

## ¿QUÉ ES UN PROCESO DE RAMIFICACIÓN?

LILIANA BLANCO(\*)

---

**Resumen.** Éste es un artículo de carácter divulgativo en el que se pretende presentar los conceptos básicos de los procesos de ramificación así como dar una visión de su amplísimo campo de aplicación.

**Abstract.** In an expository article we present some basic concepts of branching processes and its applications.

**Keywords.** Branching processes, population models.

Los procesos de ramificación ofrecen un modelo matemático para el estudio de poblaciones en las que los individuos viven, se reproducen y mueren independientemente unos de otros. Con esta herramienta numerosas preguntas pueden ser propuestas y resueltas, como por ejemplo: ¿cuáles poblaciones se extinguen?, ¿cómo se comportan aquellas que no lo hacen?, si ellas crecen más allá de todo límite ¿qué tasa de crecimiento poseen?, etc.

Aún cuando el campo de aplicación de los procesos de ramificación parece muy limitado, se ve que el modelo en sí mismo es fundamental en el desarrollo de la investigación matemática y, para poblaciones tales como las células que se reproducen asexualmente, muchas de las respuestas obtenidas son de suma importancia. Los primeros indicios de problemas propuestos relacionados con los procesos de ramificación se encuentran en el siglo XVIII en el famoso libro de tres volúmenes del investigador Thomas Malthus, titulado *Essay on the Principles of Populations*, en el cual se plantea que una población no controlada

---

(\*)Texto recibido 1/3/96, revisado 29/1/97. Liliana Blanco, Departamento de Matemáticas y Estadística, Universidad Nacional de Colombia.

debe crecer exponencialmente. Malthus relata también que, en un pueblo de Berna, de las 487 familias burguesas existentes, 379 se extinguieron en el lapso de dos siglos (1583-1783); sin embargo, él no estudia el porqué de esta aparente paradoja. El primero en tratar de explicar el problema fué el matemático francés I.J. Bienayme (1796-1878). Parece ser (desafortunadamente no hay constancia escrita) que él fue capaz de relacionar correctamente la probabilidad de extinción con el promedio de hijos varones de cada individuo.

Independientemente de Bienayme, el matemático inglés Sir Francis Galton (1822-1911) aportó precisiones al problema, el cual puede formularse de la siguiente manera. En una población supondremos que el número de hijos de cada individuo  $x$  es una variable aleatoria  $\xi_x$  con distribución  $p := (p_n)_{n \in \mathbb{N}}$ , siendo  $p_n :=$  "probabilidad de  $n$  hijos" y con  $\sum_{n=0}^{\infty} p_n = 1$  (para evitar situaciones triviales, supondremos  $p_0 + p_1 < 1$ ). El número de hijos que cada individuo tiene es independiente de su historia familiar y del número de hijos de los demás individuos. Los hijos tienen sus propios hijos con la misma distribución  $p = (p_n)_{n \in \mathbb{N}}$ . Supongamos que se inicia el proceso con un individuo quien constituye la 0-ésima generación, sus hijos forman la primera generación, sus nietos la segunda y así sucesivamente. Si la variable aleatoria  $Z_n$  denota el número de individuos en la  $n$ -ésima generación, entonces tenemos que  $Z_0 = 1$  y que  $Z_1$  tiene distribución  $p$ . Si en la  $n$ -ésima generación hay  $i \geq 1$  individuos, entonces en la  $(n+1)$ -ésima generación habrá  $Z_n^{(1)} + Z_n^{(2)} + \dots + Z_n^{(i)}$  individuos, donde  $Z_n^{(k)}$  denota el número de hijos del  $k$ -ésimo individuo de la  $n$ -ésima generación,  $k = 1, 2, \dots, i$ .

Como los  $Z_n^{(k)}$  son independientes y tienen la misma distribución  $p$ , entonces la probabilidad de transición

$$\pi_{ij} := P(Z_{n+1} = j | Z_n = i) = p_j^{*i}$$

es la  $j$ -ésima componente de la  $i$ -ésima convolución de  $p$ . En otras palabras, la distribución de  $Z_{n+1}$  dado que  $Z_n = i$ , es igual a la distribución de una suma de variables aleatorias independientes e igualmente distribuidas con distribución  $p$ . Esta última es la propiedad que caracteriza a los procesos de ramificación.

Galton se pregunta por la probabilidad de extinción del proceso  $(Z_n)_{n \in \mathbb{N}}$ , esto es, calcula  $q := \lim_{n \rightarrow \infty} P(Z_n = 0) = P(Z_n \rightarrow 0)$ . Tenemos que si el individuo inicial tiene  $k$  hijos, entonces su descendencia se extingue con probabilidad  $q^k$ , esto es,  $q = \sum_{k=0}^{\infty} P(\text{"}k \text{ hijos"} ) P(\text{"extinción"} | \text{"}k \text{ hijos"} ) = \sum_{k=0}^{\infty} p_k q^k$ , es decir,  $q$  es solución de la ecuación  $s = f(s)$  donde  $f(s) = Es^{Z_1}$  con  $0 \leq s \leq 1$ . Así,  $f(s)$  denota la función generadora de la distribución de la variable aleatoria  $Z_1$ . Para  $0 \leq s \leq 1$  se tiene que la función  $f(s)$  es continua, estrictamente creciente y convexa; además  $f(0) = P(Z_1 = 0) = p_0$  y  $f(1) = 1$ .

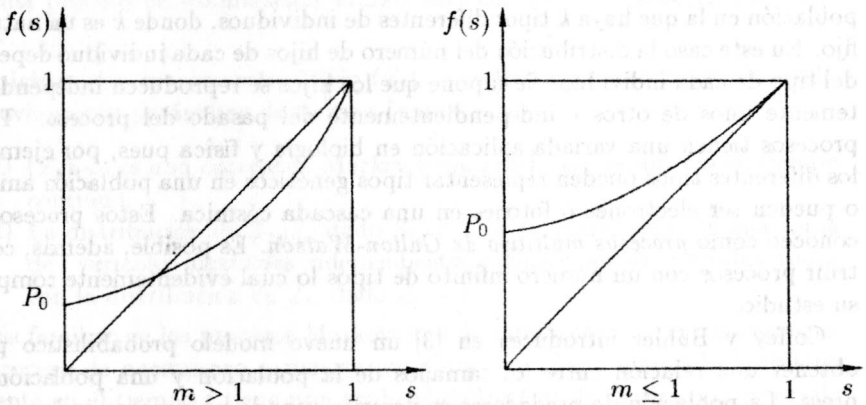
La única respuesta al problema planteado por Galton fué dada por el clérigo Watson quien concluyó que  $q = 1$ . Esta solución no es del todo correcta, como se deduce de un análisis detenido de  $f(s)$ .

Puesto que  $p_0 + p_1 < 1$  entonces debe existir por lo menos un  $k \geq 2$  con  $p_k > 0$ . Por lo tanto

$$f'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1}$$

es estrictamente creciente, de donde se deduce que la ecuación  $f(s) = s$  tiene a lo más dos soluciones. Tenemos que:

- i) Si  $m := f'(1) \leq 1$ , entonces  $f'(s) < 1$  para  $0 \leq s < 1$ . Por lo tanto, toda la gráfica de  $f(s)$  en  $(0,1)$  se encuentra encima de la diagonal. Esto implica que 1 es la única solución de la ecuación  $f(s) = s$ , esto es,  $q = 1$ .
- ii) Si  $m = f'(1) > 1$ , entonces  $f'(s_0) > 1$  para algún  $s_0 < 1$  suficientemente cercano a 1. Se tiene que en  $(s_0, 1)$  la gráfica de  $f(s)$  debe estar por debajo de la diagonal. Si  $p_0 > 0$ , entonces debe existir  $t$ ,  $0 < t < 1$ , tal que  $f(t) = t$ . Puesto que  $q$  es la menor solución no negativa de la ecuación  $f(s) = s$  se deduce que  $0 < q < 1$ . En el caso en que  $p_0 = 0$  se tiene que  $f(0) = 0$  y por tanto  $q = 0$ .



$m :=$  "número esperado de hijos de cada individuo"

(Fig.1)

El danés J.F. Steffensen publicó por primera vez en el año 1929 un análisis completo de la probabilidad de extinción. Steffensen demostró que  $q = 1$  si y sólo si  $m \leq 1$ . El proceso  $(Z_n)_{n \in \mathbb{N}}$  anteriormente descrito se conoce en la literatura bajo el nombre de proceso Galton-Watson. La definición exacta es la siguiente.

**Definición.** Un proceso Galton-Watson  $(Z_n)_{n \in \mathbb{N}}$  es una cadena de Markov homogénea sobre los enteros no negativos. Sus probabilidades de transición se definen en términos de una distribución dada  $\{p_k, k = 1, 2, \dots\}$ ,  $p_k \geq 0$ ,  $\sum p_k = 1$ :

$$\pi_{ij} := P(Z_{n+1} = j | Z_n = i) = \begin{cases} p_j^{*i}, & i \geq 1, \quad j \geq 0 \\ \delta_{0j}, & i = 0, \quad j \geq 0, \end{cases}$$

donde  $\delta_{ij}$  es el símbolo de Kronecker y  $\{p_k^{*i}, k = 0, 1, 2, \dots\}$  es la  $k$ -ésima convolución de  $\{p_k, k = 0, 1, 2, \dots\}$ .

Sin pérdida de generalidad, se puede suponer que  $Z_0 = 1$ , ya que el proceso  $\{Z_n^{(i)}, n = 0, 1, 2, \dots\}$  en el que  $Z_0 = i$ , es la suma de  $i$  copias independientes del proceso  $\{Z_n, n = 0, 1, 2, \dots\}$ . El proceso original de Galton-Watson y sus generalizaciones están relacionados con los trabajos de Niels Abel acerca de las ecuaciones funcionales y la iteración de funciones. También está relacionado con varias líneas de desarrollo de la teoría de procesos estocásticos.

Hasta ahora hemos considerado una población en la que los individuos son todos del mismo tipo. Para muchas aplicaciones es conveniente considerar una población en la que haya  $k$  tipos diferentes de individuos, donde  $k$  es un natural fijo. En este caso la distribución del número de hijos de cada individuo depende del tipo de cada individuo. Se supone que los hijos se reproducen independientemente unos de otros e independientemente del pasado del proceso. Tales procesos tienen una variada aplicación en biología y física pues, por ejemplo, los diferentes tipos pueden representar tipos genéticos en una población animal o pueden ser electrones o fotones en una cascada cósmica. Estos procesos se conocen como *procesos multitypo de Galton-Watson*. Es posible, además, construir procesos con un número infinito de tipos lo cual evidentemente complica su estudio.

Coffey y Bühler introducen en [3] un nuevo modelo probabilístico para obtener una relación entre los tamaños de la población y una población de presa. La población de predadores es descrita usando un proceso ordinario de Galton-Watson. En la población de presa se supone que cada individuo sobreviviente de la  $n$ -ésima generación produce un número aleatorio de individuos, independientemente de todos los otros individuos. Sin embargo, no todos los hijos de esos individuos se reproducen. En su lugar se tiene que cada miembro de la  $(n+1)$ -ésima generación de predadores consume un número aleatorio de hijos de la población de presa. Coffey y Bühler dan condiciones necesarias y suficientes para tener una probabilidad positiva de supervivencia indefinida para ambas poblaciones. Se demuestra que si  $X_n :=$  "número de predadores en la  $n$ -ésima generación",  $Y_n :=$  "tamaño de la  $n$ -ésima generación de la población de presa",  $\mu :=$  "número promedio de hijos de un predador" y  $m :=$  "número promedio de hijos de un presa", entonces se tienen los resultados:

(a) Si  $m \leq \mu$ , entonces para todo tamaño inicial  $x_0$  de la población de predadores y para todo tamaño inicial  $y_0$  de la población de presa se satisface:

$$P(Y_n \rightarrow 0 | X_n \rightarrow 0) = 1.$$

(b) Si  $m > \mu$ , entonces para todo tamaño inicial  $x_0$  de la población de predadores existe un tamaño inicial  $y_0$  de la población de presa tal que

$$P(Y_n \rightarrow 0, X_n \rightarrow 0) > 0.$$

El tamaño de la población de presa  $Y_n$  no constituye un proceso de Galton-Watson pues no se puede asumir sin pérdida de generalidad que  $y_0 = 1$ .

En los procesos de Galton-Watson la longitud de vida  $\lambda_x$  de cada individuo  $x$  es una unidad de tiempo. Una generalización natural es permitir que dichas longitudes de vida sean variables aleatorias. En este caso se estudia un proceso estocástico  $(Z_t)_{t \geq 0}$ , donde  $Z_t$  denota el número de individuos en el tiempo  $t$ . Estos procesos no son en general Markovianos, a menos que las longitudes de vida de los individuos sean variables aleatorias independientes e igualmente distribuidas con distribución exponencial. Cuando esto ocurre, el proceso  $(Z_t)_{t \geq 0}$  se llama proceso de ramificación Markoviano con parámetro de tiempo continuo.

**Definición.** Un proceso estocástico  $(Z_t)_{t \geq 0}$  se llama proceso de ramificación Markoviano con parámetro de tiempo continuo si:

- (1)  $(Z_t)_{t \geq 0}$  es una cadena de Markov homogénea con parámetro de tiempo continuo.
- (2) La distribución de  $Z_{t+t_0}$ , dado  $Z_{t_0} = i$ , es la distribución de una suma de  $i$  variables aleatorias independientes e igualmente distribuidas, todas con la distribución de  $Z_t$ , dado  $Z_0 = 1$ .

El más familiar de los procesos Markovianos de ramificación es el caso especial del proceso de nacimiento y muerte en el que se supone que cada individuo existente en el tiempo  $t$  tiene una probabilidad  $\mu(t)dt$  de morir en el intervalo de tiempo  $(t, t+dt)$  y una probabilidad  $\lambda(t)dt$  de desaparecer y ser reemplazado por dos nuevos individuos en el mismo intervalo de tiempo. Se obtiene en este caso que  $p_0 = \frac{\mu(t)}{(\lambda+\mu)(t)}$ ,  $p_2 = \frac{\lambda(t)}{(\lambda+\mu)(t)}$ .

Cuando las longitudes de vida son variables aleatorias independientes con una distribución general  $G$  (no necesariamente exponencial), entonces el proceso  $(Z_t)_{t \geq 0}$  se llama proceso de Bellman-Harris, en honor a R. Bellman y T. Harris quienes lo introdujeron por primera vez en el año 1952. En el modelo presentado por Harris y Bellman, los individuos son independientes unos de otros, viven un tiempo aleatorio y luego tienen un número aleatorio de hijos, independientemente de la longitud de vida del padre. Una generalización posible es permitir que el individuo se reproduzca en varios instantes aleatorios

durante su vida y no necesariamente al morir. En este caso  $\xi_x$  representa un número aleatorio de puntos sobre  $\mathbb{R}_+$  con localizaciones aleatorias no necesariamente distintas. De nuevo se supone independencia entre los individuos. Dicho modelo se conoce como modelo Sevast'yanov.

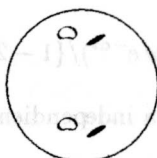
Una característica básica de los procesos de ramificación discutidos hasta ahora es la homogeneidad, en el tiempo, de la distribución de descendencia. Sin embargo puede resultar más realista suponer que la distribución de descendencia depende de un medio de desarrollo, el cual cambia con el tiempo. Los primeros en estudiar tal situación fueron Smith y Wilkinson en el año 1969, quienes presentaron un modelo para tal situación y estudiaron la probabilidad de extinción. Karlin y Athreya (1971) reformularon y generalizaron el modelo Smith-Wilkinson tratando el problema de extinción y obteniendo extensiones de los teoremas límites básicos del proceso de Galton-Watson. Este proceso de ramificación con medio aleatorio es utilizado por Altenburg, en el año de 1985, para desarrollar nuevas estrategias en la administración de recursos. En [1] Altenburg plantea el problema de que el uso de pesticidas en el control de plagas provoca numerosas consecuencias importantes y efectos desventajosos en el medio ambiente como, por ejemplo, aniquilación de enemigos naturales, incremento del nivel de resistencia de la plaga, contaminación del agua, etc. Por lo tanto se tendría que el objetivo principal del control de la plaga sería el de maximizar la probabilidad de extinción de la plaga, minimizando simultáneamente el tiempo requerido hasta la extinción de la misma. En [1] Altenburg utiliza la teoría de los procesos de ramificación para dar condiciones que permitan maximizar la probabilidad de extinción, minimizando simultáneamente el tiempo esperado hasta la extinción de la plaga.

Como anotábamos al comienzo, los procesos de ramificación ofrecen un modelo matemático adecuado para la descripción de poblaciones que se reproducen asexualmente, tales como las células. Es bien sabido que los problemas de crecimiento y división de las células son fundamentales en la biología. Desde hace ya unas décadas se sabe que hay en la vida de una célula reproductora (no todas lo son, por ejemplo, las neuronas del sistema nervioso central son un tipo especial de células que nunca se dividen) un periodo de tiempo en el cual el DNA es sintetizado. Una célula recién nacida entra inicialmente en un periodo presintético llamado fase  $G_1$ . Durante este periodo la célula se prepara para la síntesis del DNA. La fase  $G_1$  es seguida de una fase de síntesis llamada fase  $S$ . En la siguiente fase llamada fase  $G_2$ , la célula se prepara para la última parte de su vida conocida como mitosis o fase  $M$ , durante la cual la célula se divide en dos células hijas, cada una de las cuales entra en la fase  $G_1$ . Gráficamente

la división celular se lleva a cabo de la siguiente manera:



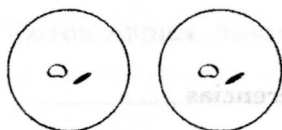
(Fig.2)



(Fig.3)



(Fig.4)



(Fig.5)

En la figura 2 los cromosomas se contraen y toman lugar a lo largo del plano ecuatorial del núcleo. En la figura 3 los cromosomas se dividen y se dirigen a los polos (figura 4). En la figura 5 la célula se divide. El ciclo celular (o tiempo de generación o ciclo mitótico) es la suma de las cuatro fases. Es necesario aclarar que no necesariamente el ciclo celular coincide con el tiempo de vida de la célula pues es posible que ésta muera antes de completar el ciclo.

Consideremos un proceso de ramificación en el cual la longitud de vida  $\lambda$  de una célula es, con probabilidad  $p$ , una variable aleatoria  $T$  (el ciclo de la célula) y con probabilidad  $(1-p)$  una variable aleatoria  $D$  (el tiempo transcurrido hasta la muerte de la célula). Si la célula muere antes de completar el ciclo, entonces  $\lambda = D$  y la célula no tiene hijas; mientras que, si la célula no muere antes de completar el ciclo, entonces  $\lambda = T$  y en tal caso la célula tiene exactamente dos hijas, quienes nacen al momento de morir la madre.

Supongamos que  $T$  tiene distribución  $G$  y que  $D$  tiene distribución  $F$ ; si  $L$  denota la distribución de  $\lambda$  se tiene que  $L = pG + (1-p)F$ . Además, si  $\mu(t) := E(\xi(t))$  donde  $\xi(t) :=$  "número de hijas del individuo inicial nacidas hasta el tiempo  $t$ ", entonces se tiene que  $\mu(t) = 2pG$ .

Si se supone que  $p > 1/2$  y que la distribución  $G$  es no reticulada, entonces se puede demostrar que existe  $\alpha > 0$  tal que

$$\int_0^{\infty} e^{-\alpha t} d\mu(t) = 1;$$

$\alpha$  recibe el nombre de parámetro Malthusiano.

El índice mitótico  $mi$  se define como la razón entre el número de células en mitosis y el número total de células. Usando la teoría de los procesos de ramificación se puede ver que

$$(*) \quad mi \approx 2p\alpha E(Me^{-\alpha t}) / \{1 - 2(1-p)E(e^{-\alpha D})\}.$$

En caso de que  $p \approx 1$  y  $M$  sea independiente del resto del ciclo, se puede ver que (\*) se reduce a

$$mi \approx \log 2 \frac{EM}{ET}.$$

### Referencias

1. H. Altenburg, "On a Branching Model for Management Resource Systems", *Statistics & Decisions* (1985), Supplement Issue 2, 111-116.
2. K. Athreya, P. Ney, *Branching Processes*, New York: Springer Verlag, 1972.
3. J. Coffey, W. Bühler, "The Galton-Watson Predator-Prey Process", *Appl. Prob.* 28 (1991), 9-16.
4. T. Harris, *The Theory of Branching Processes*, New York: Dover, 1989.
5. P. Jagers, *Branching Processes with Biological Applications*, New York: Wiley, 1975.