

**Herrera, Luis A. ; Salvatierra, Lucas M. ; Razzitte, Adrián C. ; Kovalevski, Laura I. ; Dammig Quiña, Pablo L. ; Irurzun, Isabel M. ; Mola, Eduardo E.**

*Análisis acerca de la distribución de Weibull con datos truncados y la distribución generalizada de Pareto aplicado a fallas en dieléctricos*

Energeia, Año 10, N° 10, 2012

Este documento está disponible en la Biblioteca Digital de la Universidad Católica Argentina, repositorio institucional desarrollado por la Biblioteca Central "San Benito Abad". Su objetivo es difundir y preservar la producción intelectual de la Institución.

La Biblioteca posee la autorización del autor para su divulgación en línea.

Cómo citar el documento:

Herrera, L. A. et al. Análisis acerca de la distribución de Weibull con datos truncados y la distribución generalizada de Pareto aplicado a fallas en dieléctricos [en línea]. *Energeia*, 10(10), 2012.

Disponible en: <http://bibliotecadigital.uca.edu.ar/repositorio/revistas/analisis-distribucion-weibull-datos.pdf> [Fecha de consulta: .....]

## Análisis acerca de la distribución de Weibull con datos truncados y la distribución generalizada de Pareto aplicado a fallas en dieléctricos

Luis A. Herrera <sup>1</sup>, Lucas M. Salvatierra <sup>1,2</sup>, Adrian C. Razzitte <sup>3</sup>, Laura I. Kovalevski <sup>1,2</sup>, Pablo L. Dammig Quiña <sup>2</sup>, Isabel M. Irurzun <sup>2</sup> y Eduardo E. Mola <sup>1,2 \*</sup>.

(1) Facultad de Química e Ingeniería, Pontificia Universidad Católica Argentina, Pellegrini 3314, Rosario, Santa Fe, Argentina.

(2) Grupo de Sistemas Complejos, Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas (INIFTA), Universidad Nacional de La Plata, CCT-CONICET La Plata, Diag. 113 y 64, La Plata, Buenos Aires, Argentina.

(3) Departamento de Química, Facultad de Ingeniería, Universidad de Buenos Aires.

(\*) Autor para correspondencia: eemola@inifta.unlp.edu.ar

**Resumen.** En trabajos anteriores [1-4] se ha presentado un enfoque de los modelos de ruptura dieléctrica desde la Termodinámica Estadística. Se estudió la relación intrínseca entre estas distintas familias de árboles mediante una distribución generalizada de Weibull con un índice entrópico  $q$ , llegándose a la conclusión que este índice  $q$  es representativo de la estructura fractal de los distintos tipos de árboles eléctricos. Para definir la denominada distribución generalizada de Weibull y este índice  $q$ , se recurrió a un formalismo propio de la Termodinámica No Extensiva desarrollado por Tsallis [5]. En este trabajo retomamos este formalismo y desarrollamos un estudio riguroso del mismo para alcanzar una mayor comprensión de la denominada "entropía de Tsallis", finalmente también exploramos la existencia de este índice entrópico  $q$  en los modelos de falla en dieléctricos que son representados por una distribución de Weibull con datos truncados y establecemos una importante equivalencia con la Distribución Generalizada de Pareto.

**PALABRAS CLAVE:** Falla Dieléctrica, Tiempo de Ruptura, Distribución de Weibull truncada, Distribución generalizada de Pareto.

### 1.- Las Distribuciones Estadísticas y el Principio de Máxima Entropía:

Algunas de las distribuciones estadísticas conocidas como la de Weibull, se pueden obtener mediante la aplicación del Principio de Máxima Entropía (PME) bajo ciertas restricciones apropiadas junto a la condición de normalización de las probabilidades. En esta formulación, el funcional de entropía empleado es el denominado Boltzmann-Gibbs-Shannon (BGS) corrientemente conocido como funcional de entropía termodinámica. No obstante, han aparecido algunas alternativas como la propuesta por Tsallis. Veremos cómo utilizando el PME podemos hallar algunas distribuciones clásicas de probabilidades, en nuestro caso, la distribución de Weibull.

Según el PME, al problema planteado para encontrar la distribución de Weibull lo consideramos como un problema de optimización con extremos condicionados, el cual es soluble mediante los multiplicadores de Lagrange. En cada caso específico,  $x$  es una variable aleatoria perteneciente a un determinado conjunto  $\Omega$  que pertenece a los números reales, que representa en este caso el "estado" del sistema y  $p(x)$  es su densidad de probabilidad de forma tal que  $p(x)dx$  es la probabilidad de que el sistema se encuentre entre los estados  $x$  y  $x+dx$ , finalmente se trata de maximizar el siguiente funcional:

$$\Phi[p(x)] = -\int_{\Omega} p(x) \ln p(x) dx - \lambda_0 \left[ \int_{\Omega} p(x) dx - 1 \right] - \sum_{k=1}^N \lambda_k \left[ \int_{\Omega} g_k(x) p(x) dx - G_k \right] \quad (1)$$

Donde, los  $\lambda_k$  son los multiplicadores de Lagrange y  $g_k(x)$  son las  $N$  restricciones (información conocida sobre el sistema) y los  $G_k$  son los valores obtenidos sobre las  $g_k$  a través de la distribución:

$$G_k = \int_{\Omega} g_k(x) p(x) dx \quad (2)$$

En el caso particular de la distribución de Weibull, debemos considerar las siguientes restricciones y tener en cuenta además que contamos con  $\Omega = (0, \infty)$  :

$$\int_0^{\infty} p(x)dx = 1; \quad \int_0^{\infty} x^{\alpha} p(x)dx = g_1; \quad \int_0^{\infty} \ln(x) p(x)dx = g_2; \quad 0 < \alpha < 1 \quad (3)$$

Específicamente, en este caso los multiplicadores de Lagrange son los siguientes:

$$\lambda_1 = \lambda^{-\alpha} ; \quad \lambda_2 = 1 - \alpha$$

por lo que finalmente obtenemos:

$$p(x) = \frac{\alpha}{\lambda^{\alpha}} e^{-\left(\frac{x}{\lambda}\right)^{\alpha}} x^{\alpha-1} \quad (4)$$

Ahora bien, la aplicación del principio de máxima entropía (PME) se puede modificar sustituyendo el funcional BGS por el propuesto por Tsallis. Así, cambiando la forma de plantear las restricciones podemos obtener las denominadas distribuciones generalizadas, que en nuestro caso particular, es la distribución generalizada de Weibull, para hallar ahora esta distribución debemos maximizar el siguiente funcional:

$$\Phi_q[p(x)] = \frac{1}{q-1} \left[ 1 - \int_{\Omega} p^q(x) dx \right] - \lambda_o \left[ \int_{\Omega} p(x) dx - 1 \right] - \sum_{k=1}^N \lambda_k \left[ \frac{\int_{\Omega} g_k(x) p^q(x) dx}{\int_{\Omega} p^q(x) dx} - \langle G_k \rangle_q \right] \quad (5)$$

donde la forma de calcular las restricciones se obtiene mediante:

$$\langle G_k \rangle_q = \int_{\Omega} g_k(x) P_{esc}(x) dx \quad (6)$$

por cuestiones de simplificación de la notación, definimos una nueva probabilidad denominada probabilidad *escort* y que la enunciamos de la siguiente forma:

$$P_{esc}(x) = \frac{p^q(x)}{\int_{\Omega} p^q(x) dx} \quad (7)$$

A partir de estas definiciones, reinterpretemos la fórmula de la distribución de Weibull, llegando a establecer de esta forma la distribución generalizada de Weibull, para la que vamos a considerar las siguientes tres restricciones:

$$\int_0^{\infty} p(x)dx = 1; \quad \int_0^{\infty} x^{\alpha} P_{esc}(x)dx = g_1; \quad \int_0^{\infty} \ln_q(x) P_{esc}(x)dx = g_2; \quad 0 < \alpha < 1 \quad (8)$$

Finalmente obtenemos la siguiente expresión:

$$p(x) = \frac{\alpha}{\lambda_{esc}^{\alpha}} x^{\alpha-1} \left[ 1 + \left( \frac{q-1}{2-q} \right) \frac{x^{\alpha}}{\lambda_{esc}^{\alpha}} \right]^{\frac{1}{1-q}} \quad (9)$$

Ahora bien, para la probabilidad **escort**, definida anteriormente tenemos:

$$p_{esc}(x) = \frac{\alpha \Gamma\left(\frac{q}{q-1}\right) \left(\frac{q-1}{(2-q)\lambda_{esc}^{\alpha}}\right)^{\frac{1+q(\alpha-1)}{\alpha}}}{\Gamma\left(\frac{1+q(\alpha-1)}{\alpha}\right) \Gamma\left(\frac{q}{q-1} + q\left(\frac{1}{\alpha} - 1\right) - \frac{1}{\alpha}\right)} \left[ 1 + \left( \frac{q-1}{2-q} \right) \frac{x^{\alpha}}{\lambda_{esc}^{\alpha}} \right]^{\frac{q}{1-q}} \quad (10)$$

Siendo en este caso,  $\lambda_{esc}$  un parámetro de escala, además verificamos que:

$$\lim_{q \rightarrow \infty} p_{esc}(x) = \frac{\alpha}{\lambda_{esc}} x^{\alpha-1} e^{\frac{-x^\alpha}{\lambda_{esc}}} \quad (11)$$

A partir de esta condición de límite podemos reinterpretar la equivalencia de la distribución de Weibull generalizada con la distribución de Weibull clásica, para lo que tenemos:

$$\lambda_{esc} = \lambda^\alpha ;$$

Entonces, si definimos un  $t_b$  como tiempo de ruptura y los dos parámetros característicos de la Distribución de Weibull tenemos las siguientes expresiones de la distribución clásica y la distribución generalizada:

$$F(t_b) = 1 - \exp \left[ - \left( \frac{t_b}{\alpha} \right)^\beta \right]; \quad F_q(t_b) = 1 - \exp_q \left[ - \left( \frac{t_b}{\alpha} \right)^\beta \right]$$

Donde la correspondencia entre estas distintas distribuciones se puede obtener mediante:

$$\ln q(x) = \frac{x^{1-q} - 1}{1 - q}$$

Contando con Modelos Computacionales que describen el fenómeno de ruptura dieléctrica [6-10] hemos podido interpretar la información experimental que se genera a partir de ensayos dieléctricos de materiales. Es así que, a través de la simulación computacional y bajo condiciones de ruptura idénticas se han generado una familia de árboles con tiempos de falla que siguen una distribución de Weibull caracterizada por los parámetros llamados  $\alpha$  y  $\beta$ . Ahora bien, al variar las condiciones de ruptura se generan otras familias de árboles con otros tiempos característicos. En resumen, desarrollamos distintos protocolos que generan familias de árboles, con distintos parámetros característicos como los siguientes:

**Tabla (1)** Parámetros de la distribución de Weibull para distintos protocolos de simulación

| PARAMETROS | PROTOCOLO I                   | PROTOCOLO II    |
|------------|-------------------------------|-----------------|
| $\alpha$   | $(1,28 \pm 0,13) \times 10^3$ | $56,2 \pm 4,00$ |
| $\beta$    | $3,45 \pm 0,02$               | $8,38 \pm 0,07$ |

Donde hemos podido determinar que el protocolo I favorece el desarrollo de estructuras ramificadas mediante la selección aleatoria del capacitor (Modelo de Capacidades Distribuidas) que va a cortocircuitar del conjunto de capacitores que está en condiciones de ruptura [6,7]. Esta selección representa una contribución entrópica al proceso de crecimiento del árbol eléctrico. En cambio, el protocolo II, favorece el desarrollo de estructuras “runaway” (explosivas), la selección de este protocolo está basada en argumentos energéticos, favorece un comportamiento de rápido crecimiento de la estructura del árbol, que cruza rápidamente el material dieléctrico.

También se ha trabajado sobre una combinación lineal de ambos protocolos con el objeto de poder representar más adecuadamente el proceso de crecimiento del árbol eléctrico y se llegó a la conclusión de que esta combinación lineal representa la energía libre termodinámica a temperatura constante, gobernando el crecimiento del árbol eléctrico. Esta energía libre es la suma de un término entrópico representado por el protocolo I y una energía interna dada por el protocolo II. Estos resultados están de acuerdo con las observaciones experimentales de patrones de ruptura y con la evidencia (experimental) que indica que a velocidades bajas de incremento de potencial, la contribución entrópica debido a fluctuaciones de campo eléctrico es dominante y determina las características geométricas de los patrones de ruptura.

Paralelamente se ha observado que la dimensión fractal de los árboles eléctricos, aumenta hasta llegar a un cierto valor de potencial a partir del cual comienza a disminuir. Esta competencia tiene una profunda relación con la termodinámica subyacente que gobierna el crecimiento del árbol. Finalmente, se ha podido observar que el proceso de ruptura dieléctrica es de naturaleza caótica, además existe una contribución entrópica que está asociada a los efectos de las fluctuaciones del campo eléctrico debido a la presencia de cargas espaciales y a la acumulación del daño ocasionado.

Esta contribución tiene características de cascada, en el sentido de que el daño eléctrico no se produce gradualmente sino abruptamente en un solo paso de potencial. En la medida que el daño sea más o menos gradual, el crecimiento competitivo generará distribuciones diferentes a la simple combinación lineal de las distintas contribuciones. Estas situaciones evidencian la necesidad de contar con un formalismo no lineal, tal como lo es la Termodinámica Estadística No Extensiva.

El fenómeno de ruptura dieléctrica conlleva correlaciones de largo alcance entre todas las partes del dieléctrico que sufre el daño de la ruptura [11]. Estos últimos sistemas, en donde el formalismo BGS parece resultar no ser adecuado, son considerados sistemas anómalos. La ruptura dieléctrica puede ser considerada como un paradigma de la no extensividad, ya que el objeto que es sometido a la ruptura dieléctrica, al cual ahora podemos considerar como un objeto en estado de “falla” (fragmentado) tiene asociada una entropía mayor que cuando se encontraba en el estado “sano” (entero), esto es una consecuencia lógica de considerar explícitamente la transferencia de energía producto de la acumulación progresiva del daño del material.

Hemos caracterizado [1-3] a cada familia de árboles generados por simulación numérica con una dimensión fractal promedio y hemos registrado la variación de la dimensión fractal para distintos patrones de ruptura, esto nos ha llevado a considerar la Multi Fractalidad del fenómeno de ruptura dieléctrica. Efectivamente, algunos conjuntos fractales no pueden ser caracterizados por una única dimensión fractal, sino con una función de distribución que nos dice como varía esta dimensión, este tipo de conjuntos son llamados Multifractales. En nuestro caso de estudio, se ha determinado la dependencia de la dimensión fractal con la tensión aplicada y se ha observado una dependencia no monótona [6,7]. Para establecer una correspondencia entre estas familias de árboles eléctricos hemos recurrido a un formalismo de mecánica estadística introducido por Tsallis, a partir de este formalismo definimos una distribución generalizada de Weibull con un índice entrópico  $q$ .

Se puede establecer una relación de correspondencia entre el índice entrópico  $q$  y la multifractalidad manifiesta de los patrones de ruptura dieléctrica. Esta relación enunciada [3] nos dice que la inversa de la diferencia del índice  $q$  con respecto a la unidad es igual a la diferencia entre los extremos del espectro multifractal del fenómeno en cuestión. Este espectro multifractal se puede obtener a partir del procesamiento digital de las imágenes obtenidas en la filmación del proceso de ruptura dieléctrica. Es importante tener en cuenta que, la clasificación del espectro multifractal aplicada al análisis de imágenes fue introducida por Jacques Lévy-Véhel [3], así como el desarrollo del software para el procesamiento digital de las imágenes (FracLab). A partir de la determinación de este índice entrópico pudimos hallar entonces la distribución Generalizada de Weibull ( $q$ -Weibull) para representar los distintos patrones de ruptura, además, para hallar los parámetros asociados a la distribución de  $q$  Weibull hemos utilizado el procedimiento de Bernard modificado [3].

### 3.- La Termodinámica Estadística No extensiva de Tsallis:

En la actualidad se ha visto en la necesidad de generalizar el concepto de Entropía (y por consiguiente de las funciones de distribución que la maximizan) para describir sistemas con un alto grado de complejidad que se han manifestado en muchos campos de la física no lineal. En 1988, el profesor Tsallis, introdujo un nuevo concepto de Entropía y conjuntamente nuevas restricciones para el correspondiente funcional, el cual se utiliza ampliamente para describir sistemas que poseen propiedades no extensivas.

Esta formulación generalizada de Entropía depende de un parámetro real llamado parámetro de no extensividad,  $q$ , el cual ha sido propuesto de acuerdo a las condiciones específicas que describen el sistema. Exploraremos a continuación la profunda relación existente entre la Entropía de Tsallis y la entropía de Boltzmann Gibbs Shannon (BGS). Recordar que las distribuciones estadísticas buscadas se han obtenido mediante el principio de máxima entropía (PME) y en el primer caso se ha empleado el funcional de entropía denominado de BGS, que es el fundamento de la mecánica estadística clásica y se define para  $p = (p_1, \dots, p_n)$  mediante:

$$S(p) = k_B \sum_{i=1}^n p_i \ln \left( \frac{1}{p_i} \right), \quad (12)$$

Donde se considera un sistema de  $n$  estados posibles, cada uno con probabilidad  $p_i$ , ( $1 \leq i \leq n$ ) tal que la suma de las probabilidades es igual a uno. Ahora bien, la expresión para el caso del continuo queda formulada de la siguiente forma:

$$S = -k \int_{\Omega} p(x) \ln p(x) dx \quad (13)$$

Si bien, la propuesta del funcional de BGS ha mostrado efectividad, su uso se restringe a los sistemas conocidos como aditivos o extensivos; tales como son aquellos que cumplen con la propiedad de aditividad  $S(A + B) = S(A) + S(B)$  donde A y B son dos sistemas independientes en el sentido probabilístico. Para sistemas no extensivos es necesario generalizar la entropía de forma tal que se cubran los casos subaditivos o superaditivos. En este sentido, la estadística no extensiva de Tsallis es considerada como un nuevo paradigma en mecánica estadística. La entropía  $S_q$ , definida por Tsallis, para cada real q, con  $q \neq 1$  está dada por:

$$S_q(p) = k \frac{1 - \sum_{i=1}^n (p_i)^q}{q-1} \quad (14)$$

Y es equivalente a la entropía de BGS cuando  $q=1$ . Nuevamente, nos queda la expresión para el caso del continuo, de la siguiente forma:

$$S_q = k \frac{1}{q-1} \left[ 1 - \int_{\Omega} p^q(x) dx \right], \quad q \in R \quad (15)$$

La entropía de Tsallis [12] verifica las siguientes propiedades:

- $S_q$  es extensión de S en el siguiente sentido, si  $q \rightarrow 1$  se tiene  $S_q \rightarrow S$ .
- $S_q(P) = 0$  sólo en el caso de que alguna  $p_i = 1$ .
- El máximo para  $S_q$ , bajo la única restricción de que la suma de todas las probabilidades sea igual a uno, se obtiene para  $p_i = 1/n$ , para  $1 \leq i \leq n$ .
- Propiedad de no extensividad:

Supóngase que se tienen dos sistemas independientes A y B en el sentido probabilístico, entonces se tiene:

$$\frac{S_q(A+B)}{k} = \frac{S_q(A)}{k} + \frac{S_q(B)}{k} + (1-q) \frac{S_q(A)S_q(B)}{k^2}$$

O bien de forma equivalente:

$$S_q(A+B) = S_q(A) + S_q(B) + \frac{1-q}{k} S_q(A)S_q(B)$$

Este nuevo paradigma en mecánica estadística, comprende una familia de formulaciones, todas ellas basadas en la medida de información no extensiva de Tsallis:

$$S_q(p) = k_B \frac{1 - \sum_{i=1}^n p_i^q}{q-1} = k_B \sum_{i=1}^n p_i \ln_q p_i \quad (16)$$

Donde,  $k_B$  es una constante positiva, P es un conjunto normalizado de probabilidades y n es el número de microestados accesibles al sistema. El parámetro real q es el índice de no extensividad y está relacionado estrechamente con la dinámica microscópica del sistema. Es importante destacar que en la expresión anterior usamos las funciones denominadas q-logaritmo y q-exponencial [13], ambas funciones son generalizaciones de las funciones ordinarias logaritmo natural y exponencial de la entropía de BGS. Tal relación se establece en las siguientes ecuaciones:

$$\ln_q x = \frac{1 - x^{1-q}}{q-1}, \quad \exp_q x = [1 + (1-q)x]^{1/(1-q)} \quad (17)$$

En el caso que describamos un sistema en contacto térmico con un reservorio, el conjunto canónico, además de la restricción de que la suma de las probabilidades sea uno, es necesaria otra que establezca cómo calcular la energía interna. En la literatura sobre termodinámica no extensiva, en este caso, nos encontramos

con tres posibles elecciones, la primera de ellas fue introducida por el propio Tsallis, [12,14] luego se introduce una modificación a la misma hecha por Curado y Tsallis:

$$\langle U \rangle_q^{TO} = \sum_{i=1}^N p_i \varepsilon_i, \quad \langle U \rangle_q^{CT} = \sum_{i=1}^N p_i^q \varepsilon_i \quad (18)$$

En las expresiones anteriores, el supra índice TO, indica Tsallis Original, el otro supra índice CT es la energía interna del sistema, la tercera elección para la restricción de la energía interna se debe a Tsallis, Mendes y Plastino (supra índice TMP) y es actualmente la más usada, en ella se pesan los autovalores del Hamiltoniano con un conjunto de probabilidades:

$$\langle U \rangle_q^{TMP} = \frac{\sum_{i=1}^n p_i^q \varepsilon_i}{\sum_{i=1}^n p_i^q} ; \quad (19)$$

Pero es a partir de la expresión (17) que podemos, establecer la correspondencia entre las distribuciones estadísticas clásicas y las distribuciones estadísticas que dependen del parámetro entrópico q (distribución exponencial con una q-exponencial y distribución de Weibull con una q-Weibull).

#### 4.- Estimadores Estadísticos y el Principio de Máxima Verosimilitud

Dentro del marco de la Teoría de las probabilidades y la Estadística [15], se trata con el problema de extraer información a partir de observaciones que pueden ser no confiables ya que pueden estar corrompidas por ruido. Cuando las perturbaciones son modeladas como procesos aleatorios, las observaciones pueden describirse como realizaciones de variables aleatorias, de esta forma se pueden describir las observaciones con la variable aleatoria  $y_N = (y_{(1)}, y_{(2)}, \dots, y_{(N)})$  que toma valores en  $R^N$ , además se asume que  $y_N$  tiene una función de densidad de probabilidad de la siguiente forma:

$$f(\theta, x_1, x_2, \dots, x_N) = f_y(\theta, x_N)$$

Resulta entonces que:

$$P(y^N \in A) = \int_{x^N \in A} f_y(\theta; x^N) dx^N$$

Donde  $\theta$  es un vector de parámetros que describe las propiedades de la variable observada. Este vector supuestamente no se conoce y el propósito de las observaciones es estimarlo a partir de  $y_N$ . Esto es realizado mediante el siguiente estimador:

$$\hat{\theta}(y^N): R^N \xrightarrow{\text{función}} R^d; \quad \theta \in R^d \quad (20)$$

Existen muchas funciones de estimación, pero particularmente estamos interesados en una que maximiza la probabilidad de que el evento observado haya sido generado por un grupo de datos es el denominado **Estimador de máxima Verosimilitud**. Esta función la podemos definir de la siguiente manera: la función de densidad conjunta para el vector aleatorio a ser observado está dada por la función de densidad de probabilidad definida anteriormente. La probabilidad de que la realización (observación) tome el valor  $y^N$ , es entonces proporcional a:

$$f(\theta, x^N)$$

Podemos visualizar que ésta es una función determinística de  $\theta$  una vez que un valor numérico  $y_N$  es introducido. Esta función es llamada **función de Verosimilitud** y refleja la “probabilidad” de que el evento observado ocurra. Nos queda entonces, seleccionar un estimador de  $\theta$  de forma tal que los eventos observados resulten “tan probables como sean posibles” de aquí que definimos:

$$\hat{\theta}_{ML}(y^N) = \arg \max_{\theta} f_y(\theta, y^N)$$

Donde la maximización es realizada para un  $y^N$  fijo. Esta función es conocida como **Estimador de Máxima Verosimilitud** (MLE - Maximun likelihood estimator). A menudo es dificultoso calcular las propiedades estadísticas de este estimador, en lugar de esto se suele realizar un análisis asintótico cuando el número de

muestras N, tiende a infinito. Veamos a continuación un resultado clásico del análisis asintótico de la estima obtenido por Wald y Crámer [15]. Suponemos que las variables aleatorias  $\{y_{(i)}\}$  son independientes y se encuentran idénticamente distribuidas de forma tal que:

$$f_y(\theta, x_1, x_2, \dots, x_N) = \prod_{i=1}^N f_{y(i)}(\theta, x_i) \quad (21)$$

Suponemos también que la distribución de  $y^N$  está dada por  $f_y(\theta_0; x^N)$  para algún valor de  $\theta_0$ . Entonces la variable aleatoria:

$$\hat{\theta}_{ML}(y^N) \xrightarrow{\text{Probabilidad_Uno}} \theta_0; \quad N \longrightarrow \infty$$

Además, la variable aleatoria siguiente:

$$\sqrt{N} \left( \hat{\theta}_{ML}(y^N) - \theta_0 \right)$$

converge en una distribución normal, con media cero y matriz de covarianza dada por la cota inferior de Crámer – Rao, definida a partir de  $M^{-1}$ . Aquí definimos a la matriz M como:

$$M = -E \left[ \frac{d^2}{d\theta^2} \log f_y(\theta; y^N) \right]_{\theta=\theta_0}$$

Ahora bien, la calidad de un estimador puede ser cuantificada mediante la matriz de covarianza correspondiente a la estima definida por:

$$P = E \left[ \left[ \hat{\theta}(y^N) - \theta_0 \right] \left[ \hat{\theta}(y^N) - \theta_0 \right]^T \right] \quad (22)$$

En este sentido estamos interesados en un estimador tal que haga P lo más pequeña posible, por lo tanto, buscar el límite inferior de P que se puede obtener con estimadores no desviados, es la denominada **cota de Crámer - Rao**. Un estimador (no desviado) se dice que es eficiente desde el punto de vista estadístico si la covarianza de la estima alcanza la cota de Cramer Rao, entonces:

$$P = E \left[ \left[ \hat{\theta}(y^N) - \theta_0 \right] \left[ \hat{\theta}(y^N) - \theta_0 \right]^T \right] \geq M^{-1} \quad (23)$$

Donde M es la matriz de información de Fischer definida de la siguiente forma:

$$M = E \left\{ \left[ \frac{d}{d\theta} \log f_y(\theta; y^N) \right] \left[ \frac{d}{d\theta} \log f_y(\theta; y^N) \right]^T \right\} \Big|_{\theta_0} = -E \frac{d^2}{d\theta^2} \log f_y(\theta; y^N) \Big|_{\theta_0} \quad (24)$$

La información de Fischer cumple con las propiedades de las medidas de información:

- a) Es no Negativa
- b) Es aditiva para observaciones independientes.

Además esta información es mayor cuanto menor es la cota de Cramer-Rao, por ejemplo: la cota de Cramer-Rao para N observaciones es N veces inferior a la de una observación y justamente se deduce de la matriz de información de Fischer [15]. Veremos a continuación la aplicación del Principio de Máxima Verosimilitud a las distribuciones estadísticas de nuestro interés y además exploraremos algunas equivalencias importantes que nos permitirán en un futuro simplificar las estimación de los parámetros, independientemente que trabajemos con un conjunto de datos truncados o censurados.

## 5.- La Distribuciones Generalizadas en los fenómenos de Ruptura Dieléctrica

La distribución de Weibull y la distribución Generalizada de Weibull (q Weibull) tiene las siguientes expresiones, donde  $t_b$  es el tiempo de ruptura y  $\alpha, \beta$  son los parámetros Weibull:

$$F(t_b) = 1 - \exp \left[ - \left( \frac{t_b}{\alpha} \right)^\beta \right]; \quad F_q(t_b) = 1 - \exp_q \left[ - \left( \frac{t_b}{\alpha} \right)^\beta \right] \quad (25)$$

Ahora bien, recordemos que cuando el parámetro  $\beta$  es igual a uno, lo cual se puede interpretar como que la tasa de fallos es constante, la q-Weibull se transforma en una q-exponencial, esta característica particular también se cumple en la relación existente entre las distribuciones Weibull y exponencial clásicas.

La forma general que adopta la q-exponencial, según la definición de Tsallis [16] y colaboradores es la siguiente:

$$P_{q,k}(X \geq x) = \left( 1 - \frac{(1-q)x}{k} \right)^{1/(1-q)} \quad (26)$$

Analizaremos ahora la relación que existe entre la distribución q-exponencial y la denominada distribución Generalizada de Pareto, para lo cual realizamos el siguiente cambio de parámetros:

$$q = 1 + \frac{1}{\theta}, \quad k = \frac{\sigma}{\theta} \quad (27)$$

A partir de esta transformación obtenemos la expresión similar a una distribución estadística que es conocida como la Distribución generalizada de Pareto

$$P_{\theta,\sigma}(X \geq x) = \left( 1 + \frac{x}{\sigma} \right)^{-\theta} \quad (28)$$

Con la función de densidad de probabilidad asociada:

$$P_{\theta,\sigma}(x) = \frac{\theta}{\sigma} \left( 1 + \frac{x}{\sigma} \right)^{-\theta-1} \quad (29)$$

La distribución Generalizada de Pareto es una distribución que permite estudiar determinados valores extremos que se encuentran por encima de un umbral alto, se puede demostrar que mediante un resultado límite, la distribución de estos excesos puede ser aproximada en forma adecuada a la función Generalizada de Pareto, ahora bien, la distribución de Pareto Clásica [15] está definida por la expresión:

$$F(x) = 1 - \left( \frac{x}{x_0} \right)^{-\alpha}; \quad x \geq x_0 > 0 \quad (30)$$

Donde, en este caso,  $\alpha$  es un parámetro de forma y además  $x_0$  es un parámetro de escala. Ahora bien, una variante de la distribución clásica de Pareto está definida por:

$$F(x) = 1 - \left( 1 + \frac{x - \mu}{x_0} \right)^{-\alpha}; \quad x \geq \mu \quad (31)$$

Es ocasionalmente llamada distribución de Pareto de tres parámetros, se estudia también un caso especial en el que un parámetro es igual a cero:

$$F(x) = 1 - \left( 1 + \frac{x}{x_0} \right)^{-\alpha}; \quad \mu = 0; \quad x \geq 0 \quad (32)$$

Esta distribución es frecuentemente denominada distribución de Pareto tipo II y en este caso los parámetros  $x_0, \alpha$  son mayores que cero. Pero retomemos ahora la expresión (26) y obtengamos la función de máxima verosimilitud en este caso particular será una **log likelihood**. La misma tiene la siguiente forma:

$$\log P_{\theta,\sigma}(x_1^n) = -n \log \sigma + n \log \theta - (\theta + 1) \sum_{i=1}^n \log \left( 1 + \frac{x_i}{\sigma} \right) \equiv l(\theta, \sigma)$$

A partir de esta función podemos obtener las estimaciones a través del método MLE y las expresiones de las estimas nos quedan de la siguiente forma:

$$\hat{\theta} = n \left[ \sum_{i=1}^n \log 1 + \frac{x_i}{\hat{\sigma}} \right]^{-1} ; \quad \hat{\sigma} = \frac{\theta+1}{n} \sum_{i=1}^n \frac{x_i}{1 + x/\hat{\sigma}} \quad (33)$$

En el caso de que apliquemos el Principio de Máxima Entropía (PME), para el caso de la distribución generalizada de Pareto (homóloga a la distribución q exponencial propuesta por Tsallis y colaboradores) tendríamos que hacer el siguiente planteo matemático: Vamos a realizar primero un cambio de variables [16], para la Distribución de Pareto, con el fin de simplificar los cálculos necesarios, entonces partimos de la siguiente expresión:

$$F(x) = 1 - \left[ 1 - \frac{a(x-c)}{b} \right]^{\frac{1}{a}} \quad a \neq 0 \quad (34)$$

La cual representa una Distribución Generalizada de Pareto, con parámetros **a, b, c**, a esta expresión debemos combinarla con la función que maximiza la entropía de acuerdo con el PME, recordemos que la expresión de la entropía tiene la forma genérica siguiente:

$$H(f) = - \int_{-\infty}^{\infty} f(x; \theta) \ln f(x; \theta) dx \quad ; \quad \int_{-\infty}^{\infty} f(x; \theta) dx \quad (34)$$

En nuestro caso particular la función de la entropía adoptará la siguiente forma:

$$H(f) = -\ln a - \ln(1-\lambda) + \ln b + \lambda E \left\{ \ln \left[ 1 - \frac{a(x-c)}{b} \right] \right\} \quad (35)$$

donde  $\lambda$  son los multiplicadores de Lagrange que nos permiten incorporar las restricciones vistas a la ecuación funcional. Tanto las ecuaciones de verosimilitud como las de máxima entropía, han sido derivadas y resueltas numéricamente, de donde hemos obtenido los Estimadores de Máxima Verosimilitud y de Máxima Entropía.

Vemos a continuación una tabla comparativa del desempeño [15] de estos estimadores. Analizamos el desempeño de las dos formas de estimar los parámetros de la Distribución de Pareto Generalizada, a través de dos índices de desempeño que definimos a continuación:

$$BIAS = \frac{E(\hat{x}) - x}{x} \quad ; \quad RMSE = \frac{E \left[ \left( \hat{x} - x \right)^2 \right]^{0.5}}{x} \quad (36)$$

Para la estimación de parámetros con los distintos métodos, se genera una matriz de datos de prueba (1000 renglones, una columna) de acuerdo con [4] para obtener un proceso estocástico del cual se estiman sus parámetros, Con esta población base, adoptamos distintos tamaños de muestras y a partir de estas muestras (20, 50, 100) calculamos los índices de desempeño (Muestreo de MonteCarlo) obteniendo los siguientes resultados:

**Tabla (2)** índices de desempeño de los estimadores para distintos tamaños de muestras

| Prob. | Muestra | Método | BIAS  | RMSE |
|-------|---------|--------|-------|------|
| 0,8   | 20      | MLE    | -0,01 | 0,12 |
|       | 20      | PME    | -0,03 | 0,13 |
|       | 50      | MLE    | 0,1   | 0,09 |
|       | 50      | PME    | 0     | 0,08 |
|       | 100     | MLE    | -0,04 | 0,15 |
|       | 100     | PME    | 0,03  | 0,15 |
| 0,9   | 20      | MLE    | -0,08 | 0,16 |
|       | 20      | PME    | -0,06 | 0,16 |
|       | 50      | MLE    | -0,01 | 0,07 |
|       | 50      | PME    | 0,06  | 0,13 |

|     |     |      |      |
|-----|-----|------|------|
| 100 | MLE | 0    | 0,06 |
| 100 | PME | 0,05 | 0,09 |

Como conclusión importante que podemos extraer de la visualización de estos índices de desempeño es que cuando el tamaño de la muestra es grande en comparación a la población, el estimador por máxima verosimilitud resulta eficiente, no así con muestras pequeñas, donde el estimador por máxima entropía muestra una considerable eficiencia. Es de destacar que estos son los primeros resultados, lo que habrá que cotejar con otras simulaciones y con datos provenientes de los ensayos experimentales. Si bien, hemos trabajado con distribuciones truncadas de amplia utilización en estadística industrial, en esta dirección, es también importante destacar que la relación que ha establecido Tsallis entre la distribución q-exponencial y la distribución generalizada de Pareto nos permitirá trabajar con distintos tipos de ensayos, contengan estos datos truncados y/o censurados.

Algunas implementaciones en donde se puede tener acceso a rutinas computacionales de estimación de los parámetros de la distribución q-exponencial o Pareto Generalizada de uso free y desarrolladas en lenguaje R se pueden encontrar en [16]. Estas rutinas fueron tomadas como base para el desarrollo del software de estimación de los parámetros de las distribuciones q-Weibull, q-exponencial y generalizada de Pareto y adaptadas para ser utilizadas en el entorno de trabajo provisto por el software STATA [17].

## 6.- Conclusiones y trabajos Futuros

Hemos podido establecer una correspondencia importante entre los distintos modelos de falla de dieléctricos que desarrollamos en trabajos anteriores. Primeramente trabajamos con una distribución de Weibull clásica de dos parámetros para caracterizar la falla en dieléctricos, estos parámetros fueron obtenidos para distintas familias de árboles eléctricos, además, con la intención de mejorar estos modelos de falla, también se ha trabajado con una distribución de Weibull de tres parámetros [2].

Por otro lado también se recurrió a un enfoque desde la termodinámica estadística no extensiva para caracterizar las fallas en los dieléctricos, este enfoque nos permitió definir una distribución generalizada de Weibull (q-Weibull) para construir los modelos de falla [1-3]. Luego trabajamos con una distribución de Weibull truncada, dado que podemos implementar ensayos acotados en el tiempo. El poder trabajar experimentalmente con ensayos registrados en función del tiempo, permite el truncamiento artificial de las muestras, deteniendo ficticiamente en poblaciones truncadas a distintos (t) y estudiando el comportamiento de los ajustes en función del truncamiento de datos [4].

Finalmente, en este trabajo retomamos el enfoque termo-estadístico y establecemos una correspondencia importante entre la distribución de Weibull truncada y la distribución obtenida a través del enfoque de Tsallis [12,14], esta correspondencia nos permite unificar los distintos modelos de falla con la distribución denominada Distribución Generalizada de Pareto, la cual puede, como vimos, ser estimada directamente de los ensayos de ruptura o bien utilizando el índice entrópico q, obtenido indirectamente, mediante el procesamiento digital de imágenes.

Los esfuerzos actuales y futuros se hallan orientados a generar y compaginar información experimental y computacional necesaria para estimar los parámetros característicos de estos distintos modelos de fallas, además de evaluar su desempeño en función del tipo de ensayo a realizar. Para esta estimación necesitamos contar ya con una gran cantidad de ensayos para poder en definitiva inferir la población a partir de las muestras obtenidas.

## Referencias:

- Herrera L. A., Irurzun I. M., Razzitte A. C., Mola E. E., XVI Congreso Argentino de Fisicoquímica y Química inorgánica, Año 2009.
- ENERGEIA, Cuaderno de Investigación, publicación Anual del Departamento de investigación Institucional, Año 7 N° 7, noviembre 2009.
- ENERGEIA, Cuaderno de Investigación, publicación Anual del Departamento de investigación Institucional, Año 8 N° 8, noviembre 2010.
- ENERGEIA, Cuaderno de Investigación, publicación Anual del Departamento de investigación Institucional, Año 9 N° 9, noviembre 2011.
- Tsallis C. "Nonextensive Statistical Mechanics and Its Applications" Lecture in Physics, Springer; Heidelberg, 2001.
- Damming Quiña P., Herrera L. A., Irurzun I. M., Mola E. E., "A Capacitive Model for Dielectric Breakdown in polymer materials" Computational material science, Ed. Elsevier, año 2008.

Herrera L. A., "Desarrollo de Modelos Computacionales aplicados a la simulación del Envejecimiento y la Ruptura dieléctrica de Materiales No homogéneos (Polímeros)", Tesis de Maestría, Facultad de Ingeniería, Universidad de Buenos Aires, 2006.

Dammig Quiña P., "Ruptura Dieléctrica en Sólidos", Tesis Doctoral, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, año 2010.

Razzitte, A. C. "Características fractales de la ruptura Dieléctrica y Respuesta en Frecuencia de Sólidos Aisladores", Tesis Doctoral, Facultad de Ciencias exactas, Universidad Nacional de La Plata, 2001.

P. L. Dammig Quiña, I. M. Irurzun, L. M. Salvatierra, E. E. Mola. Physical Review E 82, 041106, 2010.

Dissado L. A. and Fothergill J.C., "Electrical Degradation and Breakdown in Polymers" Peregrinus Press for IEEE, 1992.

Gell-Mann M. and Tsallis C., "Nonextensive Entropy - Interdisciplinary Applications", Oxford University Press, 2004.

<http://fraclab.saclay.inria.fr/works/signal-and-image-processing>

Tsallis C. "Introduction to Nonextensive Statistical Mechanics, Approaching a Complex World", Springer; New York, 2009.

Abernethy B. Robert "The New Weibull Handbook" Ed. Abernethy, 2004.

<http://cscs.umich.edu/~crshalizi/research/tsallis-MLE/>

<http://www.stata.com/site-index>