El proceso de Galton-Watson aplicado a la modelación del crecimiento de células

Antonio Murillo Salas
anmurillos@ugto.mx

Departamento de Matemáticas
Universidad de Guanajuato

Marzo 12, 2014

- Sea $\left(\xi_i^n\right)_{i,n=0}^{\infty}$ una colección de v.a. i.i.d tales que

$$p_k := \mathbb{P}(\xi_i^n = k) \ \text{ y } \ \sum_{k=0}^{\infty} p_k = 1.$$

- Sea $\left(\xi_i^n\right)_{i,n=0}^{\infty}$ una colección de v.a. i.i.d tales que

$$p_k := \mathbb{P}(\xi_i^n = k) \;\text{ y }\; \sum_{k=0}^{\infty} p_k = 1.$$

Función generadora de probabilidades: para $|s| \leq 1$,

$$G(s) := \mathbb{E}[s^{\xi_1^1}] = \sum_{k=0}^{\infty} p_k s^k.$$

- Sea $\left(\xi_i^n\right)_{i,n=0}^{\infty}$ una colección de v.a. i.i.d tales que

$$p_k := \mathbb{P}(\xi_i^n = k) \text{ y } \sum_{k=0}^{\infty} p_k = 1.$$

Función generadora de probabilidades: para $|s| \leq 1$,

$$G(s) := \mathbb{E}[s^{\xi_1^1}] = \sum_{k=0}^{\infty} p_k s^k.$$

- Sea $Z_0 = 1$ y defina

$$Z_{n+1} := \sum_{k=1}^{Z_n} \xi_k^n,$$

para $n = 0, 1, \cdots$.

- $Z = (Z_n)_{n=0}^{\infty}$ es una cadena de Markov con espacio de estados $\{0, 1, 2, \cdots\}$ y es conocida como proceso de **Galton-Watson.**

- $Z = (Z_n)_{n=0}^{\infty}$ es una cadena de Markov con espacio de estados $\{0, 1, 2, \cdots\}$ y es conocida como proceso de **Galton-Watson.**

- El proceso Z se puede pensar como un modelo para el crecimiento de alguna población.

- $Z = (Z_n)_{n=0}^{\infty}$ es una cadena de Markov con espacio de estados $\{0, 1, 2, \cdots\}$ y es conocida como proceso de **Galton-Watson.**

- El proceso Z se puede pensar como un modelo para el crecimiento de alguna población.

- La pregunta natural es: **¿qué le pasa a la población cuando $n \to \infty$?**

FGP de Z

- Sea
$$G_n(s) := \mathbb{E}(s^{Z_n}).$$
Notemos que $G_1(s) = G(s)$.

FGP de Z

- Sea
$$G_n(s) := \mathbb{E}(s^{Z_n}).$$
Notemos que $G_1(s) = G(s)$.

- Recuerde que
$$Z_{n+1} := \sum_{k=1}^{Z_n} \xi_k,$$

FGP de Z

- Sea
$$G_n(s) := \mathbb{E}(s^{Z_n}).$$
Notemos que $G_1(s) = G(s)$.

- Recuerde que
$$Z_{n+1} := \sum_{k=1}^{Z_n} \xi_k,$$
condicionando sobre Z_1 se tiene que para $|s| \leq 1$
$$G_{n+1}(s) = G_n(G(s)) = G(G_n(s)), \text{ para todo } n \geq 1.$$

Sea $m := \mathbb{E}[\xi_1]$. Puede probarse que

$$\mathbb{E}[Z_n] = m^n,$$

Sea $m := \mathbb{E}[\xi_1]$. Puede probarse que

$$\mathbb{E}[Z_n] = m^n,$$

y

$$\mathrm{Var}(X_n) = \begin{cases} \sigma^2 m^{n-1} \frac{m^n - 1}{m-1} & \text{si } m \neq 1, \\ n\sigma^2 & \text{si } m = 1. \end{cases}$$

Ejemplo: ramificación geométrica

Ejemplo: ramificación geométrica

- Supongamos que $G(s) = q(1 - ps)^{-1}$ $(p + q = 1)$, $|s| < \frac{1}{p}$, es decir, $\mathbb{P}(X = k) = qp^k, \; k \geq 0$.

Ejemplo: ramificación geométrica

- Supongamos que $G(s) = q(1 - ps)^{-1}$ $(p + q = 1)$, $|s| < \frac{1}{p}$, es decir, $\mathbb{P}(X = k) = qp^k, \; k \geq 0$.

- En tal caso se puede ver que

$$G_n(s) = \begin{cases} \frac{n-(n-1)s}{n+1-ns}, & p = \frac{1}{2}, \\ \frac{q[p^n - q^n - ps(p^{n-1} - q^{n-1})]}{p^{n+1} - q^{n+1} - ps(p^n - q^n)}, & p \neq \frac{1}{2}. \end{cases}$$

Probabilidad de extinción

- Probabilidad de extinción al tiempo n, es decir, $\mathbb{P}(Z_n = 0) = G_n(0)$;

Probabilidad de extinción

- Probabilidad de extinción al tiempo n, es decir, $\mathbb{P}(Z_n = 0) = G_n(0)$;
- o bién, $\lim_{n\to\infty} \mathbb{P}(Z_n = 0)$.

Probabilidad de extinción

- Probabilidad de extinción al tiempo n, es decir, $\mathbb{P}(Z_n = 0) = G_n(0)$;
- o bién, $\lim_{n\to\infty} \mathbb{P}(Z_n = 0)$.
- En el caso de ramificación geométrica:

$$\mathbb{P}(Z_n = 0) = \begin{cases} \frac{n}{n+1}, & p = q, \\ \frac{q(p^n - q^n)}{p^{n+1} - q^{n+1}}, & p \neq q. \end{cases}$$

Probabilidad de extinción

- Probabilidad de extinción al tiempo n, es decir, $\mathbb{P}(Z_n = 0) = G_n(0)$;
- o bién, $\lim_{n\to\infty} \mathbb{P}(Z_n = 0)$.
- En el caso de ramificación geométrica:

$$\mathbb{P}(Z_n = 0) = \begin{cases} \frac{n}{n+1}, & p = q, \\ \frac{q(p^n - q^n)}{p^{n+1} - q^{n+1}}, & p \neq q. \end{cases}$$

- Por lo tanto,

$$\lim_{n\to\infty} \mathbb{P}(Z_n = 0) = \begin{cases} 1, & p \leq q, \\ \frac{q}{p}, & p > q. \end{cases}$$

Probabilidad de extinción

- Probabilidad de extinción al tiempo n, es decir, $\mathbb{P}(Z_n = 0) = G_n(0)$;
- o bién, $\lim_{n\to\infty} \mathbb{P}(Z_n = 0)$.
- En el caso de ramificación geométrica:

$$\mathbb{P}(Z_n = 0) = \begin{cases} \frac{n}{n+1}, & p = q, \\ \frac{q(p^n - q^n)}{p^{n+1} - q^{n+1}}, & p \neq q. \end{cases}$$

- Por lo tanto,

$$\lim_{n\to\infty} \mathbb{P}(Z_n = 0) = \begin{cases} 1, & p \leq q, \\ \frac{q}{p}, & p > q. \end{cases}$$

- **Observación:**
  $\lim_{n\to\infty} \mathbb{P}(Z_n = 0) = 1$ cuando $\mathbb{E}[Z_1] = \frac{p}{q} \leq 1$;
  $\lim_{n\to\infty} \mathbb{P}(Z_n = 0) < 1$ cuando $\mathbb{E}[Z_1] = \frac{p}{q} > 1$.

Clasificación de procesos de Galton-Watson

Clasificación de procesos de Galton-Watson

Los procesos de Galton-Watson se clasifican de la siguiente manera:

- subcrítico si $m < 1$,
- crítico si $m = 1$,
- supercrítico si $m > 1$.

Comportamiento asintótico de Z

Comportamiento asintótico de Z

Teorema

El proceso de Galton-Watson se extingue si y sólo si la ramificación es crítica o subcrítica.

Comportamiento asintótico de Z

Teorema

El proceso de Galton-Watson se extingue si y sólo si la ramificación es crítica o subcrítica.

Lema

La probabilidad de extinción es la menor raíz positiva de la ecuación

$$G(s) = s, \; 0 \leq s \leq 1.$$

Aplicación: ciclo de la célula con muerte e inactividad

Kimmel, M. and Axelrod, D.E. 1991. *Unequal cell division, growth regulation and colony size of mammalian cells: A mathematical model and analysis of experimental data*. Journal of Theoretical Biology 153: 157-180.

Fenómeno biológico

Suponga que se tiene el siguiente modelo de reproducción celular: se inicia con una célula prolífica (es la que va a producir nueva células al momento de la división celular).

Fenómeno biológico

Suponga que se tiene el siguiente modelo de reproducción celular: se inicia con una célula prolífica (es la que va a producir nueva células al momento de la división celular).

El proceso de división celular consiste en la replicación de la célula en dos copias independientes, la cuales pueden ser de tres típos:

Fenómeno biológico

Suponga que se tiene el siguiente modelo de reproducción celular: se inicia con una célula prolífica (es la que va a producir nueva células al momento de la división celular).

El proceso de división celular consiste en la replicación de la célula en dos copias independientes, la cuales pueden ser de tres típos:

(i) **célula prolífica**, con probabilidad p_2;

Fenómeno biológico

Suponga que se tiene el siguiente modelo de reproducción celular: se inicia con una célula prolífica (es la que va a producir nueva células al momento de la división celular).

El proceso de división celular consiste en la replicación de la célula en dos copias independientes, la cuales pueden ser de tres típos:

(i) **célula prolífica**, con probabilidad p_2;

(ii) **célula inactiva**, con probabilidad p_1;

Fenómeno biológico

Suponga que se tiene el siguiente modelo de reproducción celular: se inicia con una célula prolífica (es la que va a producir nueva células al momento de la división celular).

El proceso de división celular consiste en la replicación de la célula en dos copias independientes, la cuales pueden ser de tres típos:

(i) **célula prolífica**, con probabilidad p_2;

(ii) **célula inactiva**, con probabilidad p_1;

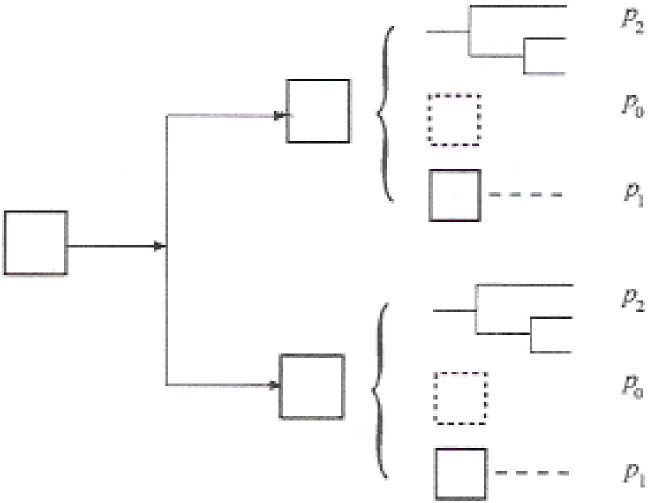
(iii) **una de la células hijas se muere**, con probabilidad p_0.

- Las células inactivas continuan en ese estado sin reproducirse o morirse.

- Las células inactivas continuan en ese estado sin reproducirse o morirse.
- En realidad, las células inactivas pueden convertirse en células prolíficas aún después de un periodo largo de tiempo.

- Las células inactivas continuan en ese estado sin reproducirse o morirse.
- En realidad, las células inactivas pueden convertirse en células prolíficas aún después de un periodo largo de tiempo.
- En el modelo no se considera esta posibilidad. Por lo tanto, se supone que

$$p_0 + p_1 + p_2 = 1.$$


p_2
p_0
p_1
p_2
p_0
p_1

El modelo matemático

El modelo matemático

El modelo matemático

- Sea Z_n el número de células prolíficas en la generación (división) n, y Q_n el número de células inactivas en la generación n.

El modelo matemático

- Sea Z_n el número de células prolíficas en la generación (división) n, y Q_n el número de células inactivas en la generación n.

- Denotemos por $Z_{n,k}^{(j)}$ ($k > n$) el número de células prolíficas en la k-ésima generación que son descendientes de la j-ésima célula prolífica en la generación n.

El modelo matemático

- Sea Z_n el número de células prolíficas en la generación (división) n, y Q_n el número de células inactivas en la generación n.

- Denotemos por $Z_{n,k}^{(j)}$ ($k > n$) el número de células prolíficas en la k-ésima generación que son descendientes de la j-ésima célula prolífica en la generación n. De manera análoga se define la correspondiente $Q_{n,k}^{(j)}$.

El modelo matemático

- Sea Z_n el número de células prolíficas en la generación (división) n, y Q_n el número de células inactivas en la generación n.

- Denotemos por $Z_{n,k}^{(j)}$ ($k > n$) el número de células prolíficas en la k-ésima generación que son descendientes de la j-ésima célula prolífica en la generación n. De manera análoga se define la correspondiente $Q_{n,k}^{(j)}$.

- Note que $Z_{1,n+1}^{(j)}$ y $Q_{1,n+1}^{(j)}$ tienen la misma distribución que Z_n y Q_n, respectivamente.

... si la población inicia con una célula prolífica,

... si la población inicia con una célula prolífica, entonces la generación $n+1$ está dada por

$$Z_{n+1} = \sum_{j=1}^{Z_n} Z_{1,n+1}^{(j)} \tag{1}$$

... si la población inicia con una célula prolífica, entonces la generación $n+1$ está dada por

$$Z_{n+1} = \sum_{j=1}^{Z_n} Z_{1,n+1}^{(j)} \tag{1}$$

y

$$Q_{n+1} = \sum_{j=1}^{Z_n} Q_{1,n+1}^{(j)} + Q_1. \tag{2}$$

Sea $G_n(s,t)$ la función generadora de probabilidades conjunta del vector (Z_n, Q_n):

Sea $G_n(s,t)$ la función generadora de probabilidades conjunta del vector (Z_n, Q_n):

$$G_n(s,t) = \mathbb{E}\left[s^{Z_n} t^{Q_n} | Z_0 = 1, Q_0 = 0\right], \;\; 0 \leq s, t \leq 1.$$

Sea $G_n(s,t)$ la función generadora de probabilidades conjunta del vector (Z_n, Q_n):

$$G_n(s,t) = \mathbb{E}\left[s^{Z_n} t^{Q_n} | Z_0 = 1, Q_0 = 0\right], \;\; 0 \leq s,t \leq 1.$$

Para encontrar $G_n(s,t)$ se condiciona sobre (Z_1, Q_1)...

Sea $G_n(s,t)$ la función generadora de probabilidades conjunta del vector (Z_n, Q_n):

$$G_n(s,t) = \mathbb{E}\left[s^{Z_n} t^{Q_n} | Z_0 = 1, Q_0 = 0\right], \;\; 0 \leq s, t \leq 1.$$

Para encontrar $G_n(s,t)$ se condiciona sobre (Z_1, Q_1)...

$$G_{n+1}(s,t) = \left(p_2 G_n(s,t) + p_1 t + p_0\right)^2 .$$

$G_n(s,t)$ contiene toda la información sobre el vector (Z_{n+1}, Q_{n+1}).

$G_n(s,t)$ contiene toda la información sobre el vector (Z_{n+1}, Q_{n+1}). Por ejemplo,

(a) si sólo estamos interesados en el comportamiento de la población de células prolíficas, es decir, en la evolución del proceso $\{Z_n, n \geq 0\}$.
Si
$$f_n(s) = \mathbb{E}\left[s^{Z_n} | Z_0 = 1\right],$$
entonces
$$f_{n+1}(s) = G_{n+1}(s,1) = (p_2 f_n(s) + p_1 + p_0)^2 .$$

$G_n(s,t)$ contiene toda la información sobre el vector (Z_{n+1}, Q_{n+1}). Por ejemplo,

(a) si sólo estamos interesados en el comportamiento de la población de células prolíficas, es decir, en la evolución del proceso $\{Z_n, n \geq 0\}$.
Si
$$f_n(s) = \mathbb{E}\left[s^{Z_n} | Z_0 = 1\right],$$
entonces
$$f_{n+1}(s) = G_{n+1}(s,1) = (p_2 f_n(s) + p_1 + p_0)^2 .$$

$\{Z_n, n \geq 0\}$ es un proceso de Galton-Watson con ley ramificación
$$f(s) = (p_2 s + p_0 + p_1)^2.$$

(b) Experimentalmente es imposible distinguir si una célula es activa o prolífica.
En el laboratorio se observa $Z_n + Q_n$.
Si $g_n(s) = \mathbb{E}\left[s^{Z_n+Q_n} | Z_0 = 1\right]$ entonces,

$$g_n(s) = G_n(s, s).$$

Aplicación a datos reales

Se consideran dos tipos de células:

(1) células de piel de ratón normales (NIH);

Se consideran dos tipos de células:

(1) células de piel de ratón normales (NIH);

(2) células NIH modificadas por la transferencia de algún cancer.

Se consideran dos tipos de células:

(1) células de piel de ratón normales (NIH);

(2) células NIH modificadas por la transferencia de algún cancer.

El objetivo del experimento fue: **establecer la diferencia en el proceso de crecimiento entre los dos tipos de células.**

Los datos

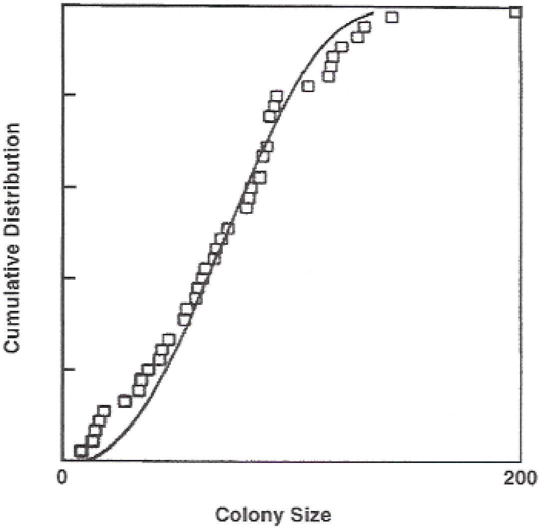
3. The Galton–Watson Process

TABLE 3.2. Colony Size Distribution: Data and Parameter Estimates

	Cell Type	
	NIH	NIH(*ras*)
Data		
Duration of the experiment (hours)	96	96
Number of colonies	52	45
Colony size (cells/colony)		
Minimum	10	8
Maximum	116	214
Median	33	70
Estimated parameters		
Number of divisions	8	8
Probability of death (p_0)	0.15	0.15
Probability of quiescence (p_1)	0.1	0

Células normales

Células modificadas



Blibliografía

- Karlin, S. and Taylor, H. An Introduction to Stochastic Modeling. Academic Press Inc.
- Kimmel, M. and Alxerod, D.E.*Branching processes in Biology*. Springer (2002)
- Kimmel, M. and Axelrod, D.E. 1991.*Unequal cell division, growth regulation and colony size of mammalian cells: A mathematical model and analysis of experimental data*. Journal of Theoretical Biology 153: 157-180.