

UN ANÁLISIS DEL TEOREMA DEL UMBRAL DE PETER WHITTLE PARA LA EPIDEMIA GENERAL ESTOCÁSTICA

*Elizabeth Doig*¹

Octubre, 2008

Resumen

En la presente exposición se pretende establecer una comparación entre el teorema del umbral de Kermack y McKendrick correspondiente a un modelo epidémico determinístico y el teorema del umbral de Peter Whittle diseñado para un modelo epidémico estocástico.

Clasificación AMS 2000: 92D25, 92D30

Palabras Clave: *Modelos Epidémicos determinísticos y estocásticos,
Procesos Estocásticos, Modelo SIR.*

1. *Sección Matemáticas, Departamento de Ciencias, PUCP.*

1. Preliminares

El Modelo Epidémico Determinístico General en una Población Homogénea

En el modelo clásico para una *epidemia general* que ahora describiremos, el tamaño de la población N se toma como una cantidad fija para el tamaño de la epidemia, y se considera que los individuos están mezclados homogéneamente. Los individuos infectos pueden morir, ser aislados, o pueden recuperarse y volverse inmunes. Se cuentan a los individuos en la población según el estado de su enfermedad, mientras $S(t)$ representa el número de susceptibles, $I(t)$ representa el número de infectos y $R(t)$ el número de removibles (muertos, aislados o inmunes), así que $S(t)$ es no-creciente, $R(t)$ no-decreciente y la suma

$$S(t) + I(t) + R(t) = N, \quad \forall t \geq 0.$$

La ecuación diferencial para $S(t)$ es:

$$\frac{dS}{dt} = -\alpha SI \tag{1.1}$$

donde $\alpha > 0$ es el parámetro de infección, y $(S, I, R)(0) = (s_0, i_0, r_0)$ con $i_0 \geq 1$, $r_0 = 0$. El número de individuos infectos simultáneamente aumenta en la misma razón como el número de susceptibles disminuye, y disminuye a través de los que se remueven (por muerte, aislamiento o inmunidad) a una razón per cápita $\beta > 0$, así que:

$$\frac{dI}{dt} = \alpha SI - \beta I \tag{1.2}$$

Finalmente el número de individuos removidos aumenta en exactamente la misma razón como la pérdida de individuos infectos, de manera que:

$$\frac{dR}{dt} = \beta I \tag{1.3}$$

Observar que:

$$\left(\frac{d}{dt}\right) (S(t) + I(t) + R(t)) = 0,$$

ya que es consistente con el tamaño de la población total que permanece fijo en N .

En el artículo titulado "Una contribución a la teoría matemática de las epidemias", que escribieron Kermack y McKendrick (1927) propusieron estas ecuaciones como un modelo simple que describe el curso de una epidemia. Podemos escribir (1.1) y (1.3) como:

$$\frac{1}{S} \cdot \frac{dS}{dt} = -\frac{\alpha}{\beta} \cdot \frac{dR}{dt} = -\frac{1}{\rho} \cdot \frac{dR}{dt}, \quad (1.4)$$

donde $\rho \equiv \beta/\alpha$ es la tasa relativa removible. Integrando directamente esta ecuación diferencial, y usando los valores iniciales s_0 y $r_0 = 0$, obtenemos:

$$S(t) = s_0 e^{-R(t)/\rho} \quad (1.5)$$

A continuación se enuncia el teorema de Kermack y McKendrick

Teorema 1.1 (Kermack y McKendrick)

Una epidemia general se desarrolla según las ecuaciones diferenciales (1.1-3) de valores iniciales $(s_0, i_0, 0)$, donde $s_0 + i_0 = N$.

- (i). (La supervivencia y el Tamaño Total) Cuando termina la infección la propagación de la epidemia cesa y hay un número de individuos susceptibles S_∞ , que no se ha infectado. El número total r_∞ de individuos que se infectaron finalmente y que fueron removidos es igual a $s_0 + i_0 - s_\infty$, y es la única raíz de la ecuación

$$N - r_\infty = s_0 + i_0 - r_\infty = s_0 e^{-r_\infty/\rho} \quad (1.6)$$

donde $i_0 < r_\infty < s_0 + i_0$, $\rho = \beta/\alpha$ la tasa relativa de removimiento.

- (ii). (El Teorema del Umbral). Una mayor propagación ocurre si y sólo si $\frac{dI}{dt}(0) > 0$, lo que sucede únicamente si el número inicial de susceptibles es $s_0 > \rho$.
- (iii). (Segundo Teorema del Umbral) Si s_0 excede a ρ por una cantidad pequeña v , y si el número inicial de infectos i_0 es relativamente pequeño respecto de v , entonces el número final de susceptibles que queda en la población es aproximadamente $\rho - v$, y $r_\infty \approx 2v$.

A partir de esta información se puede definir convenientemente la Intensidad de la Epidemia, de la siguiente manera:

$$I_E = \frac{\text{Número total de Removidos}}{N}; \quad N = \text{Tamaño Total de la Población.}$$

Así por ejemplo si $I_E = 0,9$, aproximadamente 90 % de la población son afectados por la enfermedad y sólo 10 % permanecen susceptibles al final de la epidemia.

2. La Epidemia Estocástica General

Consideremos un modelo SIR donde la población total $N_S + N_I$ está subdividida en $S(t)$ individuos susceptibles, $I(t)$ individuos infectos, y $R(t)$ individuos inmunes o recuperados, con $(S, I, R)(0) = (N_S, N_I, 0)$ y

$$S(t) + I(t) + R(t) = N_S + N_I \tag{2.1}$$

Considerar que $\{(S, I)(t) : t \geq 0\}$ es un proceso de Markov bivariado; (2.1) asegura que $R(t) = N_S + N_I - S(t) - I(t)$ es conocido cuando $(S, I)(t)$ es conocido. Sea i el número de individuos susceptibles y j el número de individuos infectos, en el instante t . Se asume de nuevo que hay mezcla homogénea para que en el intervalo de tiempo $(t, t + \Delta t)$ las infecciones ocurran con la tasa $\alpha ij \Delta t$ y recuperaciones con la tasa

$\beta j \Delta t$, manteniéndose para las probabilidades de transición infinitesimal

$$\begin{aligned} P[(S, I)(t + \Delta t) = (i - 1, j + 1) | (S, I)(t) = (i, j)] &= \alpha i j \Delta t + o(\Delta t), \\ P[(S, I)(t + \Delta t) = (i, j - 1) | (S, I)(t) = (i, j)] &= \beta j \Delta t + o(\Delta t), \\ P[(S, I)(t + \Delta t) = (i, j) | (S, I)(t) = (i, j)] &= 1 - (\alpha i + \beta) j \Delta t + o(\Delta t), \end{aligned}$$

donde α es el parámetro de infección y β es el parámetro de recuperación. Entonces si escribimos las probabilidades de estado como:

$$p_{ij}(t) = P[(S, I)(t) = (i, j) | (S, I)(0) = (N_S, N_I)], \quad (2.2)$$

se puede obtener rápidamente *las ecuaciones de Kolmogorov* hacia adelante en la forma

$$\begin{aligned} \frac{dp_{N_S N_I}(t)}{dt} &= -N_I(\alpha N_S + \beta)p_{N_S N_I}(t) \\ \frac{dp_{ij}(t)}{dt} &= \alpha(i + 1)(j - 1)p_{i+1, j-1}(t) - j(\alpha i + \beta)p_{ij}(t) + \\ &\quad + \beta(j + 1)p_{i, j+1}(t) \end{aligned}$$

$$(0 \leq i + j \leq N_S + N_I, 0 \leq i \leq N_S, 0 \leq j \leq N_S + N_I), \quad (2.3)$$

sujetas a las condiciones iniciales $p_{N_S N_I}(0) = 1, p_{ij}(0) = 0$ en otro caso. Para un uso posterior se definirá a la distribución del tamaño final de la epidemia $\{P_n\}_{(n \geq 0)}$; es decir, la distribución del número de individuos susceptibles inicialmente que se infecta finalmente como:

$$P_n = P \left[\lim_{t \rightarrow \infty} (S, I)(t) = (N_S - n, 0) | (S, I)(0) = (N_S, N_I) \right] = p_{N_S - n, 0}(\infty)$$

Se obtiene una representación más simple de la ecuación si se define $t' = \alpha t$ así que $d/dt = \alpha d/dt'$. Con $\rho = \frac{\beta}{\alpha}$ como la tasa de recuperación relativa, las ecuaciones (2.3) se convierten en (escribiendo $t = t'$ por

conveniencia)

$$\frac{dp_{N_S N_I}(t)}{dt} = -N_I(N_S + \rho)p_{N_S N_I}(t),$$

$$\begin{aligned} \frac{dp_{ij}(t)}{dt} = & (i+1)(j-1)p_{i+1,j-1}(t) - j(i+\rho)p_{ij}(t) + \\ & + \rho(j+1)p_{i,j+1}(t) \end{aligned}$$

$$(0 \leq i+j \leq N_S + N_I, 0 \leq i \leq N_S, 0 \leq j \leq N_S + N_I), \quad (2.4)$$

sujeto a la misma condición inicial $p_{N_S N_I}(0) = 1$. Podríamos ahora intentar resolver estas ecuaciones recursivamente, empezando con $p_{N_S N_I}(t) = e^{-(N_S + \rho)N_I t}$. Aunque también se puede usar una aproximación sistemática como la que se basa en una combinación de la transformada de Laplace y los métodos de f.g.p.

3. Teorema del Umbral de Whittle para la Epidemia Estocástica General

Similar en naturaleza, al teorema de Kermack y McKendrick para el modelo determinístico de la epidemia general, Peter Whittle construye un teorema para el umbral estocástico, con el objetivo de ver el desarrollo general de la epidemia; es decir, observar si la intensidad de la epidemia aumenta o disminuye. Empecemos considerando al valor N_S de la población susceptible como un valor muy grande [Bartlett 1978], así que

$$P[(S, I)(t + \Delta t) = (i - 1, j + 1) | (S, I)(t) = (i, j)] \approx N_S j \Delta t + o(\Delta t)$$

en lugar del valor exacto $ij \Delta t + o(\Delta t)$, y

$$P[(S, I)(t + \Delta t) = (i, j - 1) | (S, I)(t) = (i, j)] = \rho j \Delta t + o(\Delta t)$$

Se denota como $\tilde{I}(t)$ al proceso marginal del número de individuos infectos que satisface estas relaciones aproximadamente, observando que $\{\tilde{I}(t) : t \geq 0\}$ con $\tilde{I}(0) = N_I$ es un proceso de nacimiento y muerte con parámetros N_S para los nacimientos y ρ para las muertes. La f.g.p. de \tilde{I} está dada por

$$\varphi(z; t) = \begin{cases} \left[\frac{\rho e^{(N_S - \rho)t}(z - 1) - (N_S z - \rho)}{N_S e^{(N_S - \rho)t}(z - 1) - (N_S z - \rho)} \right]^{N_I}, & \text{si } N_S \neq \rho, \\ \left[\frac{1 - (\rho t - 1)(z - 1)}{1 - \rho t(z - 1)} \right]^{N_I}, & \text{si } N_S = \rho. \end{cases} \quad (3.1)$$

Notar que el resultado para $N_S = \rho$ puede derivarse fácilmente de que para $N_S \neq \rho$ seleccionando $N_S = \rho + \varepsilon$ en este caso y haciendo que $\varepsilon \rightarrow 0$. Estamos interesados en la probabilidad de extinción de $\tilde{I}(t)$, es decir en $\lim_{t \rightarrow \infty} P[\tilde{I}(t) = 0] = \lim_{t \rightarrow \infty} \varphi(0, t)$. De la primera de las ecuaciones de (3.1) vemos que el comportamiento difiere entre $N_S > \rho$ y $N_S < \rho$, a saber,

$$\lim_{t \rightarrow \infty} \varphi(0; t) = \begin{cases} \lim_{t \rightarrow \infty} \left[\frac{\rho - \rho e^{-(N_S - \rho)t}}{N_S - \rho e^{-(N_S - \rho)t}} \right]^{N_I} = \left[\frac{\rho}{N_S} \right]^{N_I} < 1, & \text{si } N_S > \rho, \\ \lim_{t \rightarrow \infty} \left[\frac{\rho - \rho e^{(N_S - \rho)t}}{\rho - N_S e^{(N_S - \rho)t}} \right]^{N_I} = 1, & \text{si } N_S < \rho. \end{cases} \quad (3.2)$$

Cuando $N_S = \rho$ la segunda ecuación en (3.1) mantiene el límite 1, el cual es igual al límite de ambas expresiones en (3.2). Así, para $N_S \leq \rho$, $\lim_{t \rightarrow \infty} P[\tilde{I}(t) = 0] = 1$ luego es probable que la propagación epidémica sea pequeña. Por otro lado, para $N_S > \rho$, el límite es positivo pero menor que 1, así que la propagación puede ser pequeña o grande. Whittle (1955) hizo un análisis más preciso que la aproximación de Bartlett; el paso esencial es acotar el número de infectos $I(t)$ en el proceso epidémico actual entre dos procesos de nacimiento y muerte $\tilde{I}(t)$.

Para cualquier ξ en $]0, 1[$ podemos preguntar si la intensidad de la epidemia, significa que la proporción de susceptibles que se han infectado, excede a ξ . Para terminar se define

$$\pi(\xi) = \lim_{t \rightarrow \infty} P[S(0) - S(t) \leq N_S \xi] = \sum_{n=0}^{\lfloor N_S \xi \rfloor} P_n \quad (3.3)$$

con P_n la probabilidad de que el tamaño final de la epidemia sea n , como se definió en (2.3). Se puede acotar la componente $I(t)$ en el proceso epidémico $\{(S, I)(t) : t \geq 0\}$ por $\tilde{I}(t)$ en dos procesos bivariados

$$\{\tilde{I}(t), U(t) = N_I + S(0) - S(t) : t \geq 0\}$$

donde $\tilde{I}(\cdot)$ es un proceso de nacimiento y muerte con parámetro de nacimiento ya sea $\lambda_1 = N_S$, como en (3.1), o $\lambda_2 = N_S(1 - \xi)$, y con parámetro de muerte $\mu_1 = \mu_2 = \rho$. Notar que $U(t)$ representa a todos los individuos que han sido infectados en un tiempo mayor que t , incluyendo a los N_I individuos inicialmente infectos. Sea

$$P_{j\ k}(t) = P\{(I, U)(t) = (j, k) | (I, U)(0) = (N_I, N_I)\}.$$

y

$$\varphi(z, w; t) = \sum_{j, k} p_{jk}(t) z^j w^k$$

Ahora las ecuaciones de Kolmogorov hacia delante para el proceso de nacimiento y muerte con el parámetro λ de nacimiento, son:

$$\begin{aligned} \frac{dp_{jk}}{dt} &= \lambda(j-1)p_{j-1, k-1} - (\lambda + \rho)jp_{jk} + \rho(j+1)p_{j+1, k} \\ &\quad (0 \leq j \leq N_S + N_I, N_I \leq k \leq N_S + N_I), \end{aligned}$$

definiendo $p_{jk} = 0$ cuando (j, k) cae fuera del rango permisible. Esto lleva a las ecuaciones diferenciales parciales

$$\frac{\partial \varphi}{\partial t} = [\lambda z^2 w + (\lambda + \rho)z + \rho] \frac{\partial \varphi}{\partial z}, \quad (3.4)$$

con la condición inicial $\varphi(z, w; 0) = z^{N_S} w^{N_I}$. La Ecuación (3.4) puede resolverse por los métodos clásicos.

Obteniéndose por el método de Lagrange:

$$\varphi(z, w; t) = w^{N_I} \left[\frac{\eta_1(z - \eta_2)e^{-\lambda z(\eta_1 - \eta_2)t} + \eta_2(\eta_1 - z)}{(z - \eta_2)e^{-\lambda z(\eta_1 - \eta_2)t} + (\eta_1 - z)} \right]^{N_I}, \quad (3.5)$$

Como $t \rightarrow \infty$, la distribución asintótica del número total de individuos $U(t)$ que finalmente se infecta está dada por:

$$\lim_{t \rightarrow \infty} \varphi(1, w; t) = w^{N_I} \eta_2^{N_I}(w) = \left[\frac{\lambda + \rho}{2\lambda} \right]^{N_I} (1 - \sqrt{1 - kw})^{N_I}, \quad (3.6)$$

donde $k = \frac{4\lambda\rho}{(\lambda + \rho)^2}$. Se extiende esta relación, usando la expansión de Lagrange para la función $\psi(s) = s^{N_I}$, donde $s(w) = 1 - \sqrt{1 - kw}$ se puede dar convenientemente como la raíz de $s^2 - 2s + kw = 0$, que satisface $\lim_{w \rightarrow 0} s(w) = 0$. Notar que

$$w = \frac{s(2 - s)}{k} = \frac{s}{k/(2 - s)};$$

por lo tanto usando la fórmula

$$\psi(s(w)) = \sum_{n=1}^{\infty} \frac{w^n}{n!} \left[\frac{d^{n-1}}{ds^{n-1}} \left[N_I s^{N_I-1} \left(\frac{k}{2-s} \right)^n \right] \right]_{s=0}$$

para la expansión de Lagrange, vemos que lleva al siguiente resultado:

$$P_n(\lambda) = \lim_{t \rightarrow \infty} P[U(t) = N_I + n] = \lim_{t \rightarrow \infty} P[S(0) - S(t) = n] =$$

$$= \begin{cases} N_I \frac{(2n + N_I - 1)!}{n!(n + N_I)!} \frac{\lambda^n \rho^{n+N_I}}{(\lambda + \rho)^{2n+N_I}}, & (0 \leq n \leq N_S - 1) \\ 1 - \sum_{m=0}^{N_S-1} P_m(\lambda), & (n = N_S) \end{cases} \quad (3.7)$$

donde $\{P_n(\lambda)\}$ es la distribución de probabilidad para el número total de individuos inicialmente susceptibles que finalmente se infectaron, para un parámetro específico de nacimiento λ .

Ahora sabemos de (3.6) que para $N_S \rightarrow \infty$,

$$\lim_{t \rightarrow \infty} \varphi(1, 1; t) = \sum_{n=0}^{\infty} P_n(\lambda) = \begin{cases} \left(\frac{\rho}{\lambda}\right)^{N_I}, & (\rho < \lambda) \\ 1, & (\rho \geq \lambda) \end{cases} \quad (3.8)$$

así que

$$\sum_{n=0}^{\infty} P_n(\lambda) = \left[\min\left(\frac{\rho}{\lambda}, 1\right) \right]^{N_I} \quad (3.9)$$

Por lo tanto para N_S suficientemente grande, en el caso de una epidemia de intensidad ξ , con limitar los valores $\lambda_1 = N_S \geq S(t) \geq \lambda_2 = N_S(1 - \xi)$ para los parámetros de nacimiento del proceso de nacimiento y de muerte involucrados, tenemos aproximadamente

$$\sum_{n=0}^{\lfloor N_S \xi \rfloor} P_n(\lambda_1) \leq \sum_{n=0}^{\lfloor N_S \xi \rfloor} P_n \leq \sum_{n=0}^{\lfloor N_S \xi \rfloor} P_n(\lambda_2),$$

o equivalentemente

$$\left[\min\left(\frac{\rho}{N_S}, 1\right) \right]^{N_I} \leq \pi(\xi) \leq \left[\min\left(\frac{\rho}{N_S(1 - \xi)}, 1\right) \right]^{N_I} \quad (3.10)$$

De este resultado se deducen las siguientes conclusiones:

Teorema 3.1 (Teorema del Umbral de Whittle). Considerar un proceso epidémico general con números iniciales de individuos susceptibles N_S e infectos N_I , y tasa relativa de removimiento ρ . Para cualquier ξ en $]0, 1[$, $\pi(\xi)$ denota la probabilidad de que a lo sumo $\lfloor N_S \xi \rfloor$ de los individuos susceptibles se infectan finalmente; es decir, que la intensidad de la epidemia no excede a ξ .

(i). Si $\rho < N_S(1 - \xi)$, entonces

$$\left(\frac{\rho}{N_S}\right)^{N_I} \leq \pi(\xi) \leq \left(\frac{\rho}{N_S(1 - \xi)}\right)^{N_I}$$

(ii). Si $N_S(1 - \xi) \leq \rho \leq N_S$, entonces:

$$\left(\frac{\rho}{N_S}\right)^{N_I} \leq \pi(\xi) \leq 1.$$

En ambos casos hay aproximadamente una probabilidad igual a $1 - (\rho/N_S)^{N_I}$, de que la epidemia logre una intensidad mayor que ξ , para un valor pequeño ξ .

(iii). Si $\rho \geq N_S$ entonces $\pi(\xi) = 1$, y la probabilidad de que la epidemia logre una intensidad mayor que cualquier ξ predeterminado en $]0, 1[$, es cero.

Referencias

- [1] Anderson, Hakan and Britton, Tom. *Stochastic Epidemic Models and Their Statistical Analysis*. Lecture Notes in Statistics. New York: Springer - Verlag; 2000.
- [2] Bartlett, M.S. *Some evolutionary stochastic processes*. J. R. Statistic. Soc.; 1949; Ser. B; Vol. 11; 211 - 229.
- [3] Bartlett, M. S. *An introduction to stochastic processes*. 3^{ra} Edición. Cambridge: Cambridge University Press; 1978.
- [4] Doig, Elizabeth. *Teorema del Umbral de Kermack y McKendrick para un modelo en tiempo discreto*. Pro Mathematica PUCP. Vol. XVII/ No. 33. 2003
- [5] Kermack, W. O. and McKendrick, A. G. *A contribution to the mathematical theory of epidemics*. London: Proc. Roy. Soc. London; 1927; Series A Vol. 115; 700 - 721.
- [6] Whittle, P. *The outcome of a stochastic epidemic-a note on Bailey's paper*. Biometrika; 1955; Vol. 42; 116 - 122.

- [7] Bailey, N. T. J. (1953). *The total size of a general stochastic epidemic* ("El tamaño total de una epidemia estocástica general") *Biometrics*, **40**, 177.
- [8] Bartlett, M. S. (1946). *Stochastic Processes* ("Procesos Estocásticos") Mimeografiado, Notas de ponencias de Carolina del Norte.
- [9] Bartlett, M. S. (1949). *Some evolutionary stochastic processes* ("Algunos procesos estocásticos evolutivos") *J.B. Statist. B*, **11**, 211.
- [10] Kendall, D. C. (1948). *On the generalized "birth-and-death" process*. ("Acerca del proceso "nacimiento - y - muerte" generalizado") *Ann. Math. Statist.* **19**, 1.

Abstract

In the present exposition we seek to establish a comparison among the Kermack and McKendrick's threshold theorem for the deterministic epidemic model and the Peter Whittle's threshold theorem for the stochastic epidemic model.

Keywords: Deterministic and Stochastic Epidemic Models, Stochastic Processes, SIR Model.

Elizabeth Doig
Sección Matemáticas,
Departamento de Ciencias,
Pontificia Universidad Católica del Perú
edoig@pucp.edu.pe