

CAPÍTULO V

RESULTADOS, CONCLUSIONES Y RECOMENDACIONES

Para realizar la optimización de la Función de Distribución Hockey Stick se siguió una metodología, la cual se presenta en este Capítulo. Esta metodología cuenta con 8 pasos los cuales se describirán aquí. También se presentarán los resultados obtenidos de la simulación y las líneas de investigación que se quedarán abiertas para futuras investigaciones.

5.1 Metodología general

Los pasos que se siguieron para la elaboración de esta tesis se enuncian en este apartado. El desarrollo y explicación de los puntos siguientes se pueden encontrar en Capítulos anteriores.

1. El primer paso es desarrollar la Función de Distribución Acumulativa de las fallas y esto se realizó con la suposición de que las fallas aleatorias tienen una distribución Exponencial, mientras que la tasa de fallas tardías cuentan con una distribución Weibull. Las siguientes ecuaciones obtenidas son: Ver Capítulo III.

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & 0 < t \leq t_a \\ 1 - e^{-\left[\lambda t_a \left(\frac{m-1}{m}\right) + \frac{t^m}{c^m}\right]} & t > t_a \\ 0 & d.o.m. \end{cases}$$

2. El segundo paso es derivar la Función de Distribución Acumulativa y obtener la Función de Densidad para la distribución.

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & 0 < t \leq t_a \\ \frac{mt^{m-1}}{c^m} e^{-\left[\lambda t_a \left(\frac{m-1}{m}\right) + \left(\frac{t}{c}\right)^m\right]} & t > t_a \\ 0 & d.o.m. \end{cases}$$

3. Como tercer paso se generó la ecuación de la Función de Máxima Verosimilitud que posteriormente se programará para estimar los valores de los parámetros de la distribución Hockey Stick. Ver Capítulo III.

$$L = \prod_{\{t_i / 0 < t_i \leq t_a\}} \lambda e^{-\lambda t_i} \prod_{\{t_j / t_j > t_a\}} \frac{mt_j^{m-1}}{c^m} e^{-\left[\lambda t_a \left(\frac{m-1}{m}\right) + \frac{t_j^m}{c^m}\right]}$$

4. El cuarto paso de ésta metodología es aplicar la ecuación 1.2 mencionada en el Capítulo I, por manejabilidad de las operaciones. Esta ecuación será la que se utilice para realizar la optimización de la Función de Distribución Hockey Stick.

$$l = \sum_{\{t_i / 0 < t_i \leq t_a\}} (\ln \lambda - \lambda t_i) + \sum_{\{t_j / t_j > t_a\}} \left[\ln \left(\frac{mt_j^{m-1}}{c^m} \right) - \lambda t_a \left(\frac{m-1}{m} \right) - \frac{t_j^m}{c^m} \right]$$

5. El quinto paso de esta metodología es realizar un programa computacional con el objetivo de encontrar los valores de los estimadores de los parámetros de la distribución Hockey Stick (Apéndice A). El Fortran es el lenguaje utilizado por las ventajas ya mencionadas en el Capítulo IV.

El programa consta de un programa principal, en el cual se introducen los siguientes datos de entrada: un punto inicial el cual se encuentra dentro de la

región factible implícita en la matriz necesaria para llevar a cabo el método simplex y una muestra aleatoria. Este programa manda llamar a un archivo que contiene la muestra de los números aleatorios que se distribuyen Hockey Stick. La subrutina amoeba, es la encargada de llevar a cabo el algoritmo Nelder Mead que optimiza la Función de Máxima Verosimilitud que se obtuvo en el paso 4 de esta metodología.

Las variables y arreglos más representativos que fueron necesarios definir en el programa son los siguientes:

- **rdatos (1500):** vector entero en el cual se almacenan los valores de la muestra aleatoria. Este vector esta definido con una capacidad de 1500, pero no existe inconveniente en que éste sea alterado para realizar pruebas con un número mayor o menor de datos en caso de que fuera necesario.
 - **p(np,mp):** matriz entera que corresponde a los valores requeridos para llevar a cabo el método simplex. El punto inicial va implícito en este arreglo.
 - **np:** es una variable entera que corresponde al número de estimadores a calcular. En este caso es tres ya que el cuarto estimador es c que se calcula dentro del programa.
 - **mp:** np+1.
 - **ndatos:** variable entera que corresponde a los valores de la muestra. Estos valores son almacenados en el vector rdatos, anteriormente descrito.
 - **ta, rm, c:** variables reales pertenecientes a los estimadores de los parámetros de la distribución Weibull, utilizada para el periodo de fallas tardías.
 - **rlambda:** variable real perteneciente al estimador del parámetro de la distribución Exponencial, utilizada para el periodo de fallas aleatorias.
6. En el sexto paso es necesario generar una muestra que siga la distribución Hockey Stick y para este propósito se realizó un programa en Excel utilizando el Teorema de la Transformación Inversa que se describió en el Capítulo IV.

6.1 Como primer paso fue establecer los valores de los parámetros siguientes:

$$\lambda = 0.01$$

$$t_a = 100$$

$$m = 2$$

6.2 Con los valores anteriores se calculó el valor de c con la siguiente igualdad:

$$c = \sqrt[m]{\frac{mt_a^{m-1}}{\lambda}}$$

Obteniendo el valor de:

$$c = 141.42$$

6.3 Se prosiguió a calcular la Función Acumulativa de Falla en el punto t_a por medio de la siguiente ecuación:

$$F(x) = 1 - e^{-\lambda t} = 0.632120$$

6.4 A continuación se generaron 1500 números aleatorios $\sim U(0, 1)$. Para este propósito se utilizó Excel y la función utilizada para este propósito fue: ALEATORIO (). En la Tabla 5.1 se muestran algunos de los números aleatorios $\sim U(0, 1)$ obtenidos.

Num. De observación	Num. $\sim U(0,1)$
1	0.71744
2	0.64694
3	0.81596
4	0.75456

5	0.77097
6	0.88922
\vdots	\vdots
1497	0.05806
1498	0.45086
1499	0.11628
1500	0.18486

Tabla 5.1 Números aleatorios $X \sim U(0, 1)$.

6.5 Las siguientes ecuaciones que fueron desarrolladas en el Capítulo IV se programan respecto a los valores de los números aleatorios en relación a FDA que tiene el valor de 0.63 y se controlaron con la función IF.

$$t = \begin{cases} -\frac{1}{\lambda} \ln(1-y) & 0 < t \leq t_a \\ c \left[-\lambda t_a \left(\frac{m-1}{m} \right) - \ln(1-y) \right]^{\frac{1}{m}} & t > t_a \\ 0 & d.o.m. \end{cases}$$

Se obtuvieron los valores que se muestran en la Tabla 5.2 con una distribución Hockey Stick.

Observación	Aleatorio	t ~ Hockey Stick
1	0.71744	123.6004
2	0.64694	104.0297
3	0.81596	154.4396
4	0.75456	134.5137

5	0.77097	139.5618
6	0.88922	184.4025
⋮	⋮	⋮
1497	0.05806	5.9817
1498	0.45086	59.9404
1499	0.11628	12.3614
1500	0.18486	20.4395

Tabla 5.2 Números aleatorios $t \sim \text{Hockey Stick}$.

7. En este séptimo paso teniendo ya calculados los valores de los parámetros de la distribución, es aquí cuando el método estadístico de Bootstrap aparece para calcular los estimadores de las muestras tomadas de la muestra aleatoria para que estos parámetros sean los mejores estimados.
8. En el octavo paso ya que se tienen los valores calculados de los estimadores se realiza un intervalo de confianza de los estimadores al 95 % usando la siguiente fórmula:

$$IC = \hat{\theta} \pm 1.96\sigma$$

Donde $\hat{\theta}$ es el valor de la estimación de los parámetros de la distribución.

σ es la desviación estándar del parámetro $\hat{\theta}$ y se calcula de la siguiente forma:

$$\sigma^2 = E(x^2) - [E(x)]^2$$

5.2 Resultados

Se realizaron corridas del programa para distintos tamaños de muestra: $n = 100$, $n = 200$, $n = 500$, $n = 1000$, $n = 1500$. En la Tabla 5.3 se muestra la estimación de los parámetros obtenidos con el tamaño de muestra correspondiente.

Parámetros	Estimadores				
	Tamaño de muestra (n)				
	n = 100	n = 200	n = 500	n = 1000	n = 1500
ta = 100	81.8117	80.753	95.9045	96.1365	99.4633
$\lambda = 0.01$	0.007853	0.009565	0.01001	0.009548	0.009539
m = 2	2.5547	1.9409	1.906	1.9491	1.9851
c = 141.42	140.4345	129.8189	137.3936	141.4737	144.2715
Función =	-541.505	-1083.16	-2720.027	-5465.665	-8205.945

Tabla 5.3 Resultados para diferentes tamaños de muestra.

Es posible observar en la Tabla 5.3 como se acercan los valores estimados a los valores de los parámetros. Esto se debe a que entre mayor cantidad de datos exista en la muestra es mejor la estimación de los parámetros. Para ver diferencia de los valores estimados con respecto a los parámetros se corrió el programa para 5 tamaños de muestra.

La Tabla 5.4 muestra los intervalos de confianza al igual que la Tabla 5.3 con los mismos cinco tamaños de muestra.

Parámetros	Estimadores					
	Tamaño de muestra (n)					
	n = 100		n = 200		n = 500	
	LI	LS	LI	LS	LI	LS
ta = 100	45.7384	117.885	57.9597	103.5463	77.633	114.176
$\lambda = 0.01$	0.006701	0.009005	0.008678	0.01045	0.008859	0.01117
m = 2	1.009	4.1005	1.4604	2.4212	1.4339	2.3781
c = 141.42	108.4357	172.4333	112.6237	147.014	125.8212	148.966

a)

Parámetros	Estimadores			
	Tamaño de muestra (n)			
	n = 1000		n = 1500	
	LI	LS	LI	LS
ta = 100	91.4607	100.8125	96.2912	102.6355
$\lambda = 0.01$	0.00892	0.0112	0.009363	0.009714
m = 2	1.9207	1.9777	1.9271	2.043
c = 141.42	139.973	142.9744	141.9084	146.6346

b)

Tabla 5.4 Intervalos de confianza para diferentes tamaños de muestra.

a) n = 100, 200, 500

b) n = 1000, 1500

En la Tabla 5.5 se presentan las diferencias entre el intervalo superior y el intervalo inferior de los parámetros estimados. Se puede observar como va disminuyendo la diferencia conforme el tamaño de muestra es mayor. Por lo que se aconseja que se tome el mayor número de elementos en una muestra cuando se desee obtener alguna estimación para que ésta sea lo mas certera posible.

Parámetros	Estimadores				
	Tamaño de muestra (n)				
	n = 100	n = 200	n = 500	n = 1000	n = 1500
ta = 100	72.1466	45.5866	36.543	9.3518	6.3443
$\lambda = 0.01$	0.0023	0.00177	0.00231	0.00228	0.00035
m = 2	3.0915	0.9608	0.9442	0.057	0.1159
c = 141.42	63.9976	34.3903	23.1448	3.0014	4.7262

Tabla 5.5 Diferencias entre intervalos según tamaño de muestra.

Gráficamente se puede apreciar en el Apéndice B en gráficas de barras para cada uno de los parámetros que se estimaron en esta tesis. La diferencia para cada uno de los intervalos se ve como va disminuyendo conforme el tamaño de la muestra n va aumentando.

5.3 Apertura de nuevas líneas de investigación

El programa que se realizó para la elaboración de esta tesis no es capaz de proporcionar una solución inicial para comenzar la búsqueda de la mejor solución. Es por esta razón que se le tiene que suministrar un punto de inicio. Una de las aportaciones que se pueden hacer a esta tesis es desarrollar una metodología para llevar a cabo el procedimiento de generar una solución inicial. En el caso del programa elaborado se procedió a darle valores alrededor de los parámetros que fueron establecidos para la generación de los números aleatorios.

Otra interesante aportación que se puede originar para un nuevo trabajo de investigación es el desarrollo de una prueba de bondad de ajuste para la distribución Hockey Stick. Debido a la naturaleza de esta tesis, los números de la muestra aleatoria fueron generados por medio del Teorema de la Transformación Inversa, con lo que se garantiza que estos datos sigan una distribución Hockey Stick; sin embargo, cuando el muestreo se realiza directamente en un sistema en estudio al tomar los tiempos de falla, es sumamente importante verificar que las muestras sigan esta distribución, por que de lo contrario no será posible utilizar el programa desarrollado en esta investigación.

Además de la aplicación en temas de confiabilidad también es importante mencionar que esta distribución que no es solo para esta aplicación sino también se puede aplicar a la vida útil de cualquier material, el cual sufra desgaste con el paso del tiempo o por el uso. Por ejemplo en la vida de los conductores eléctricos, en la vida de cualquier maquinaria, y no solo en la industria ya que se puede aplicar en el riesgo de adquirir cáncer, en los efectos ambientales que se acentúan con el tiempo.