www.simiode.org
SIMIODE Systemic Initiative for Modeling Investigations and Opportunities with Differential Equations

STUDENT VERSION

Stochastic Birth/Death/Immigration Processes

STATEMENT

We develop a mathematical model of a death and immigration process using m&m's as a stochastic process with the help of probability generating functions (pgf). We start with 50 m&m's in a bag.

The process: We start each step by shaking our bag of m&m's.

We empty the bag onto a plate.

The m&m's that land with their m showing up are considered to have died and are removed from the plate.

The m&m's without the m showing are considered to have survived the step and are put back into the bag.

The step ends by our putting an additional 10 "immigrating" m&m's into the bag.

We assume that the probability that an m&m will land with an m facing upwards is 1/2.

The question: What is the probability distribution for the number of m&m's that are alive at the end of the n^{th} step and how many m&m's should we expect to be alive at the end of the n^{th} step?

Solution.

The Random Variable: Let b_n be the random variable that counts the total number of m&m's in the bag at the end of the n^{th} step. Since we are starting with 50 m&m's we know that $P(b_0 = 50) = 1$.

Restating question in terms of b_n: Find the distribution of the random variable b_n, i.e. $P(b_n = k)$ for $k = 0, 1, 2, \ldots$ and its expected value $E(b_n) = \sum_{k=0}^{\infty} kP(b_n = k)$. Note that $\max b_n = 50 + 10^n$.

Method of solution. We can recursively find the distribution of b_n using probability generating functions.

Generally speaking, if b is a random variable that counts something, then the pgf of b will be the power series

$$B(x) = \sum_{k=0}^{\infty} P(b=k)x^k$$

See the appendix for more about the theory of pgf's.

Let's consider a single m&m that starts a particular step alive. If that m&m fails to survive that step it will contribute 0 to the end-of-step count. On the other hand, if that m&m survives, it will contribute 1. With this in mind, we define the random variable g as follows:

$$g = \begin{cases} 0, & \text{if that m\&m dies during the step;} \\ 1, & \text{if that m\&m survives the step.} \end{cases}$$

The probability that $g=0$ and $g=1$ are both $\frac{1}{2}$. So, the probability generating function of g is:

$$G(x) = \ P(g=0)\, x^0 + P(g=1)\, x^1 \ = \ \frac{1}{2} + \frac{1}{2}x \ = .5(1+x)$$

Since our initial conditions are that we start with 50 m&m's, the $P(b_0 = 50) = 1$, so the pgf for the distribution of b_0 is the monomial:

$$B_0(x) = x^{50}$$

Since all the m&m's are identical and independent, the pgf of the distribution of the number of m&m's to survive to the end of the n^{th} step (not counting the immigrants that arrive just before the end of the n^{th} step), will be $B_n(G(x))$. To include (and count) the 10 new immigrants we just multiply by x^{10} to get:

$$B_{n+1}(x) = B_n(G(x))\ x^{10}$$

Let's look for a pattern:

$$\begin{aligned}
B_0(x) &= x^{50} \\
B_1(x) &= B_0(G(x))\ x^{10} = (G(x))^{50}\ x^{10} \\
B_2(x) &= B_1(G(x))\ x^{10} = (G(G(x)))^{50}\ (G(x))^{10}\ x^{10} \\
B_3(x) &= B_2(G(x))\ x^{10} = (G^{(3)}(x))^{50}\ (G^{(2)}(x))^{10}\ (G^{(1)}(x))^{10}\ x^{10}
\end{aligned}$$

So, the pattern for $B_n(x)$ is clear. If we let $G^{(0)}(x) = x$, we can write $B_n(x)$ as:

$$B_n(x) = (G^{(n)}(x))^{50}\ (G^{(n-1)}(x))^{10}\ \cdots\ (G^{(2)}(x))^{10}\ (G^{(1)}(x))^{10}\ (G^{(0)}(x))^{10} \tag{1}$$

Our formula for $B_n(x)$, Equation (1), is obvious in retrospect. The pgf $(G^{(k)}(x))^j$ gives the distribution[1] for the number of m&m's that will be alive after k steps if we start with a cohort of j

[1] Focus on a single m&m that is "alive" in the current step. Let g_k be 1 if that m&m is alive k steps later, and 0 if it is not. Then $G^{(k)}(x)$ is the pgf for the distribution of g_k. Since the m&m's live and die independently of each other $(G^{(k)}(x))^j$ will be the pgf for how many m&m's in a cohort of j m&m's survive k steps.

m&m's. So the $(G^{(n)}(x))^{50}$ term tracks the original 50 m&m's and each of the $(G^{(k)}(x))^{10}$ terms tracks a cohort of immigrants.

See footnote[2] and appendix for more details on expectation E in terms of pgf's.

We can calculate the expectation of b_n, $E(B_n)$ using the following formulas for expectation in terms of pgf's. If G, G_v, and G_w are pgf's, then the following formulas hold:

$$E(G) = G'(1), \qquad E(G_v\, G_w) = E(G_v) + E(G_w), \qquad E(G^j) = j\, E(G), \qquad E(G^{(k)}) = (E(G))^k$$

So,

$$\begin{aligned}
E(B_n(x)) &= E\left((G^{(n)}(x))^{50} \;\; (G^{(n-1)}(x))^{10} \;\cdots\; (G^{(2)}(x))^{10} \;\; (G^{(1)}(x))^{10} \;\; (G^{(0)}(x))^{10}\right) \\
&= E\left((G^{(n)}(x))^{50}\right) + \sum_{j=0}^{n-1} E\left((G^{(j)}(x))^{10}\right) \\
&= 50\, E\left((G^{(n)}(x))\right) + \sum_{j=0}^{n-1} 10 E\left((G^{(j)}(x))\right) \\
&= 50\, (E(G))^n + 10 \sum_{j=0}^{n-1} E(G)^j
\end{aligned}$$

Since $G(x) = \frac{1}{2} + \frac{1}{2}x$ it follows that $E(G) = G'(1) = \frac{1}{2}$. So we have:

$$\begin{aligned}
E(B_n(x)) &= 50\, (E(G))^n + 10 \sum_{j=0}^{n-1} E(G)^j \\
&= 50 \left(\frac{1}{2}\right)^n + 10 \sum_{j=0}^{n-1} \left(\frac{1}{2}\right)^j \\
&= 50 \left(\frac{1}{2}\right)^n + 10 \sum_{j=0}^{n-1} \left(\frac{1}{2}\right)^j
\end{aligned} \tag{2}$$

We can simplify (2) using the geometric series formula

$$\sum_{j=0}^{n-1} p^j = \frac{1 - p^n}{1 - p}$$

Letting $p = \frac{1}{2}$ we get:

$$\sum_{j=0}^{n-1} p^j = \frac{1 - \left(\frac{1}{2}\right)^n}{1 - \frac{1}{2}} = 2\left(1 - \left(\frac{1}{2}\right)^n\right)$$

[2]Suppose that G is the generating function for a counting random variable g. Then the expectation of $g = E(g) = \sum_{k=0}^{\infty} kP(g = k) = \left.\frac{d}{dx}\, G(x)\right|_{x=1} = G'(1) = E(G)$, where the last equality can be taken as the definition of $E(G)$.

and so:

$$\begin{aligned} E(B_n(x)) &= 50\left(\frac{1}{2}\right)^n + 10\sum_{j=0}^{n-1}\left(\frac{1}{2}\right)^j \\ &= 50\left(\frac{1}{2}\right)^n + 10\left(2\left(1-\left(\frac{1}{2}\right)^n\right)\right) \\ &= 50\left(\frac{1}{2}\right)^n + 20\left(1-\left(\frac{1}{2}\right)^n\right) \end{aligned} \tag{3}$$

So, two things are clear regarding the expectation: (1) the initial population of 50 m&m's has only a transient effect on the expectation of b_n and, (2) regardless of the initial population size, the expected population size will eventually be 20. This doesn't mean that $\lim_{n\to\infty} b_n = 20$ since after all, at each step we are emptying our bag of m&m's and removing those that have their m facing upwards.

Calculating the coefficients of $B_n(x)$.

It is easy enough to calculate the coefficients of $G^{(n)}(x)$. First of all, $G(x) = .5 + .5x$ so it follows that $G^{(n)}(x) = (1-q) + qx$ for some $q \in [0, 1]$. But then $E(G^{(n)}) = q$ (since $E(G^{(n)}) = \frac{d}{dx} G^{(n)}(x)\big|_{x=1} = q$). On the other hand $E(G^{(n)}) = (E(G))^n = (.5)^n$. So $q = (.5)^n$ and $G^{(n)}(x) = (1-(.5)^n) + (.5)^n x$.

However there isn't such a nice compact formula for $B_n(x)$ itself and so it is more convenient to have a computer calculate the coefficients for us. However, keep in mind that $B_n(x)$ is a $50 + 10n$ degree polynomial.

Appendix I: A note on modeling the death and immigration process as a differential equation.

The m&m scenario can also be modeled recursively by

$$b(n+1) = \frac{1}{2}\, b(n) + 10 \quad \text{with IC} \quad b(0) = 50, \tag{4}$$

where $b(n)$ represents the expected number of m&m's to be alive at the end of the n^{th} step. Iterating Equation (4) yields the same expectations as Equation (3), i.e.

$$b(n) = 50\left(\frac{1}{2}\right)^n + 20\left(1-\left(\frac{1}{2}\right)^n\right)$$

Equation (4) can be turned into a difference equation by subtracting $b(n)$ from both sides yielding

$$b(n+1) - b(n) = -\frac{1}{2}\, b(n) + 10 \quad \text{with IC} \quad b(0) = 50,$$

which also gives the slope between $b(n)$ and $b(n+1)$ from which we get the differential equation

$$b' = -0.5b + 10 \quad \text{with IC} \quad b(0) = 50. \tag{5}$$

Although the DE in Equation (5) is separable, it is most easily solved by the method of undetermined coefficients by rewriting it as

$$b' + 0.5b = 10 \quad \text{with IC} \quad b(0) = 50. \tag{6}$$

from which it is obvious that the solution is $b(t) = ce^{-0.5t} + 20$, where we've replaced the discrete n by the continuous t. The IC of $b(0) = 50$ immediately tells us that $c = 30$ and so

$$b(t) = 30e^{-0.5t} + 20$$

which for small values of t, say $t < 10$ will slightly differ from the recursive solution. See Table 4. of [2] for a numerical comparison.

Appendix II: A Brief Introduction to Probability Generating Functions

Random Variables and Counts

Let v be a non-negative integer valued random variable. We can think of v as counting something. So we'll call such random variables "counts".

Example: We flip a coin. Let v count heads (H). So v can equal 0 or 1.

Example: We flip two coins. Let v count how many heads we get. So v can equal $0, 1,$ or 2.

Example: We roll one die (die is the singular of dice). Let v count how many dots are facing upwards. So v can equal $1, 2, 3, 4, 5,$ or 6.

Example: We roll a pair of dice. Let v count how many dots are facing upwards. So v can equal $2, 3, 4, \ldots, 11, 12$.

Example: We look through a telescope. Let v count how many stars we see. So v can equal $0, 1, 2, \ldots$

Example: We take a water sample. Let v count how many bacteria are present. So v can equal $0, 1, 2, \ldots$

Probability Generating Functions

If the random variable v is a count, we can represent its probability distribution as the power series $G_v(x)$:

$$G_v(x) \; = \; \sum_{k=0}^{\infty} P(v = k)\, x^k \; = \; \sum_{k=0}^{\infty} p_k\, x^k$$

where, in the second power series, we are letting $P(v = k) = p_k$. We call $G_v(x)$ a probability generating function (pgf). So $G_v(x)$ represents the distribution of the random variable v as a pgf.

The x in $G_v(x)$ has no special meaning, we could have also used the letter y or z, etc, instead of x, it would mean the same thing.

Example: We flip a coin. Let v count heads. So, if we get tails (T), $v = 0$; if we get heads (H), $v = 1$. If the coin is fair, meaning if $P(H) = P(T) = \frac{1}{2}$, we would represent the distribution of v as the pgf $G_v(x) = \frac{1}{2} + \frac{1}{2}\,x$

Example: We flip two coins. Let v count how many heads we get. If the coins are fair (and independent of each other) then $G_v(x) = \frac{1}{4} + \frac{2}{4}\,x + \frac{1}{4}\,x^2$. Why? Because:

there is a $\frac{1}{4}$ chance that we'll get 0 heads, i.e. $\{TT\}$, so the coefficient of x^0 is $\frac{1}{4}$.
there is a $\frac{2}{4}$ chance that we'll get 1 head, i.e. $\{TH, HT\}$, so the coefficient of x^1 is $\frac{1}{2}$.
there is a $\frac{1}{4}$ chance that we'll get 2 heads, i.e. $\{HH\}$, so the coefficient of x^2 is $\frac{1}{4}$.

Multi-step Processes

Suppose we have a process that we can break down into a number of steps. The number of steps we consider a process to have is largely a matter of convenience.

Example: We toss three coins. We can think of this as being three steps, where in each step we flip one of the coins. We could also think of this as being a one step process, where in that step we flip all three coins.

Example: We toss three coins and roll a pair of dice. We can think of this as being a two step process, where in one step we toss the three coins and in the other step we roll the pair of dice. We can also think of this as being a 5 step process. For three steps we toss a single coin and for two steps we roll a single die. We could also think of this as being a one step process, where in that one step we flip all three coins and roll both dice.

Multi-step Processes and Probability Generating Functions

Let's suppose that we have a two step process: with v being the count random variable for the first step and w being the count random variable for the second step. What can we say about the pgf for the distribution of $v + w$, which we we'll denote as $G_{v+w}(x)$?

Since,

$$P(v + w = j) = \sum_{i=0}^{j} P(v = i, w = j - i)$$

it follows that

$$G_{v+w} \;=\; \sum_{j=0}^{\infty} P(v + w = j)\, x^j \;=\; \sum_{j=0}^{\infty} \left(\sum_{i=0}^{j} P(v = i, w = j - i) \right) x^j \tag{7}$$

If v and w are independent then

$$P(v=i, w=j-i) = P(v=i)\, P(w=j-i) \quad \text{for all } i, j \in 0, 1, 2, \ldots$$

and Equation (7) becomes:

$$G_{v+w} \;=\; \sum_{j=0}^{\infty} \left(\sum_{i=0}^{j} P(v=i)\, P(w=j-i) \right) x^j \tag{8}$$

and since $x^j = x^i\, x^{j-i}$, we can rewrite Equation (8) as

$$\begin{aligned} G_{v+w} \;&=\; \sum_{j=0}^{\infty} \left(\sum_{i=0}^{j} P(v=i)\, x^i \;\, P(w=j-i)\, x^{j-i} \right) \\ &=\; \left(\sum_{k=0}^{\infty} P(v=k)\, x^k \right) \left(\sum_{k=0}^{\infty} P(w=k)\, x^k \right) \\ &=\; G_v(x)\, G_w(x) \end{aligned}$$

Example: We toss three fair coins. So the $P(H) = P(T) = \frac{1}{2}$. We can think of this as a two step process. In the first step we toss one of the coins. In the second step we toss the other two coins. Let v count the number of heads in the first step, i.e, for the first coin, so $v = 0$ or 1; and let w count the number of heads in the second step, i.e, for the second and third coins, so $w = 0, 1,$ or 2. The random variable $v + w$ will count how many heads we get all together, from the three coins, so $v + w = 0, 1, 2,$ or 3. The pgf for v is

$$G_v(x) \;=\; \frac{1}{2} + \frac{1}{2}\, x \;=\; p_0 + p_1 x$$

where $p_i = P(v = i)$ for $i = 0, 1$; and the pgf for w is

$$G_w(x) \;=\; \frac{1}{4} + \frac{2}{4}\, x + \frac{1}{4}\, x^2 \;=\; q_0 + q_1 x + q_2 x^2$$

where $q_i = P(w = i)$ for $i = 0, 1, 2$. So

$$\begin{aligned} G_{v+w} &= G_v(x)\, G_w(x) \\ &= (p_0 + p_1 x)\,(q_0 + q_1 x + q x^2) \\ &= p_0 q_0 + (p_0 q_1 + p_1 q_0) x + (p_0 q_2 + p_1 q_1) x^2 + p_1 q_2 x^3 \\ &= \frac{1}{2}\,\frac{1}{4} + \left(\frac{1}{2}\,\frac{2}{4} + \frac{1}{2}\,\frac{1}{4} \right) x + \left(\frac{1}{2}\,\frac{1}{4} + \frac{1}{2}\,\frac{2}{4} \right) x^2 + \frac{1}{2}\,\frac{1}{4}\, x^3 \\ &= \frac{1}{8} + \frac{3}{8}\, x + \frac{3}{8}\, x^2 + \frac{1}{8}\, x^3 \end{aligned}$$

Example: We toss three fair coins. Just like in the previous example. But now we'll think of this as a three step process. In each step we toss one of the coins. Let v_i count the number of heads in the i^{th} step, for $i = 1, 2, 3$, so $v_i = 0$ or 1; The random variable $v_1 + v_2 + v_3$ will count how many heads we get all together, from the three coins. The pgf's are

$$G_{v_1}(x) = G_{v_2}(x) = G_{v_3}(x) = G(x) \;=\; \frac{1}{2} + \frac{1}{2}\, x \;=\; p_0 + p_1 x$$

where $p_0 = p_1 = \frac{1}{2}$. So

$$\begin{aligned} G_{v_1+v_2+v_3} &= (G(x))^3 \\ &= (p_0 + p_1 x)^3 \\ &= p_0p_0p_0 + (p_0p_0p_1 + p_0p_1p_0 + p_1p_0p_0)x + (p_0p_1p_1 + p_1p_0p_1 + p_1p_1p_0)x^2 + p_1p_1p_1x^3 \\ &= \frac{1}{8} + \frac{3}{8}\,x + \frac{3}{8}\,x^2 + \frac{1}{8}\,x^3 \end{aligned}$$

Alternately,

$$\begin{aligned} G_{v_1+v_2+v_3} &= (G(x))^3 = \left(\frac{1}{2} + \frac{1}{2}\,x\right)^3 = \left(\frac{1}{2}\;(1+x)\right)^3 = \left(\frac{1}{2}\right)^3\;(1+x)^3 \\ &= \frac{1}{8}\;\left(1 + 3x + 3x^2 + 1\right) \\ &= \frac{1}{8} + \frac{3}{8}\,x + \frac{3}{8}\,x^2 + \frac{1}{8}\,x^3 \end{aligned}$$

Same as before.

Note: $(1+x)^n = \sum_{k=0}^{n} C(n,k)x^k$, where $C(n,k) = \dfrac{n!}{k!(n-k)!} = n$ choose k, which is just the n^{th} row in Pascal's triangle.

Composition of Probability Generating Functions Applied to Birth/Death Processes

Let us suppose that we have self replicating copies or clones of something [3] that live, die, and replicate independently of each other, but with identical probabilities. We'll represent each copy's "per step reproductive distribution" as the pgf

$$G(x) = \sum_{k=0}^{\infty} q_k\, x^k$$

where q_k is the probability that in that time step the copy will split into k copies of itself, with $k = 0$ representing the copy dying, and $k = 1$ representing the copy not dying, but not splitting either.

If we have j copies alive at the start of a step, then the distribution of the number of copies alive at the end of the step will be given by $(G(x))^j$.

We can write

$$(G(x))^j = \sum_{k=0}^{\infty} q_{j,\,k}\, x^k$$

[3]The word "copies" could refer to viruses, bacteria, any living organism or non-living entity that can "replicate" itself, "immigrate", and/or "die". In these notes, we will assume that every copy is a perfect copy of the original (so no mutation, no aging, etc). We will also assume that, for all intents and purposes, that replication/death/immigration happens in perfectly synchronized, discrete steps (which we'll sometimes call "time steps"). These assumptions simplify the math greatly. Most organisms do not have perfectly synchronized life-cycles, however sometimes they almost do. For example, perennial flowers have a one year life cycle: seeds sprout in the Spring, the flower is dead by winter.

where $q_{j,\,k}$ is the probability that if we start a time step with j copies alive, we will finish that time step with k copies alive. So $q_{j,\,k}$ is the conditional probability

$$q_{j,\,k} = P(k \text{ copies will be alive at the end of the time step} \mid j \text{ copies are alive at the start of the time step})$$

Note that $q_{1,k} = q_k$.

Let b_n be the random variable that counts the number copies that are "alive" at the end of the n^{th} step. We will represent the distribution of b_n as the probability generating function $B_n(x)$:

$$B_n(x) = \sum_{j=0}^{\infty} P(b_n = j)\, x^j = \sum_{j=0}^{\infty} p_{n,j}\, x^j$$

where $P(b_n = j) = p_{n,j}$.

It turns out that if all copies are identical and independent, and there is no immigration in or out, then $B_{n+1}(x) = B_n(G(x))$.

The following calculation and note shows this.

$$\begin{aligned}
B_n(G(x)) &= \sum_{j=0}^{\infty} p_{n,j}\ (G(x))^j \\
&= \sum_{j=0}^{\infty} p_{n,j}\ \left(\sum_{k=0}^{\infty} q_{j,\,k}\, x^k\right) \\
&= \sum_{j=0}^{\infty} \left(\sum_{k=0}^{\infty} p_{n,j}\, q_{j,\,k}\, x^k\right) \\
&= \sum_{k=0}^{\infty} \left(\sum_{j=0}^{\infty} p_{n,j}\, q_{j,\,k}\, x^k\right) \\
&= \sum_{k=0}^{\infty} \left(\sum_{j=0}^{\infty} p_{n,j}\, q_{j,\,k}\right)\ x^k \\
&= \sum_{k=0}^{\infty} (p_{n+1,k})\ x^k = B_{n+1}(x)
\end{aligned}$$

Note, the last three equalities follow from:

$p_{n,j}\, q_{j,\,k} =$ (the probability that at the end of the n^{th} step we'll have j copies alive) $\times$
(the probability that if we start a step with j copies alive, we'll end with k copies alive).

Since the end of the n^{th} step is the start of the $n+1$ step, if we sum $p_{n,j}\, q_{j,\,k}$ over j, we get:

$$\sum_{j=0}^{\infty} p_{n,j}\, q_{j,\,k} = p_{n+1,k} = \ \text{probability that the } n+1 \text{ step ends with } k \text{ copies alive}$$

The scenario of the replicating copies, discussed above, is an examples of a birth/death processes. When modeled by pgf's, such processes are often called Galton-Watson.

Expectation and Probability Generating Functions

The expectation of a random variable v, which takes on countably many distinct values $v_i,\ i = 0, 1, 2, \ldots$, is defined as

$$E(v) = \sum_{i=0}^{\infty} v_i\, P(v = v_i)$$

The expectation is a fancy way to say average. Suppose v is a count random variable so that

$$G_v(x) = \sum_{k=0}^{\infty} P(v = k)\, x^k$$

is its pgf. Then

$$\frac{d}{dx}\, G_v(x) = \sum_{k=1}^{\infty} kP(v = k)\, x^{k-1}$$

is its derivative. Plugging in $x = 1$ yields the expectation:

$$\frac{d}{dx}\, G_v(x)\bigg|_{x=1} = G'_v(1) = \sum_{k=1}^{\infty} kP(v = k) = E(v)$$

Example: We toss n fair coins. Let v count how many heads. Let $G(x) = \frac{1}{2} + \frac{1}{2}\,x$ Then:

$$\begin{aligned}
G_v &= (G(x))^n \\
&= \left(\frac{1}{2}\right)^n (1 + x)^n \\
G'_v(x) &= \left(\frac{1}{2}\right)^n n\,(1 + x)^{n-1} \\
E(v) = G'_v(1) &= \left(\frac{1}{2}\right)^n n\,(1 + 1)^{n-1} = \left(\frac{1}{2}\right)^n n\ (2)^{n-1} = \frac{n}{2}
\end{aligned}$$

So, if we toss n fair coins, we should expect to get $\dfrac{n}{2}$ heads, i.e., we should expect to get heads 1/2 the time.

Expectation of the Product of Probability Generating Functions

Suppose we have two step process, with v the count random variable for the first step and w the count random variable for the second. We showed above that if v and w are independent and if the $G_v(x)$ is the pgf for v and $G_w(x)$ is the pgf for w then $G_{v+w}(x) = G_v(x)G_w(x)$ will be the pgf for $v + w$. Then the expectation of $v + w$ is given by:

$$E(v{+}w) = E(G_{v+w}) = E(G_v\, G_w) = G'_{v+w}(1) = G'_v(1)G_w(1){+}G_v(1)G'_w(1) = G'_v(1){+}G'_w(1) = E(G_v){+}E(G_w)$$

where we have use the product rule from differential calculus: $(fg)' = f'g + fg'$ and the fact that if $G(x)$ is any pgf, then $G(1) = 1$.

If $G_v = G_w = G$ we have have the nice formula $E(G^2) = 2E(G)$ and, more generally, $E(G^j) = jE(G)$

Example: We toss n fair coins. Let v count how many heads. Find $E(v)$.

Solution: $G_v(x) = (G(x))^n$ where $G(x) = \frac{1}{2} + \frac{1}{2}x$. Since $E(G) = G'(1) = \frac{1}{2}$ we have:

$$E(v) = E(G_v) = E(G^n) = n\,E(G) = \frac{n}{2}$$

as before.

Expectation of the Composition of Probability Generating Functions

If $B_{n+1}(x) = B_n(G(x))$ then $B'_{n+1}(1) = B'_n(G(1))G'(1)$. But if G is any pgf then $G(1) = 1$ and so

$$E(B_{n+1}) = B'_{n+1}(1) = B'_n(1)G'(1) = E(B_n)E(G),$$

where $E(B_n)$ and $E(G)$ are the expectations of the distributions B_n and G.

Notation for composition: It is convenient to denote G composed with itself n times as $G^{(n)}(x)$. So $G(G(x)) = G^{(2)}(x)$. We also let $G^{(0)}(x) = x$.

The following calculation shows that

$$E((G^{(n)})) = (E(G))^n$$

First:

$$\frac{d}{dx}\,G^{(n)}(x) = G'(G^{(n-1)}(x))\ \ G'(G^{(n-2)}(x))\ \ \cdots\ \ G'(G(x))\ \ G'(x)$$

Then, since $G(1) = 1$, we have $G^{(k)}(1) = 1$ for $k = 0, 1, 2, \ldots$, and so:

$$\begin{aligned}
E((G^{(n)})) &= \frac{d}{dx}\,G^{(n)}(1) \\
&= G'(G^{(n-1)}(1))\ \ G'(G^{(n-2)}(1))\ \ \cdots\ \ G'(G(1))\ \ G'(1) \\
&= G'(1)\ \ G'(1)\ \ \cdots\ \ G'(1)\ \ G'(1) \\
&= (G'(1))^n \\
&= (E(G))^n
\end{aligned}$$

Example (a death process): We start each step by tossing or flipping our "live" coins. Those that come up tails we consider to be "dead" and are removed. Those that come up heads we consider to be "alive" and are counted. We start the process with 9 "living" coins. Find the distribution $B_n(x)$ of the number of coins that will be alive at the end of the n^{th} step. Also find the expectation of $B_n(x)$. Assume the coins are fair.

Solution. $B_0(x) = x^9$. Let $G(x) = .5(1+x)$. then:

$$\begin{aligned} B_1(x) &= (G(x))^9 \\ B_2(x) &= (G(G(x)))^9 \\ \vdots \qquad & \qquad \vdots \\ \text{So, } \ B_n(x) &= (G^{(n)}(x))^9 \end{aligned}$$

Since $G(x) = .5(1+x)$ it follows that $E(G) = G'(1) = .5$, and

$$\begin{aligned} G(G(x)) &= .5(1+.5(1+x)) = (.5) + (.5)^2 + (.5)^2 x \\ G(G(G(x))) &= .5(1+.5(1+.5(1+x))) = (.5) + (.5)^2 + (.5)^3 + (.5)^3 x \\ \vdots \qquad & \qquad \vdots \\ \text{So, } \ G^{(n)}(x) &= \left(\sum_{k=1}^{n} (.5)^k \right) + (.5)^n x = (1-(.5)^n) + (.5)^n x \end{aligned}$$

We can can now find the expectation:

$$\begin{aligned} E(B_n) &= \frac{d}{dx} B_n(x) \bigg|_{x=1} \\ &= \frac{d}{dx} (G^{(n)}(x))^9 \bigg|_{x=1} \\ &= 9(G^{(n)}(x))^8 \ \frac{d}{dx} G^{(n)}(x) \bigg|_{x=1} \\ &= 9(G^{(n)}(1))^8 (G'(1))^n \\ &= 9 \ (1)^8 \ (.5)^n \\ &= 9 \ (.5)^n \end{aligned}$$

For a more in depth treatment of probability generating functions see [1]. It is free to download for educational purposes from `https://www.math.upenn.edu/~wilf/gfologyLinked2.pdf`.

REFERENCES

[1] Wilf, H. S. 1994. *generatingfunctionology*/ San Diego CA: Academic Press, Inc.

[2] Winkel, B. J. 2014. *m&m Death and Immigration*. `https:/www.simiode.org/resources/132`. SIMIODE. `www.simiode.org`. Accessed 6 February 2016.