqrt{k(m^2 - 1)/12}} \right) - \Phi \left( \frac{n - 0.5 - k \frac{m+1}{2}}{\sqrt{k(m^2 - 1)/12}} \right) \right)$$

Where 0.5 is the continuity correction factor, and the other terms (except $n$) are the mean and standard deviation of $k$ numbers of $UD(1, m)$ distributed R.V.'s. This gives us an approximation to $p_m(k, n)$ for larger $k$-values. Furthermore, since $p_m(k, n)$ is integer valued a perhaps more useful approximation is obtained by rounding to the nearest integer. $\Box$