CAPÍTULO V

Optimización de la Función de Distribución Tina de Baño.

CAPITULO V

OPTIMIZACIÓN DE LA FUNCIÓN DE DISTRIBUCIÓN TINA DE BAÑO

5.1 Metodología general

Los pasos que se siguieron para la elaboración de esta tesis se enuncian a continuación. Cabe mencionar que el desarrollo y explicación de los puntos que enseguida se mencionarán se pueden encontrar en capítulos anteriores.

1. Se dedujo la Función Acumulativa de falla suponiendo que las fallas tempranas y tardías tienen una distribución Weibull, y las fallas aleatorias una Exponencial. Llegando a las siguientes ecuaciones. Ver Capítulo III, sección 3.1.

$$
F(x)\begin{cases}
=1-\exp\left(-\left(\dfrac{t}{c_a}\right)^{m_a}\right) & 0<t<t_a \\
=1-\exp\left(-\,l\left[t_a\left(\dfrac{1-m_a}{m_a}\right)+t\right]\right) & t_a<t<t_b \\
=1-\exp\left(-\left[l\left(t_a\left(\dfrac{1-m_a}{m_a}\right)+t_b\left(\dfrac{m_b-1}{m_b}\right)\right)+\left(\dfrac{t}{c_b}\right)^{m_b}\right]\right) & t_b<t
\end{cases}
$$

2. Se procedió a derivar las Funciones de Distribución Acumulativa para de esta manera obtener las funciones de densidad correspondientes:

$$f(t)\begin{cases} = \dfrac{m_a t^{m_a-1}}{ca^{m_a}} \exp\left(-\left[\left(\dfrac{t}{ca}\right)^{m_a}\right]\right) & 0 < t < t_a \\[2ex] = l \exp\left(-l\left[t_a\left(\dfrac{1-m_a}{m_a}\right)+t\right]\right) & t_a < t < t_b \\[2ex] = \dfrac{m_b t^{m_b-1}}{cb^{m_b}} \exp\left(-\left[l\left[t_a\left(\dfrac{1-m_a}{m_a}\right)+t_b\left(\dfrac{m_b-1}{m_b}\right)\right]+\left(\dfrac{t}{cb}\right)^{m_b}\right]\right) & t_b < t \end{cases}$$

3. Se generó la ecuación de la función de Máxima Verosimilitud que más adelante se programará para estimar los valores de los parámetros de la función Tina de Baño. Ver Capítulo III, sección 3.2.

$$L = \prod_{\{t_i/0\langle t_i\langle t_a\}} \frac{m_a t_i^{m_a-1}}{ca^{m_a}} \exp\left(-\left[\left(\frac{t_i}{ca}\right)^{m_a}\right]\right) \prod_{\{t_j/t_a\langle t_j\langle t_b\}} l \exp\left(-l\left[t_a\left(\frac{1-m_a}{m_a}\right)+t_j\right]\right)$$

$$\prod_{\{t_k/t_b\langle t_k\}} \frac{m_b t_k^{m_b-1}}{cb^{m_b}} \exp\left(-\left[l\left[t_a\left(\frac{1-m_a}{m_a}\right)+t_b\left(\frac{m_b-1}{m_b}\right)\right]+\left(\frac{t_k}{cb}\right)^{m_b}\right]\right)$$

4. Como en el Capítulo III, sección 3.2 se había explicado, por manejabilidad de la función, se aplicó la siguiente igualdad: $l = \ln(L)$. La ecuación 3.3 que a continuación se muestra fue la utilizada para realizar la optimización de la Función de Distribución Tina de Baño.

$$l = \sum_{\{t_i / 0 \langle t_i \langle t_a\}} \left[\ln \frac{m_a t_i^{m_a - 1}}{c_a^{m_a}} - \left[\left(\frac{t_i}{c_a} \right)^{m_a} \right] \right] + \sum_{\{t_j / t_a \langle t_j \langle t_b\}} \left[\ln\ l - l \left[t_a \left(\frac{1 - m_a}{m_a} \right) + t_j \right] \right] +$$
$$\sum_{\{t_k / t_b \langle t_k\}} \left[\ln \frac{m_b t_k^{m_b - 1}}{c_b^{m_b}} - \left[l \left[t_a \left(\frac{1 - m_a}{m_a} \right) + t_b \left(\frac{m_b - 1}{m_b} \right) \right] + \left(\frac{t_k}{c_b} \right)^{m_b} \right] \right]$$

5. Se procedió a realizar un programa computacional cuyo objetivo era encontrar los valores de los estimadores de los parámetros de la función Tina de Baño (Apéndice A). El lenguaje de programación que se utilizó fue Fortran, ver Capítulo IV, sección 4.3. Se realizó un programa principal, en el cuál se introdujeron los siguientes datos de entrada: un punto inicial que estuviera dentro de la región factible implícita en la matriz necesaria para llevar a cabo el método simplex, la muestra aleatoria. Este programa hace un llamado al archivo que contiene la muestra de los números aleatorios que se distribuyen Tina de Baño, Numeros.for; a la subrutina amoeba, proporcionada por la M.C. Rina Ojeda, que es la encargada de llevar a cabo el algoritmo Nelder Mead que nos permitió optimizar la Función de Máxima Verosimilitud anteriormente desarrollada, ver Capítulo IV sección 4.2. Las variables y arreglos más representativos que fueron necesarios definir en el programa principal, fueron los siguientes:

- rdatos (1200): Vector real en el cuál se almacenan los valores de la muestra aleatoria. Aunque definido con una capacidad de 1200, no existe inconveniente en alterar el tamaño del vector en caso de que fuera necesario hacer pruebas con un mayor número de observaciones.

- p(np,mp): Matriz entera correspondiente a los valores requeridos para llevar a cabo el método simplex. El punto inicial va implícito en este arreglo.

- np: Variable entera que corresponde al número de estimadores a calcular.

- mp: np + 1.

- ndatos: Variable entera que corresponde a los valores de la muestra. Estos valores son almacenados en el vector Rdatos, anteriormente descrito.

- rma, ta, ca : Variables reales correspondientes a los estimadores de los parámetros de la distribución Weibull utilizada para el período de fallas tempranas.

- rlamda: Variable reales correspondiente al estimador del parámetro de la distribución Exponencial utilizada para el período de fallas aleatorias.

- rmb, tb, cb : Variables reales correspondientes a los estimadores de los parámetros de la distribución Weibull utilizada para el período de fallas tardías.

6. Posteriormente, fue necesario generar números que siguieran la distribución “Tina de Baño”. Para este propósito se realizó un programa en Excel utilizando el Teorema de la Transformación Inversa. Ver Capítulo IV, sección 4.4.1.

6.1 Primeramente se establecieron los valores de siguientes parámetros:

m_a = 0.5

c_a = 10

t_a = 10

m_b = 2

t_b = 40

6.2 Con estos valores se calculo λ y c_b por medio de las siguientes ecuaciones:

$$\lambda = \frac{m_a t_a^{m_a - 1}}{c_a}$$

$$c_b = \frac{m_b t_b^{m_b - 1}}{\lambda}$$

Obteniéndose los siguientes resultados:

λ = 0.05

c_b = 40

6.3 Se procedió a calcular la Función Acumulativa de Falla en los puntos t_a y t_b por medio de las siguientes ecuaciones, desarrolladas con anterioridad:

$$F(t_a) = 1 - \exp\left(-\left(\frac{ta}{ca}\right)^{m_a}\right) = 0.632120559$$

y:

$$F(t_b) = 1 - \exp\left(-\left[I\left(t_a\left(\frac{1-m_a}{m_a}\right) + t_b\left(\frac{m_b-1}{m_b}\right)\right) + \left(\frac{tb}{cb}\right)^{m_b}\right]\right) = 0.917915001$$

6.4 Se generaron 1200 números aleatorios ~ U(0,1). Para este propósito se utilizó el programa Excel. La función utilizada para este propósito fue: RAND ( ). A continuación, en la tabla 5.1, se muestran algunos de los números aleatorios distribuidos uniformemente 0,1 obtenidos.

Tabla 5.1. Números aleatorios X~U(0,1)

Número de observación	No. Aleatorio X~U(0,1)
1	0.66
2	0.614588447
3	0.084462237
4	0.714960755
⋮	⋮
⋮	⋮
1200	0.840904317

Elaboración propia.

6.5 Se programaron las siguientes ecuaciones desarrolladas en el Capítulo IV, sección 4.4.1. La ecuación que se debía utilizar de acuerdo a las condiciones respecto a los valores de x en relación con $F(t_a)$ y $F(t_b)$ se controlaron con la función IF.

$$
t\begin{cases}
= \left(-\ln(1-x)\right)^{1/m_a} * c_a & 0 < x < F(t_a) \\[2ex]
= \dfrac{-\ln(1-x)}{l} - t_a\left(\dfrac{1-m_a}{m_a}\right) & F(t_a) < x < F(t_b) \\[2ex]
= \left[-\ln(1-x) - l\left[t_a\left(\dfrac{1-m_a}{m_a}\right) + t_b\left(\dfrac{m_b-1}{m_b}\right)\right]\right]^{1/m_b} * c_b & F(t_b) < x
\end{cases}
$$

Obteniéndose los valores que se muestran en la tabla 5.2 con una distribución Tina de Baño.

Tabla 5.2. Números aleatorios t~Tina de Baño

Número de observación	No. Aleatorio X~U(0,1)	No. Aleatorio t~Tina de Baño
1	0.66	11.57619323
2	0.614588447	9.090545956
⋮	⋮	⋮
1200	0.840904317	26.7649896

Elaboración propia.

5.2 Resultados

Se corrió el programa utilizando los siguientes tamaños de muestra: n=100, 200, 400, 800 y 1200. En la tabla 5.3 se muestran las estimaciones de los parámetros con el tamaño de muestra correspondiente.

Tabla 5.3 Estimadores calculados al correr el programa con n=100, 200, 400, 800 y 1200.

Parámetros		Estimadores				
		Tamaño muestral (n)				
		100	200	400	800	1200
m_a	**0.5**	5.78E-01	5.76E-01	5.61E-01	4.96E-01	5.11E-01
c_a	**10**	10.52183	8.659636	9.750351	9.965721	9.861924
t_a	**10**	29.86664	19.19946	11.68696	8.903655	11.23713
m_b	**2**	5.304743	4.58483	3.059899	2.660617	2.486932
t_b	**40**	48.41754	48.24279	46.94837	47.10644	42.19753
λ	**0.05**	3.54E-02	4.74E-02	5.31E-02	5.27E-02	4.86E-02
c_b	**40**	59.90548	56.13425	50.20029	48.35099	45.61061
FUNCIÓN	**-11170.43568**	-335.6219	-635.8582	-1293.609	-2517.405	-3824.848

Elaboración propia.

Se calcularon los errores correspondientes a cada tamaño de muestra. La manera en que fueron calculados los errores relativos fue:

A. Se obtuvo el valor absoluto correspondiente a la diferencia entre el estimador con cada tamaño de muestra y el parámetro.

B. El resultado del inciso A se dividió entre el valor del parámetro.

C. Finalmente se multiplico por cien para expresar el error relativo en valor porcentual.

En la tabla 5.4, que a continuación se muestra, se registraron los errores que se generaron al correr el programa con distintos tamaños de muestra.

Tabla 5.4. Errores con diferentes tamaños de muestra.

Parámetros	Errores (%)				
	Tamaño muestral (n)				
	100	200	400	800	1200
m_a	15.69	15.19	12.13	0.77	2.12
c_a	5.22	13.40	2.50	0.34	1.38
t_a	198.67	91.99	16.87	10.96	12.37
m_b	165.24	129.24	52.99	33.03	24.35
t_b	21.04	20.61	17.37	17.77	5.49
λ	29.17	5.10	6.20	5.39	2.86
c_b	49.76	40.34	25.50	20.88	14.03
FUNCION	97.00	94.31	88.42	77.46	65.76
Error Porcentual	81.17	54.09	20.37	12.57	9.14

Elaboración propia.

En el Apéndice B se anexa una gráfica en la que se puede apreciar la forma en que el error va disminuyendo a medida que se va incrementando el tamaño de muestra (n).

5.3 Apertura de nuevas líneas de investigación.

El programa realizado para la elaboración de esta tesis necesita que se le suministre un punto de inicio. Sería posible desarrollar una metodología para llevar a cabo este procedimiento. En este caso se procedió a darle valores alrededor de los parámetros que fueron establecidos para la generación de los números aleatorios. Una sugerencia del procedimiento a seguir es el siguiente, sin embargo, cabe mencionar que sería necesario realizar la justificación y pruebas del mismo.

1. Determinar el rango en los que se encuentran las observaciones de la muestra.

   Rango = Mayor tiempo de falla – Menor tiempo de falla.

2. Se establecen los siguientes valores:

$$m_a = 0.5$$

$$m_b = 2.0$$

3. Estimar λ a través de la siguiente ecuación, desarrollada en el Capítulo I, sección 1.4:

$$\lambda = \frac{n}{\sum_{i=1}^{n} t_i}$$

4. Determinar t_a y t_b como sigue:

$$t_a = 0.33 * \text{Rango}$$

$$t_b = 0.66 * \text{Rango}$$

5. Una vez que se cuentan con los estimadores m_a, m_b, t_a, t_b y λ; es posible calcular c_a y c_b aprovechando la igualdad de las funciones de riesgo de falla en los tiempos críticos t_a y t_b. De esta manera se llega a las siguientes ecuaciones, previamente explicadas:

$$c_a = \frac{m_a t_a^{m_a - 1}}{\lambda} \qquad c_b = \frac{m_b t_b^{m_b - 1}}{\lambda}$$

Otro interesante tópico que podría originar un nuevo trabajo de investigación es el desarrollo de una prueba de Bondad y Ajuste para la distribución Tina de Baño. 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 garantizó que seguían una distribución Tina de Baño; sin embargo, cuando se realiza el muestreo de los tiempos de falla directamente del sistema en estudio, es de suma importancia verificar que las muestras sigan esta distribución, ya que, de lo contrario, no es posible utilizar el programa desarrollado en esta investigación.