

Reunión de Discusión Nº 79

Fecha: 13/ 4/1994

Hs.: 16

Mal de Chagas: Estimación de las Tasas de Infectación(*)

EUSEBIO CLETO del REY

1. Introducción

Para estimar los beneficios brutos totales que se derivan de la prevención del mal de Chagas, a partir de los beneficios pro-

(*) . This investigation received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR). Este trabajo es una versión mejorada de una parte de: DEL REY, E. C.: "Cálculo del Stock de Salvados del Mal de Chagas", Reunión de Discusión Nº 77, I. I. E., F. C. E. J. y S., U. N. Sa., 19/12/1993, y surgió del Proyecto Nº 357/4 del Consejo de Investigación de la UNSa.

medio por enfermo¹, necesitamos conocer la cantidad de personas salvadas de este mal gracias a la prevención, existentes en cada uno de los años correspondientes al lapso que empieza en el momento en que se realiza la fase de ataque y continúa hasta infinito.

Una forma de conocer esa cantidad de personas, a la que denominamos "stock de salvados del mal de Chagas", es estimar la tasa anual de infectación con la enfermedad, correspondiente a una población protegida por la prevención, y la tasa anual de infectación de una población no protegida por esa prevención.

Aplicando cada una de las mencionadas tasas a la población de cada edad, no infectada al momento de realizarse el ataque², encontraremos el número estimado de no infectados con Chagas que corresponde a cada edad de la población protegida, y el que corresponde a esa misma edad y esa misma población sin protección, y repetiremos este proceso para cada año, a partir del momento del ataque. Restamos luego nuestra estimación de la cantidad de no infectados para el caso con protección de aquella que corresponde a la misma población sin protección, y obtenemos el stock de salvados del Chagas para cada año y cada edad. Por sumatoria sobre las edades calcularemos entonces el stock total de salvados de cada año.

El problema considerado en este trabajo consiste en esti-

1. Cuyo cálculo se está realizando, con resultados parciales detallados en: DE REY, E. C., BASOMBRIÓ, M. A., ROJAS, C. L. Y FAINGUERSCH, D. P.: "Costos de los Tratamientos de Mal de Chagas", Asociación Argentina de Economía Política: Anales: XXVIII Reunión Anual, Tucumán, 1993, Tomo II, pág. 455/478.

2. Obtenida en base a datos censales de 1991.

mar: a) La tasa de infectación anual correspondiente a la población considerada, si esta se encuentra protegida por el rociado de las viviendas, tendiente a controlar el vector; b) La tasa de infectación anual correspondiente a la población considerada, si ella no está protegida por el rociado antes mencionado.

El modelo que nos permite realizar la estimación se basa en las siguientes presunciones: 1) Dada una cohorte de personas de la misma edad, residentes en la misma área geográfica, originariamente (esto es, en el momento de su nacimiento) no infectada con Chagas, suponemos que ella se infecta con ese mal a una tasa anual i_n , si no está protegida por rociado, o a una tasa anual i , si está protegida por ese medio, ambas aplicables en forma continua ("capitalización" continua); 2) No se registra migración.

En la Sec. 3 presentamos el modelo, en la Sec. 4 desarrollamos el método de estimación y en la Sec. 5 hacemos las consideraciones finales.

2. Simbología

En este trabajo emplearemos la simbología que se define a continuación¹:

x edad en años cumplidos

P_x población de edad x

1. Recomendamos al lector pasar directamente a la Sec. 3, y consultar la 2 cuando necesite conocer el significado de los símbolos.

i tasa anual promedio de infectación de la población, si está protegida por rociado

N_x número de no infectados con Chagas existente en la población de edad x

i_n tasa anual promedio de infectación de la población, si no está protegida por rociado

a número de años transcurridos desde el momento en que se realizó la fase de ataque, en el área geográfica estudiada, y aquel en el que se lleva cabo la observación

$n_x = \frac{N_x}{P_x}$ proporción de no infectados con Chagas, existente en la población de edad x

x' edad mínima, con datos utilizables, comprendida en la muestra

P_x^m número de personas comprendidas en la muestra, que tiene x años de edad

N_x^m número de no infectados con Chagas, de edad x , comprendidos en la muestra

u_x valor que toma la variable aleatoria u (a la que suponemos distribuida con media nula y varianza σ^2 , igual para todas las edades), en la muestra correspondiente a la edad x

$n_x^m = \frac{N_x^m}{P_x^m}$ proporción de no infectados con Chagas, de x años de edad, comprendidos en la muestra

n_p media geométrica de los n_x^m , para $x \leq a$.

$$\bar{x} = \frac{\sum_{x=x'}^a x}{a-x'+1} \quad \text{edad promedio de la muestra de personas protegidas, o sea para } x \leq a^1$$

$$\bar{u} = \frac{\sum_{x=x'}^a u_x}{a-x'+1} \quad \text{promedio muestral de los desvíos aleatorios, para } x \leq a$$

$\hat{i}$ estimador de la tasa de infectación i

x'' edad máxima, con datos utilizables, comprendida en la muestra

n_p' media geométrica de los n_x^m , para $x > a$.

$$\bar{x}^* = \frac{\sum_{x=a+1}^{x''} x}{x''-a} \quad \text{edad promedio de la muestra de personas parcialmente protegidas, o sea para } x > a$$

$$\bar{u}' = \frac{\sum_{x=a+1}^{x''} u_x}{x''-a} \quad \text{promedio muestral de los desvíos aleatorios, para } x > a$$

$\hat{i}_n$ estimador de la tasa de infectación i_n

1. En este caso, y en todos aquellos en los que empleamos sumatoria o productorio, suponemos que hay observaciones para todas las edades intermedias, entre la mínima (x' en este caso) y la máxima (a en este caso). Si ello no fuera cierto sería necesario un subíndice j que reemplace al subíndice x , y en el caso de las edades escribiríamos: x_j . Luego sumaríamos o haríamos el productorio sobre j .

3. El Modelo

Una población de edad x , P_x , que se estuvo infectando con el mal de Chagas a una tasa anual constante i , durante toda su vida, tiene el siguiente número de no infectados con la enfermedad:

$$N_x = P_x e^{-xi} \quad (1)$$

Donde: e es la base de los logaritmos naturales.

Si las condiciones, relacionadas con la enfermedad, a las que estuvo sometida esa población cambiaron en algún momento de su vida, de modo que se pueda esperar la existencia de dos tasas anuales de infectación, una de ellas, i_n , vigente durante los primeros $(x-a)$ años, y la segunda, i , vigente durante los a restantes, la cantidad de no infectados sería¹:

$$N_x = P_x e^{-(x-a)i_n - ai} \quad (2)$$

Esto también se puede escribir:

$$n_x = e^{-xi} \quad (3)$$

y

$$n_x = e^{-(x-a)i_n - ai} \quad (4)$$

respectivamente.

Para el caso particular en el que estamos interesados, aplicaríamos este modelo de la siguiente manera: En el área geográfica en la que se encuentra radicada esta población se realizó la fase de ataque del control del mal por rociado de

1. Podemos suponer que cada año cambian las circunstancias referentes a la enfermedad, y en tal caso el número de no infectados sería:

$$N_x = P_x e^{-\sum_j i_j} \quad \text{para: } j = 1, 2, \dots, x.$$

Pero a nosotros nos interesa el caso en el que hubo sólo un cambio en las condiciones.

viviendas a años antes del momento en que realizamos la observación y se estuvo en fase de vigilancia a lo largo de esos a años. Si $x \leq a$, la población estuvo infectándose a una tasa constante i , correspondiente al caso de gente que estuvo protegida por el rociado durante toda su vida. Si $x > a$, la población estuvo infectándose a una tasa anual constante i_n durante los años de su vida anteriores al rociado, y a una tasa i (que cabe esperar que sea menor que la anterior) en los últimos a años.

4. Estimación de las Tasas

Mediante un muestreo a realizarse en el área geográfica bajo estudio, determinaremos, por análisis serológico, la cantidad de no infectados con mal de Chagas (N_X^m). Tales datos, así como los respectivos tamaños de muestra (P_X^m), serán clasificados por edad de las personas observadas. Conociendo el momento en que tuvo lugar la fase de ataque en esa área, sabremos la cantidad de años transcurridos desde entonces hasta la fecha de nuestra observación, o sea a.

Con esos datos estamos en condiciones de estimar las tasas de infectación, como puede verse en las Sec. 4.1 y 4.2.

4.1. Tasa Anual de Infectación de Personas Protegidas del Mal

Para estimar esta tasa, debemos trabajar con datos referentes a personas (incluidas en la muestra) cuya edad sea menor o igual al tiempo transcurrido desde el momento en el que se realizó la fase de ataque en su lugar de residencia hasta el momento

en el que se realiza la observación. Tales personas estuvieron toda su vida protegidas por el rociado.

Por lo tanto, emplearemos los datos que corresponden a $x = x', x'+1, x'+2, \dots, a^1$.

La muestra de P_x^m personas, de cada una de esas edades, se estuvo infectando a la tasa i hasta el momento en que realizamos la observación, quedando, en ese momento, N_x^m de sus miembros sin infectar. Entonces tendremos:

$$N_x^m = P_x^m e^{-xi+u_x} \quad (5)$$

De donde:

$$n_x^m = e^{-xi+u_x} \quad (6)$$

Definimos la media geométrica², en este caso, como:

$$n_p = \left[\frac{a}{\pi} \int_{x=x'} n_x^m \right]^{1/(a-x'+1)} \quad (7)$$

1. Bajo el supuesto de que no faltan datos correspondientes a edades intermedias. Véase nota al pie 1, pág. 5.

2. Emplease la media geométrica no sólo porque ello es aconsejable cuando se trabaja con proporciones (como lo son, en este caso, los n_x^m), sino porque el estimador de i al que llegamos en esta sección es equivalente a calcular:

$$\hat{i}_x = - \frac{\ln n_x^m}{x} \quad \text{para: } x = x', x'+1, \dots, a$$

y luego hacer el promedio, ponderado por edad, de los $\hat{i}_x$:

$$\hat{i} = \frac{\sum_{x=x'}^a x \hat{i}_x}{\sum_{x=x'}^a x}$$

Empleando la media geométrica de la ecuación (7), y la ecuación (6), obtenemos:

$$n_p = \left[\pi_{x=x'}^a e^{-xi+u_x} \right]^{[1/(a-x'+1)]} \quad (8)$$

De donde:

$$n_p = \left[e^{-i \left(\sum_{x=x'}^a x \right) + \sum_{x=x'}^a u_x} \right]^{[1/(a-x'+1)]} \quad (9)$$

y

$$\ln n_p = \frac{-i \left[\sum_{x=x'}^a x \right] + \sum_{x=x'}^a u_x}{(a-x'+1)} \quad (10)$$

Donde: $\ln$ simboliza logaritmo natural. Reemplazando (Ver: Sec. 2):

$$\ln n_p = -i \bar{x} + \bar{u} \quad (11)$$

Finalmente resulta:

$$i = \frac{-\ln n_p + \bar{u}}{\bar{x}} \quad (12)$$

Como no conocemos $\bar{u}$, estimamos i del siguiente modo:

$$\hat{i} = - \frac{\ln n_p}{\bar{x}} \quad (13)$$

Nótese que:

$$\hat{i} = i - \frac{\bar{u}}{\bar{x}} \quad (14)$$

4.2. Tasa Anual de Infección de Personas no Protegidas del Mal

En este caso, debemos trabajar con datos referentes a personas (incluidas en la muestra) cuya edad sea mayor que el tiempo transcurrido desde el ataque por rociado hasta que se tomó la muestra. Tales personas estuvieron parcialmente protegidas por el rociado, ya que parte de sus vidas transcurrió antes de que se realizara la fase de ataque.

Por lo tanto, emplearemos los datos correspondientes a las edades $x = a + 1, a + 2, \dots, x^{11}$. Utilizaremos también la tasa $\hat{i}$ estimada en la sección anterior.

La gente de edad $x > a$, incluida en la muestra, a la que suponemos originariamente no infectada con Chagas: P_x^m , se irá infectando, hasta el momento del rociado, a una tasa anual i_n , y, a partir de ese momento, lo seguirá haciendo a una tasa anual i . A la edad mencionada, tal población tendrá N_x^m de sus miembros sin infectar, número que será:

$$N_x^m = P_x^m e^{-(x-a)i_n - ai + u_x} \quad (15)$$

o, lo que es lo mismo:

$$n_x^m = e^{-(x-a)i_n - ai + u_x} \quad (16)$$

1. Bajo el supuesto de que no faltan datos correspondientes a edades intermedias. Véase nota al pie 1, pág. 5.

Redefiniendo la media geométrica¹ de la ecuación (7), tendremos:

$$n_p' = \left[\prod_{x=a+1}^{x''} n_x^m \right]^{1/(x''-a)} \quad (17)$$

De (16) y (17) obtenemos:

$$n_p' = e^{-ai} \left[\prod_{x=a+1}^{x''} e^{-(x-a)i_n + u_x} \right]^{1/(x''-a)} \quad (18)$$

Tomando logaritmos:

$$\ln n_p' = -a i - (\bar{x}^* - a) i_n + \bar{u}' \quad (19)$$

De donde resulta:

$$i_n = \frac{-\ln n_p' - a i + \bar{u}'}{\bar{x}^* - a} \quad (20)$$

Para estimar esta tasa empleamos:

1. La equivalencia señalada en la nota al pie 2 de la pág. 8 se da, en este caso, entre el estimador de i_n al que arribamos en esta sección, y el promedio ponderado calculado del siguiente modo:

Primero obtenemos:

$$\hat{i}_{nx} = \frac{-\ln n_x^m - a \hat{i}}{x - a}$$

Y luego promediamos:

$$\hat{i}_n = \frac{\sum_{x=a+1}^{x''} (x-a) \hat{i}_{nx}}{\sum_{x=a+1}^{x''} (x-a)}$$

Nótese que para calcular todas y cada una de las $\hat{i}_{nx}$ se usa el mismo valor de $\hat{i}$, o sea su promedio muestral.

$$\hat{i}_n = \frac{-\ln n_p' - a \hat{i}}{\bar{x}^* - a} \quad (21)$$

Nótese que:

$$\hat{i}_n = i_n - \frac{\bar{u}'}{\bar{x}^* - a} + \frac{a}{\bar{x}(\bar{x}^* - a)} \bar{u} \quad (22)$$

5. Observaciones Finales y Problemas a Resolver

Consideramos aquí las siguientes cuestiones: En la Sec. 5.1, la unidad de observación; en la 5.2 el tamaño de muestra en cada edad; el problema de las características de los estimadores en la Sec. 5.3, y el de la migración en 5.4.

5.1. Unidad Muestral

No es fácil ver de inmediato cuál es la unidad muestral o unidad de observación en este problema, sino que resulta necesario considerar el asunto con cuidado. En efecto, existen tres unidades alternativas: 1) Cada una de las personas cuya sangre analizamos; 2) Cada una de las edades en las que clasificamos a esas personas; 3) Cada año de exposición al riesgo de contraer la enfermedad, correspondiente a un grupo de individuos de la misma edad.

La alternativa 1) es importante desde el punto de vista de la calidad de n_x^m , puesto que si P_x^m es demasiado pequeño, nuestra estimación de la proporción de no infectados de x años de edad sería muy variable (de muestra a muestra), y, por lo tanto, "poco

representativa".

Pero nosotros trabajamos con individuos agrupados por edad, y nos preocupamos porque esté bien estimada la proporción de no infectados para cada valor de x . Ello parece indicar que la unidad de observación es cada edad. Así, si hemos observado 100 individuos, dispersos en 50 edades diferentes, todas ellas con pocas observaciones, la muestra nos resulta inútil. Si, por el contrario, los 100 individuos se dividen aproximadamente por igual entre sólo dos edades, podríamos emplear los datos.

Por otra parte, si tenemos en cuenta las observaciones hechas en la notas al pie 2 de la pág. 8 y 1 de la pág. 11, donde vemos que nuestro método de estimación es equivalente a hacer un promedio ponderado de la tasa de infectación muestral de cada edad, usando como ponderador el número de años de esa edad, podemos concluir que nuestra unidad de observación es el año de exposición al riesgo de infectación, o sea la alternativa 3. El problema radica, en este caso, en que no observamos a las unidades muestrales (años de exposición) de una en una, sino que lo hacemos en bloque. Así, por ejemplo, al observar el grupo de 10 años de edad estamos tomando diez unidades muestrales juntas: primer año de exposición, segundo año de exposición, ..., décimo año de exposición.

5.2. Número de Individuos en cada Edad

En la Sec. 5.1 dijimos que el tamaño de la muestra de personas de cada edad, P_{x^m} , debe ser suficientemente grande como para que la proporción de no infectados se estabilice. La pregunta que

ahora nos hacemos es: ¿Cuántos individuos constituyen una muestra que merezca el calificativo de "suficientemente grande"?

Para poder contestar esta pregunta disponemos de los datos de los tamaños de las poblaciones a muestrear, que nos pueden ayudar a establecer el tamaño de muestra. En efecto, contamos con el dato censal de cantidad de personas existentes en Anta, clasificadas por edad y municipio¹. Lo que es necesario definir es que proporción de la cantidad de habitantes de cada población constituye un tamaño de muestra aceptable. La Encuesta Permanente de Hogares trabaja con un 2 % de la población, en tanto que, cuando se realiza un muestreo sobre datos censales, se emplea como porcentaje un 20 % de los mismos. Estos porcentajes² pueden servirnos de guía en nuestra elección del tamaño de la muestra.

5.3. Características de los Estimadores

El único supuesto que hicimos en la Sec. 2, respecto a la variable aleatoria u , es que ella se distribuye con media nula y varianza σ^2 , igual para todas las edades. Ese supuesto nos basta para afirmar que nuestras estimaciones de las tasas de infección son insesgadas. En efecto, siendo, por el mencionado supuesto:

$$E(u) = 0 \quad (23)$$

Será, por aplicación de la ecuación (23) a la (14):

$$E(\hat{i}) = i \quad (24)$$

1. Censo de 1991. Datos provistos por la Dirección de Estadísticas y Censos de la Provincia de Salta.

2. Que nos fueron sugeridos por el Lic. Jorge A. Paz, a quien agradecemos por ello.

y, por su aplicación a la ecuación (22):

$$E(\hat{i}_n) = i_n \quad (25)$$

Sin embargo, el mencionado supuesto no nos basta para conocer otras posibles características de estos estimadores. Podríamos avanzar en este sentido si suponemos que u se distribuye como $N(0, \sigma^2)$. Tal supuesto nos permitiría, además, conocer el sesgo de nuestra estimación del stock de salvados del mal de Chagas, que toma forma exponencial¹. Así, por ejemplo, el estimador de la cantidad de no infectados de edad x que existen en el año considerado, estando bajo protección (N_x^p), es:

$$N_x^p = P_x e^{-x\hat{i}} \quad (26)$$

Si u es $N(0, \sigma^2)$, la distribución de N_x^p será lognormal, y este estimador tendrá un sesgo conocido, al que es posible corregir.

Pero dada la forma en que se generan nuestros datos ¿Es aceptable suponer normalidad?

5.4. Migración

Nuestro modelo supone que no existe migración (Véase Sec. 1, punto 2). Pero si en la realidad ella existe, y no fuera tomada en cuenta, puede sesgar nuestros resultados. Así, por ejemplo, si numerosas personas con $x \leq a$, comprendidas en la muestra, vivieron largo tiempo en una zona donde no se realizó rociado y se

1. Véase: DEL REY, Eusebio Cleto: "Cálculo del Stock de Salvados del Chagas", op. cit., pág. 8/10. Ténganse en cuenta las diferencias en la simbología.

trasladaron luego a la zona bajo estudio, nuestro n_p tenderá a ser una subestimación, y lo será también $\hat{i}$.

Una forma de evitar ese tipo de problemas es hacer, a cada individuo incluido en la muestra, preguntas del siguiente tipo: "¿Nació Ud. en esta localidad?", y ante una respuesta positiva agregar: "¿Vivió siempre en ella?". Una nueva respuesta positiva implica que el sujeto no presenta problemas de migración, y debe ser incluido en la muestra. Sin embargo, la respuesta negativa a una de las dos preguntas no necesariamente implica que deba descartarse a la persona. Por ejemplo, si un individuo tiene 10 años de edad, nació fuera de la localidad, vino a ella de pocos meses y no volvió a emigrar, tiene en él información lo suficientemente buena como para no desperdiciarla. Se hace necesario, por lo tanto, preparar preguntas adicionales para contemplar esos casos.

Puede ocurrir que en la tarea de campo nos encontremos con el hecho de que la migración no es muy frecuente y, por lo tanto, que descartando a todo el que migró podemos lograr una muestra de tamaño aceptable. En tal caso, puesto que hacer las preguntas adicionales tiene un costo en términos de tiempo y esfuerzo, es conveniente excluir, sin más consideración, a los que migraron. Pero vale la pena hacer esas preguntas si el número de inmigrados es grande, y su exclusión puede conducirnos a costos aún mayores (por ejemplo, a incluir una escuela más para lograr el tamaño de muestra deseado). El método podría ser formular las dos primeras preguntas, a todos los individuos disponibles, antes de iniciar la extracción de sangre, y, de acuerdo a los resultados, decidir si se excluirá a todos los que estuvieron fuera de la localidad o se formularán otras preguntas, que permitan saber si la inclusión

de cada individuo es o no aceptable.

BIBLIOGRAFIA:

1. DEL REY, Eusebio Cleto: "Cálculo del Stock de Salvados del Mal de Chagas", Reunión de Discusión Nº 77, I. I. E., F. C. E. J. y S., U. N. Sa., 1º/12/1993.
2. DEL REY, Eusebio Cleto y BASOMBRIIO, Miguel Angel: "Análisis de Costos y Beneficios de la Prevención del Mal de Chagas - Metodología", Asociación Argentina de Economía Política: Anales: XXVI Reunión Anual, Santiago del Estero, 1991, Tomo I, pág. 339/67.
3. DEL REY, Eusebio Cleto y BASOMBRIIO, Miguel Angel: "Costos Y Beneficios de la Prevención del Mal de Chagas. Una Aproximación Metodológica", Estudios, Año XV, Nº 61, Enero/Marzo 1992, pág. 3/12.
4. DEL REY, E. C., BASOMBRIIO, M. A., ROJAS, C. L. y FAINGUERSCH, D. P.: "Costos de los Tratamientos del Mal de Chagas", Asociación Argentina de Economía Política: Anales: XXVIII Reunión Anual, Tucumán, 1993, Tomo II, pág. 455/78.
5. GÜRTLER, R. E., KRAVETZ, F. O., PETERSEN, R. M., LAURICELLA, M.A. and WISNIVERSKY-COLLI, C.: "The Prevalence of Trypanosoma cruzi and the demography of dogs populations after insecticidal spraying of houses: a predictive model", Annals of Tropical Medicine and Parasitology, Vol. 84, Nº 4, 1990, pag. 313/23.

Universidad Nacional de Salta
Facultad de Ciencias Económicas,
Jurídicas y Sociales
Instituto de Investigaciones Económicas
Buenos Aires 177
4400 Salta
Argentina

REUNIONES DE DISCUSION

<u>Nº</u>	<u>Fecha</u>	<u>Autor</u>	<u>Título</u>
70	28/10/92	Eduardo Antonelli	"Aspectos Microeconómicos de la Macroeconomía Keynesiana"
71	17/ 3/93	Eduardo Antonelli	"Una Modelización del Plan de Convertibilidad"
72	18/ 7/93	Eusebio Cleto del Rey y Carlos Luis Rojas	"Costos de los Tratamientos del Mal de Chagas"
73	18/ 9/93	Eduardo Antonelli	"La Demanda Agregada"
74	8/ 9/93	Juan Carlos Cid	"Clasificación de Areas Geográficas de la Ciudad de Salta en Base a Resultados Censales"
75	6/10/93	Carlos Luis Rojas	"Distintos Enfoques en la Evaluación Económica de Proyectos"
76	24/11/93	Eduardo Antonelli	"Equilibrio y Brechas Económicas"
77	18/12/93	Eusebio Cleto del Rey	"Cálculo del Stock de Salvados del Mal de Chagas"
78	16/ 2/94	Eduardo Antonelli	"La Política Económica en Salta en el Período 1976-1983"
79	13/ 4/94	Eusebio Cleto del Rey	"Mal de Chagas: Estimación de las Tasas de Infectación"