

MODELO ESTOCÁSTICO DE LA PROPAGACIÓN DE LA ATAXIA ESPINOCEREBELOSA TIPO 2

Karen García*,¹; Vivian Sistachs, **; Aymee Marrero, **;

Gonzalo Joya*** & Luis Enrique Almager****

*Instituto de Neurología y Neurocirugía "Dr. Rafael Estrada Gonzáles", Cuba

**Universidad de La Habana, Cuba

***Universidad de Málaga, España

****Clínica para la Investigación y Rehabilitación de las Ataxias Hereditarias, Cuba.

ABSTRACT

The spinocerebellar ataxia type 2 (SCA-2) is a neurodegenerative disease which is characterized by the progressive destruction of the nervous cells of the spinal marrow and of the cerebellum, presenting an autosomal and dominant genetic transmission. No consulting investigation has been carried out to model the ataxia progressive evolution in Cuba, where more of the ataxia cases have been found, therefore it would be interesting to estimate the growing parameter r of the disease and analyze their behavior for predict his evolutionary future.

KEY WORDS: Branching processes, BGW, spinocerebellar ataxia type 2

RESUMEN

La ataxia espinocerebelosa tipo 2 (SCA-2) es una afección neurodegenerativa que se caracteriza por una destrucción progresiva de las células nerviosas de la médula espinal y del cerebelo, además presenta un patrón de herencia autosómico dominante. Dado que ningún estudio consultado sobre la ataxia hereditaria SCA-2, modela la evolución epidemiológica de dicha enfermedad en Cuba y por ser este el país con mayor cantidad de casos reportados, resulta de interés simular la evolución de la ataxia hereditaria SCA-2 en Cuba, estimar su parámetro de crecimiento y analizar su comportamiento a través de generaciones para predecir su futuro evolutivo.

PALABRAS CLAVE: Procesos de ramificación, BGW, ataxia espinocerebelosa tipo 2

1. INTRODUCCIÓN

La ataxia espinocerebelosa tipo 2 (SCA-2) pertenece a un grupo de ocho afecciones neurodegenerativas que se caracterizan por presentar un patrón de herencia autosómico dominante, lo cual significa que puede afectar con igual probabilidad tanto a hombres como a mujeres [1]. Esta enfermedad conlleva a una destrucción progresiva de las células nerviosas de la médula espinal y del cerebelo trayendo consigo trastornos de la coordinación, la marcha, el habla y los movimientos oculares, confinando al paciente a estar en una cama o en una silla de ruedas (minusvalía). Los síntomas a menudo empiezan entre los 20 y los 49 años, precisamente en la etapa de mayor desarrollo social, profesional y personal del individuo, de ahí que estos enfermos se conviertan pronto en una carga para la familia y la sociedad.

A nivel mundial, Cuba cuenta con la mayor cantidad de casos reportados con la afección SCA-2 con 575 enfermos que pertenecen a 120 familias. Notemos que en Cuba la prevalencia general para las ataxias hereditarias es de 7.18 casos por 100 000 habitantes, el segundo país con mayor prevalencia es Islandia con 2.7 casos por 100 000 habitantes, seguido de 2.1 casos por 100 000 habitantes en Suecia y 1.1 casos por 100 000 habitantes en Suiza [1].

Sobre la ataxia SCA-2 se han realizado estudios descriptivos en donde se analizan las características epidemiológicas de la enfermedad [1], así como la relación existente entre la edad de inicio de la enfermedad y el tamaño de la cadena expandida. Cabe destacar, en este último aspecto, el trabajo del Dr. Luis Enrique Almaguer Mederos en el cual se estima el riesgo de desarrollar la enfermedad SCA-2 a una edad en particular como función del número de repeticiones del trinucleótido CAG en el primer exón del gen [2].

¹ kgmmath@gmail.com

Sin embargo, no se tienen referencias sobre investigaciones que modelen la evolución epidemiológica de la enfermedad. Por eso, con este estudio se desea simular la evolución de la ataxia hereditaria SCA-2 en Cuba. Para ello se basa nuestro modelo en un proceso de ramificación del tipo Bienaymé-Galton-Watson, con el objetivo de estimar el parámetro de crecimiento r de la enfermedad, así como analizar su comportamiento a través de generaciones para finalmente predecir su futuro evolutivo.

2. DESCRIPCIÓN DEL MODELO

2.1 Descripción del modelo de Bienaymé- Galton-Watson:

El proceso de Bienaymé-Galton-Watson (BGW) es una cadena de Markov $(X_n)_{n=0,1,\dots}$ donde X_n es el número de individuos presentes en la n -ésima generación. Su escala de tiempo está dada por las generaciones, no se consideran las duraciones de vida de los individuos. Visto de otra forma cada individuo vive una unidad de tiempo [3]. Los individuos se reproducen independientemente siguiendo una ley p_j es decir la probabilidad de producir j individuos [4]. La función generadora de esta ley es:

$$f(s) = \sum_{j=1}^{\infty} p_j s^j \quad (1)$$

El valor medio de la ley de reproducción se define como:

$$m = \sum_{j=1}^{\infty} j p_j = f'(1-) \quad (2)$$

que representa el número medio de individuos generados en una reproducción.

La probabilidad de extinción depende en gran medida de este valor esperado m de la ley de reproducción. Esto originó la clasificación de los BGW en los siguientes tres tipos: supercrítico ($m > 1$), crítico ($m = 1$) y subcrítico ($m < 1$) [4].

En el caso subcrítico y crítico el proceso se extingue casi seguramente para algún valor de n , es decir

$$P(X(n) = 0 | X(0) = i) = 1 \quad (3)$$

En el caso supercrítico la población sobrevive con probabilidad positiva y

$$P(X(\infty) = +\infty | X(0) = i) = 1 - (q^i) \quad (4)$$

donde q es la única raíz en $]0,1[$ de la ecuación $f(s) = s$.

2.2. Descripción del modelo epidemiológico para la ataxia SCA-2

El modelo epidemiológico que se muestra es un proceso Markoviano conocido como proceso de Bienaymé-Galton -Watson (BGW), que comienza con n individuos que introducen el gen enfermo dentro de sus n respectivas familias en g años anteriores. Este modelo puede ser empleado para cualquier enfermedad genética con un patrón de herencia autosómico dominante [Ver Figura 1].

Se asume que todos los individuos enfermos pueden reproducirse y de esta manera procrear nuevos individuos enfermos. Este proceso tiene una distribución Poisson con parámetro r , siendo r el número medio de hijos enfermos por persona enferma. Se tiene en cuenta que si una persona no tuvo hijos, entonces no tuvo hijos enfermos.

Las variables en este modelo son:

- t : Número de generación.
- n_t : Número de familias con el gen en la generación t .
- $X_i(t)$: Número de personas enfermas de la familia i en la generación t .
- $Y_{i,j}(t)$: Número de descendientes enfermos del individuo j de la familia i en la generación t . Distribuye Poisson de parámetro r . ($Y_{i,j}(t) \sim \text{Poisson}(r)$)
- $X(t+1)$: Número total de individuos con el gen en la generación $t+1$:

$$X(t+1) = \sum_{i=1}^{n_t} \sum_{j=1}^{X_i(t)} Y_{i,j}(t) \quad (5)$$

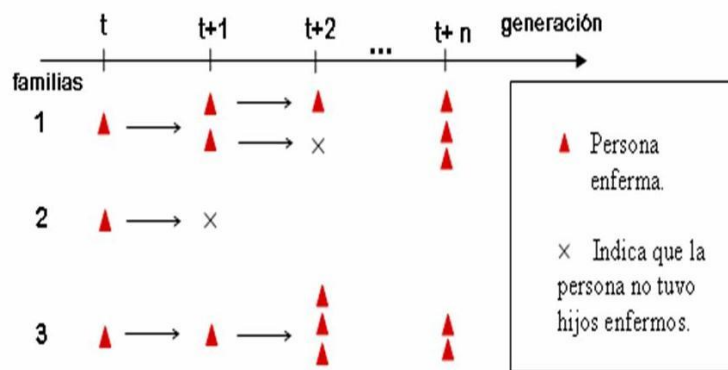


Figura1 Modelo de la ataxia SCA-2

Para implementar el modelo se creo un programa en Matlab que llamamos ASIM cuyos detalles aparecen en el anexo 1. Se trabajó en Matlab 7.11 ya que es una herramienta de alto nivel para desarrollar aplicaciones técnicas y una de sus capacidades más atractivas es la de realizar una amplia variedad de gráficos. El programa confeccionado para esta investigación (ASIM) permite graficar el comportamiento de la cantidad de nacidos con la mutación causante de la ataxia SCA-2 a través de generaciones. Además permite modificar el valor del parámetro r , así como la cantidad de generaciones que se desean simular. Para un mayor desenvolvimiento de las simulaciones en las siguiente sección se estima un intervalo de confianza sobre r .

3. RESULTADOS

3.1. Estimación del parámetro de la distribución Poisson

A partir de los datos ofrecidos por el Centro para la Investigación y Rehabilitación de las Ataxias Hereditarias cito en Holguín se tiene el número de hijos por paciente con ataxia espinocerebelosa tipo 2 y su probabilidad en una muestra de 200 individuos [Ver Tabla 1]. Se quiere buscar un intervalo de confianza de un 95% sobre r .

No. de hijos (k)								
Probabilidad (p_k)	.05	.195	.385	.205	.095	.035	.03	.005

Tabla 1 Probabilidad del número de hijos en pacientes con SCA-2.

Sea la variable aleatoria:

$$Z = \frac{\sum_{i=1}^n X_i - nr}{\sqrt{nr}} \quad (6)$$

donde $n = 200$ y $X_i, i \in [1, 200]$ es una muestra de variables aleatorias con distribución Poisson y media desconocida r . De acuerdo con el teorema central del límite, la variable aleatoria Z distribuye normal con esperanza 0 y varianza 1.

Por lo tanto es posible hallar los valores cuantiles de $z_{\alpha/2}$ y $z_{1-\alpha/2}$ [5], para la cual

$$P\left(\frac{z_{\alpha}}{2} < Z < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha \quad (7)$$

Se sustituye y despeja \sqrt{r}

$$P\left(-\frac{\sqrt{\sum X_i + \left(\frac{z_{1-\frac{\alpha}{2}}}{2}\right)^2} - \frac{z_{1-\frac{\alpha}{2}}}{2}}{\sqrt{n}} < \sqrt{r} < \frac{\sqrt{\sum X_i + \left(\frac{z_{\frac{\alpha}{2}}}{2}\right)^2} - \frac{z_{\frac{\alpha}{2}}}{2}}{\sqrt{n}}\right) = 1 - \alpha \quad (8)$$

Entonces para un coeficiente de confianza del 95% se tiene que $\alpha = 0.05$ y $z_{1-\frac{\alpha}{2}} = 1.96$. Con base a los datos muestrales el valor de $\sum_{i=1}^{200} X_i = 970$. Entonces el intervalo de confianza al 95% para el parámetro r de la distribución Poisson es:

$$I = [2.14, 2.57] \quad (9)$$

Se conoce que un buen estimador puntual del parámetro r de la distribución Poisson es la media, pues esta satisface una serie de propiedades deseables: su esperanza es igual a la del parámetro que se está estimando, es cada vez mejor conforme crece el tamaño de la muestra y es un estimador insesgado cuya varianza se apega a la cota inferior de Cramer-Rao [5]. Además la media es una estadística suficiente para el parámetro de la distribución Poisson [Ver Figura 2], nótese que la media $\bar{x}=2.35$ pertenece al intervalo I .

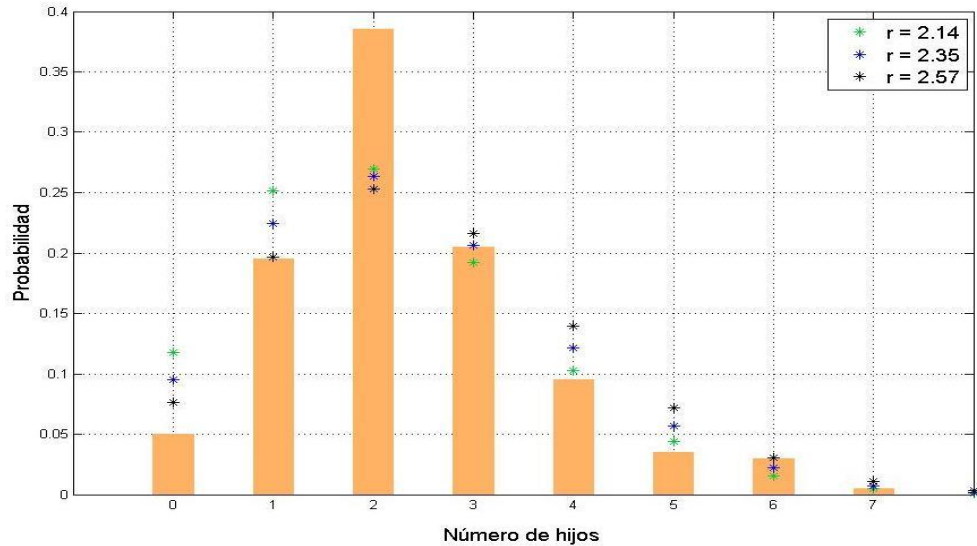


Figura 2 Probabilidad del número de hijos por persona con la mutación causante de la ataxia espinocerebelosa tipo 2. (*) Representa la distribución Poisson de parámetro r .

3.2. Validación del modelo

Anteriormente se supuso que la ley de reproducción del modelo epidemiológico se comporta según una distribución Poisson con parámetro r , siendo r el número medio de hijos enfermos por persona enferma. Además se halló para el parámetro de la distribución Poisson un intervalo de confianza a un 95 %. Ahora se comprobará la validez de tales suposiciones utilizando una prueba Ji-cuadrado. Comprobaremos la hipótesis de que $H: r = r_0$. El procedimiento consiste en hacer comparaciones entre el número real de observaciones y el número esperado de las mismas (esperados bajo la suposición). La estadística Ji-cuadrado se calculará de acuerdo con la ecuación:

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} \quad (10)$$

donde o_i representa el i -ésimo número observado y e_i representa el i -ésimo número esperado. Los grados de libertad serán $k - p - 1$, donde k representa el número de observaciones y p la cantidad de parámetros estimados para la muestra estadística. Entonces para el nivel de significación escogido α , si $\chi^2 \geq \chi_{1-\alpha}^2(k - p - 1)$ la hipótesis H será rechazada.

En nuestro modelo se conoce que el parámetro r de la distribución Poisson se encuentra en el intervalo $[2.14, 2.57]$, por lo tanto para cada valor fijo de r se tendrá un valor de la estadística Ji-cuadrado. Observemos la Figura 3 en donde se tiene el comportamiento de la estadística Ji-cuadrado conforme varía el parámetro r de la distribución Poisson:

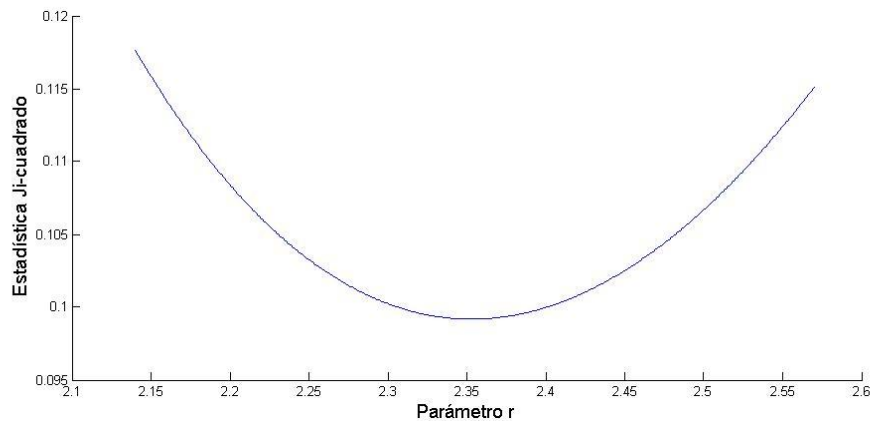


Figura 3 Comportamiento de la estadística Ji-cuadrado conforme varía el parámetro r de la distribución Poisson.

Fijando $\alpha = 0.001$ encontramos que los valores tomados por la estadística Ji-cuadrado son todos menores que $\chi_{0.001}^2(6) = 0.872$ la hipótesis de que los datos se comportan según una distribución Poisson con parámetro r no puede ser rechazada.

4. FUTURO EVOLUTIVO DE LA ENFERMEDAD

4.1. Simulación de la evolución epidemiológica de la enfermedad

Con el objetivo de analizar el comportamiento de la cantidad de nacidos con la mutación causante de la ataxia SCA-2, se elaboró el programa ASIM en Matlab 7.11[ver Anexo 1], el cual utiliza el modelo estocástico considerado anteriormente. Se realizaron 100 simulaciones en 5 generaciones para los siguientes valores del parámetro r : 2.14; 2.35; 2.46 y 2.57, que se corresponden a los extremos del intervalo, el valor del estimador media y un valor al azar del intervalo. Como se observa en la Figura 4 la cantidad de nacidos con la mutación causante de la ataxia SCA-2 se incrementa conforme aumenta la cantidad de generaciones.

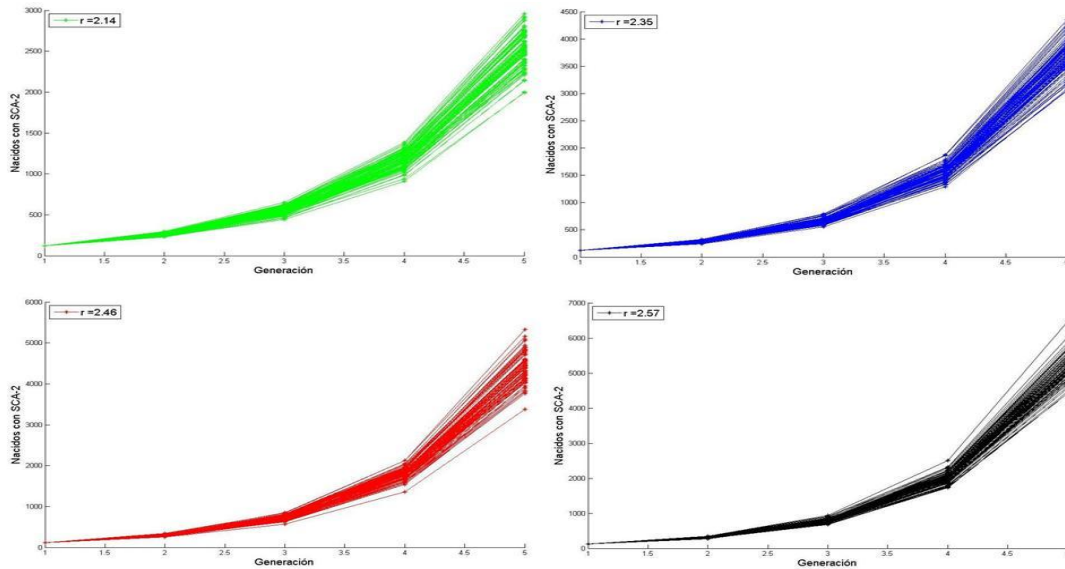


Figura 4 Cantidad de nacidos con la mutación causante de la ataxia SCA-2 en 5 generaciones para diferentes valores del parámetro r .

4.2. Probabilidad con que la población de enfermos sobreviva

En el modelo mostrado, la reproducción sigue una ley Poisson con parámetro $r \in [2.14, 2.57]$, cuya función generadora f de esta ley según (1) es $f(x) = e^{-r(1-s)}$ [6]. Entonces por (2) $m = f'(1) = r$, como $m > 1$ estamos en el caso supercrítico. De aquí que la población sobreviva con probabilidad positiva y se cumple (4), es decir

$$P(X(\infty) = +\infty | X(0) = i) = 1 - (q^i)$$

donde q es la única raíz en $]0,1[$ de $f(s) = s$.

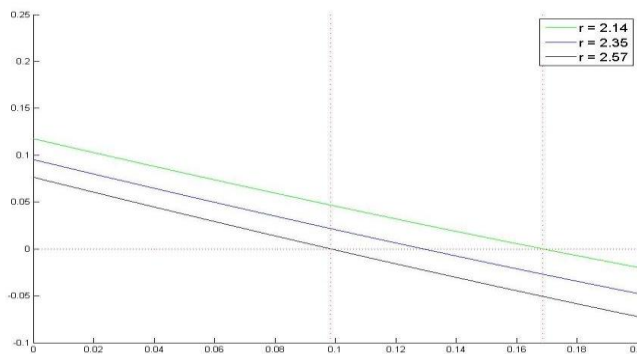


Figura 5 Gráfico de la función $e^{-r(1-s)}=s$ para diferentes valores del parámetro r . Las líneas discontinuas indican la banda en donde varía el valor de la raíz q .

El parámetro r varía entre 2.14 y 2.57 entonces la raíz q en $]0, 1[$ de $e^{-r(1-s)} = s$, para cada valor fijo de r se encontrará en la banda comprendida entre los números 0,0986 y 0,1688 [Ver Figura 5].

Existen 120 familias con la mutación causante de la ataxia espinocerebelosa tipo 2, entonces en la generación 0 hay 120 personas con el gen, es decir, $X(0) = 120$. De aquí que la población de personas con el gen SCA-2 sobreviva casi seguramente con probabilidad 1.

5. CONCLUSIONES

En el presente trabajo se desarrolló un modelo que simula la evolución de la ataxia espinocerebelosa SCA-2 basándose en un proceso BGW. Este modelo sigue una ley de ramificación que distribuye Poisson con parámetro r (número medio de hijos enfermos por persona enferma). Según la bibliografía consultada hasta el momento, es la primera vez que se utiliza un modelo de este tipo para analizar el comportamiento evolutivo de la enfermedad en Cuba. Este puede aplicarse, además, en el estudio de cualquier enfermedad con patrón de herencia autosómico dominante.

Para desarrollar el modelo se halló el parámetro de crecimiento r de la ataxia SCA-2, el cual está entre 2.14 y 2.57. Al utilizar el valor del parámetro r y las propiedades de los BGW se demostró que la evolución epidemiológica crece progresivamente y, por lo tanto, la enfermedad no se extingue.

6. RECOMENDACIONES

La predicción se basa en el parámetro de crecimiento actual, no obstante, este parámetro puede variar en un futuro, teniendo en cuenta los cambios demográficos, sociales, económicos, etc, lo cual puede ocasionar cambios en la predicción del futuro evolutivo de la enfermedad. Puesto que la tendencia en estos días es que la natalidad disminuya progresivamente, habrá que estudiar si la enfermedad en un futuro lejano varía su comportamiento disminuyendo su presencia.

Agradecimientos: A la AECID (Agencia Española de Cooperación Internacional para el Desarrollo), Project D/023835/09 por el apoyo para la realización de este trabajo.

REFERENCIAS

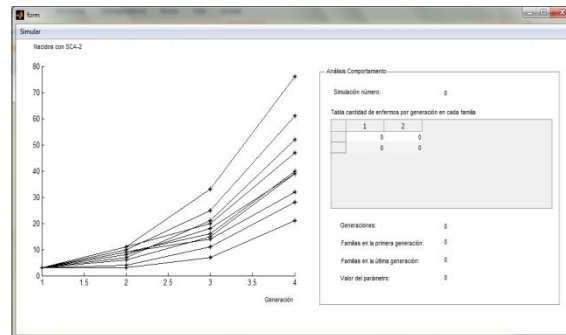
- [1] VELÁZQUEZ-PÉREZ, L., SANTOS, F. N., GARCÍA, R., PANEQUE, H. M. & HECHAVARRÍA, P. R., (2001): Epidemiología de la ataxia hereditaria cubana, **Rev Neurol**; 32, 606-611.
- [2] ALMAGUER-MEDEROSA, L. E., FALCÓN, N. S., ALMIRA, Y. R., ZALDIVAR, Y. G., ALMARALES, D. C., GÓNGORA, E. M., HERRERA, M. P., BATALLÁN, K. E., ARMIÑÁN, R. R., MANRESA, M. V., CRUZ, G. S., LAFFITA-MESA, J., CYUS, T. M., CHANG, V., AUBURGER, G., GISPERT, S., & PÉREZ, L. V., (2010): Estimation of the age at onset in spinocerebellar ataxia type 2 Cuban patients by survival analysis, **Rev ClinGenet**; 78, 169-174.
- [3] VATUTIN, V. A., (2005): **Branching processes and their applications**, Notas. <http://ismi.math.uni-frankfurt.de/vatutin/BranchingProcIndex.shtml>.
- [4] ATHREYA, K. B. & NEY, P. E., (1972): **Branching processes**, Springer, New-York.
- [5] CANAVOS, G., (1988): **Probabilidad y estadística: Aplicaciones y métodos**, McGraw Hill, México.
- [6] PENISSON, S., (2010): Conditional limit theorem for multitype branching processes and illustration in epidemiological risk analysis, **Tesis doctoral Universidad de Postdam**. <http://opus.kobv.de/ubp/volltexte/2010/4530/>

ANEXO 1

El programa ASIM (Ataxia Simulation) fue implementado en Matlab 7.11 y tiene como objetivos:

- Graficar la cantidad de enfermos por generación en tantas simulaciones como desee el usuario.
- Analizar por separado el comportamiento de cada una de estas simulaciones.

Parámetros de inicio y graficar: Estos se introducen al dar clic en la pestaña graficar en el menú simulación. El usuario introduce la cantidad de generaciones que se desean simular, la cantidad de familias con el gen en la primera generación, el número de simulaciones que se desea desarrollar y el valor del parámetro.



Análisis del comportamiento de una de las simulaciones: Una vez introducidos los datos el programa permite analizar el comportamiento de cualquiera de las simulaciones seleccionadas por el usuario. Esta se coloreará de rojo y en el lado derecho de la ventana se mostraran valores relacionados con la misma.

