

NUEVOS MODELOS DE DISTRIBUCIONES DE EXTREMOS BASADOS EN APROXIMACIONES EN LAS RAMAS

*Castillo E.
Moreno E. y Puig-Pey J.
Dpto. de Matemáticas Aplicadas
E.T.S. de Ingenieros de Caminos, Canales y Puertos
Universidad de Santander*

RESUMEN

En este trabajo se presenta una metodología que permite clasificar funciones de distribución absolutamente continuas unidimensionales atendiendo a sus ramas. La idea básica es que, en las ramas la función de distribución difiere en un infinitésimo del valor uno o cero dependiendo de la rama de interés. La principal ventaja de esta clasificación es su aplicación a la teoría de distribuciones de extremos. En esta línea se obtienen nuevas familias de distribuciones de extremos. Entre ellas, las clásicas de Gumbel, Fréchet y Weibull surgen como casos particulares, poniendo de manifiesto la insuficiencia de esas tres familias para resolver todos los casos prácticos. La teoría presentada se aplica al análisis de datos de oleaje.

PALABRAS CLAVE

Aproximación. Distribución en las ramas. Distribuciones de valores extremos. Distribuciones asintóticas. Modelos.

1. INTRODUCCION

En el tratamiento estadístico de problemas de valores extremos se utilizan, con general aceptación, tres familias de distribuciones cuyas expresiones para máximos y mínimos son las siguientes:

(*) Recibido, Mayo, 1981

MAXIMOS

MINIMOS

$$\begin{array}{ll}
 \text{I} & F(x) = \exp\{-\exp[-a(x-b)]\}, \quad F(x) = 1 - \exp\{-\exp[a(x-b)]\}, \\
 & \qquad \qquad \qquad a > 0 \qquad \qquad \qquad \qquad \qquad \qquad a > 0 \\
 \\
 \text{II} & F(x) = \exp[-(b/x)^a], \quad \left\{ \begin{array}{l} a > 0 \\ b > 0 \\ x \geq 0 \end{array} \right. \quad F(x) = 1 - \exp[-(b/x)^a], \quad \left\{ \begin{array}{l} a > 0 \\ b < 0 \\ x \leq 0 \end{array} \right. \\
 \\
 \text{III} & F(x) = \exp\{-[(w-x)/b]^a\}, \quad \left\{ \begin{array}{l} a > 0 \\ b > 0 \\ x \leq w \end{array} \right. \quad F(x) = 1 - \exp\{-[(x-w)/b]^a\}, \quad \left\{ \begin{array}{l} a > 0 \\ b > 0 \\ x \geq w \end{array} \right.
 \end{array}$$

Sin embargo los planteamientos matemáticos basados en la estabilidad asintótica, dominios de atracción, etc. que han conducido a la obtención de estas tres familias de distribuciones son confusos, por no surgir estrictamente de razones físicas. Incluso se manejan conceptos, como es el de estabilidad, con significado diferente para diversos autores.

Ya algunos investigadores han hecho mención de este hecho y así por ejemplo Green (10) dice textualmente: “La estabilidad de los máximos es un artificio matemático. No hay razón para creer que las distribuciones experimentales tengan ramas tales que las distribuciones de sus máximos converjan a una de las llamadas familias estables. De hecho los máximos pueden tener cualquier distribución e incluso para una misma distribución la distribución del máximo puede variar completamente al variar el tamaño muestral”.

Por otra parte, en la literatura existente, son varios los autores que muestran ejemplos en los que ninguna de las tres distribuciones clásicas da resultados aceptables (Weinstein (17), Copeiro (4), etc...) e incluso algunos que ajustan con éxito distribuciones del mínimo para máximos (Copeiro (4), Houmb et al (12), Battjes (1), etc).

Todo ello demuestra que existe una gran confusión en el tema y que se hace necesaria una revisión y clarificación de conceptos.

2. DISCUSION DEL PROBLEMA

Las funciones de distribución del máximo y del mínimo de una muestra aleatoria simple de tamaño n , procedente de una variable aleatoria cuya función de distribución es $F(x)$, vienen dadas por $[F(x)]^n$ y $1 - [1 - F(x)]^n$, respectivamente. Sin embargo, en muchos casos prácticos ello no resulta de gran utilidad, puesto que $F(x)$ es desconocida.

Uno de los primeros problemas que se planteó en el desarrollo de la teoría de valores extremos fue el de la estabilidad.

Sea una familia de variables aleatorias con función de distribución

$$\{F(x; \theta^*); \theta^* \in \Omega \subset \mathbb{R}^m\}$$

donde $\theta^* = (\theta_1, \theta_2, \dots, \theta_m)$ representa un parámetro m -dimensional de dicha familia. Se dice que la familia es estable respecto de los máximos si se verifica que el máximo de una muestra aleatoria simple de tamaño n procedente de la familia pertenece a ella, es decir, si:

$$\forall \theta_1^* \in \Omega, \exists \theta_2^* \in \Omega / F^n(x; \theta_1^*) = F(x; \theta_2^*); \forall x \in \mathbb{R} \quad (1)$$

La solución de este problema no ofrece gran complicación ya que cualquier familia es ampliable a una familia estable sin más que incorporar el parámetro n como una componente más del vector paramétrico de la familia, es decir, considerar la nueva familia

$$H(x; \theta^*, n) = F^n(x; \theta^*); \theta^* \in \Omega \subset \mathbb{R}^m, n \in \mathbb{N} \quad (2)$$

El trabajo con familias estables facilita la resolución de problemas prácticos, pero ello exige el conocimiento de $F(x)$, lo que no es habitual.

Con objeto de poder ajustar la distribución del máximo de una muestra aleatoria simple sin necesidad de conocer $F(x)$, diversos investigadores han estudiado las condiciones bajo las cuales la sucesión de máximos (mínimos), para tamaños de muestra crecientes, convergen en distribución a una familia dada. No obstante, la convergencia plantea algunos problemas, especialmente para distribuciones no limitadas en la rama de interés. Así, dada una muestra aleatoria simple (x_1, x_2, \dots, x_n) , procedente de una distribución $F(x)$, como la distribución del máximo

$X_{m\acute{a}x}$ es $[F(x)]^n$, cuando n tiende a infinito, para cualquier valor de x dado se cumplirá

$$\lim_{n \rightarrow \infty} F_{X_{m\acute{a}x}}(x) = \begin{cases} 1 & \text{si } F(x) = 1 \\ 0 & \text{si } F(x) < 1 \end{cases} \quad (3)$$

Por esta razón surge el concepto de dominio de atracción de máximos de una variable aleatoria. Se dice que una distribución $F(x)$ pertenece al dominio de atracción de máximos de una distribución $G(x)$ si existen sucesiones $\{a_n\}$, $a_n > 0$ y $\{b_n\}$ tales que

$$F^n(a_n x + b_n) \xrightarrow{D} G(x) \quad (4)$$

Entonces, supuesto $G(x)$ continua para todo x y n suficientemente grande, es posible sustituir la función de distribución del máximo $F^n(x)$ por $G\left(\frac{x - b_n}{a_n}\right)$. Para las tres familias clásicas citadas, no es necesario conocer a_n y b_n pues estas constantes forman parte de los parámetros de la familia.

Sin embargo, la convergencia en distribución exigida en la definición de dominio de atracción es demasiado fuerte pues exige convergencia puntual en todos los puntos de continuidad de la distribución límite. Ello da lugar a que no todas las funciones de distribución pertenecen al dominio de atracción de máximos de una de las tres familias clásicas, lo cual equivale a decir que existen distribuciones tales que las distribuciones de sus máximos (mínimos), no pueden ser aproximadas por ninguna de esas tres familias. Por otra parte, cuando n es pequeño, dicha sustitución no es correcta, pues está basada en un resultado asintótico.

3. APROXIMACION DE DISTRIBUCIONES EN LA RAMA DERECHA

El problema anteriormente expuesto puede resolverse satisfactoriamente si se tienen en cuenta las siguientes consideraciones:

- a) La función de distribución del máximo se utiliza en la mayoría de los casos prácticos para valores grandes de X , lo cual equivale

a decir que sólo la rama derecha de dicha distribución es de interés. Ahora bien, dicha rama viene perfectamente determinada por la misma rama de la distribución inicial, por lo que no será necesario conocer $F(x)$ en su totalidad, sino solamente en la rama de interés.

- b) Para valores grandes de X , la función de distribución del máximo difiere en un infinitésimo del valor uno. Por tanto se trata de aproximar $1 - F_{X_{\max}}(x)$ por cierto infinitésimo $\epsilon(x)$ para x grandes e identificar qué funciones de distribución pueden aproximarse utilizando un mismo infinitésimo.

3.1. Aproximación de distribuciones no limitadas en la rama derecha

Con objeto de analizar las funciones de distribución según su rama derecha, se enuncian seguidamente una serie de teoremas que permiten clasificar el conjunto \mathbf{F} de las funciones de distribución absolutamente continuas unidimensionales, y que se pueden enunciar análogamente para el conjunto de todos los infinitésimos de una variable.

TEOREMA I.— Sea \mathbf{F} el conjunto de las funciones de distribución absolutamente continuas unidimensionales. La relación, R_1 , definida mediante

$$F_1(x) R_1 F_2(x) \Leftrightarrow \begin{cases} \lim_{x \rightarrow \infty} \frac{1 - F_1(x)}{1 - F_2(x)} = 1 & \text{si } \begin{cases} F_1(x) < 1; \forall x < \infty \\ F_2(x) < 1; \forall x < \infty \end{cases} & \text{(a)} \\ \exists K \in \mathbb{R} / F_1(x) = F_2(x) = 1 & ; \quad \forall x > K & \text{(b)} \end{cases} \quad (5)$$

es una relación de equivalencia y particiona el conjunto \mathbf{F} , en clases disjuntas tales que la sustitución en la rama derecha de una distribución puede hacerse sólo por otra de la misma clase. La sustitución es válida ya que entonces $1 - F_1(x) \simeq 1 - F_2(x) \Rightarrow F_1(x) \simeq F_2(x)$ en el caso (a) y $F_1(x) = F_2(x)$ en el (b).

Evidentemente esta clasificación engloba en una misma clase a todas las distribuciones limitadas en la rama derecha, y por tanto sólo servirá para clasificar las distribuciones no limitadas en dicha rama.

Dos distribuciones que pertenecen a la misma clase del conjunto cociente \mathbf{F}/R_1 se dice que son equivalentes en su rama derecha.

Una definición equivalente a ésta aparece ya en Resnick (14), sin embargo este autor se limita a estudiar las relaciones entre la clasificación resultante de este nuevo concepto y la asociada al concepto de dominio de atracción.

El siguiente teorema tiene por objeto definir una clasificación más general de los subconjuntos de \mathbf{F} .

TEOREMA II.— Sea \mathbf{F} el conjunto de las funciones de distribución absolutamente continuas unidimensionales. La relación, R_2 , definida mediante

$$F_1(x) R_2 F_2(x) \Leftrightarrow \begin{cases} \lim_{x \rightarrow \infty} \frac{1 - F_1(x)}{1 - F_2(x)} = c; c \in \mathbb{R}^+ \text{ si } \begin{cases} F_1(x) < 1; \forall x < \infty \\ F_2(x) < 1; \forall x < \infty \end{cases} \quad (a) \\ \exists K \in \mathbb{R}/F_1(x) = F_2(x) = 1; \forall x > K \quad (b) \end{cases} \quad (6)$$

es una relación de equivalencia y por tanto particiona el conjunto \mathbf{F} .

Como en el caso anterior, todas las distribuciones limitadas en la rama derecha, quedan contenidas en una misma clase, que pertenece tanto al conjunto cociente \mathbf{F}/R_2 como al \mathbf{F}/R_1 .

Dos distribuciones que pertenecen a la misma clase de \mathbf{F}/R_2 se dice que son del mismo orden en su rama derecha. Cada clase de \mathbf{F}/R_2 constituida por distribuciones no limitadas en su rama derecha contiene al menos una clase de \mathbf{F}/R_1 , precisamente la clase de las distribuciones tales que $\lim_{x \rightarrow \infty} \frac{1 - F_1(x)}{1 - F_2(x)} = 1$.

En \mathbf{F}/R_2 puede definirse una relación en orden parcial estricto como sigue: $F^*(x) \in \mathbf{F}/R_2$ es de orden inferior a $G^*(x) \in \mathbf{F}/R_2$ si se tiene

$$\lim_{x \rightarrow \infty} \frac{1 - F(x)}{1 - G(x)} = \infty \quad (7)$$

donde $F(x)$ y $G(x)$ pertenecen a $F^*(x)$ y $G^*(x)$ respectivamente.

COROLARIO I.— La distribución del máximo de una muestra aleatoria simple de tamaño n , procedente de una población con función de distribución no limitada en su rama derecha, y la distribución inicial son del mismo orden en su rama derecha.

En efecto, como para todo x es $F(x) < 1$, y haciendo $V(x) = 1 - F(x)$, se tiene:

$$\begin{aligned} \frac{1 - F(x)}{1 - F(x)^n} &= \frac{V(x)}{1 - (1 - V(x))^n} = \\ &= \frac{V(x)}{1 - 1 + n V(x) - \binom{n}{2} V(x)^2 + \dots + (-1)^{n+1} V(x)^n} = \\ &= \frac{1}{n - \binom{n}{2} V(x) + \dots + (-1)^{n+1} V(x)^{n-1}} \end{aligned}$$

Entonces, teniendo en cuenta que $V(x) \rightarrow 0$ cuando $x \rightarrow \infty$ se deduce que:

$$\lim_{x \rightarrow \infty} \frac{1 - F(x)}{1 - F(x)^n} = \frac{1}{n} \quad (8)$$

Seguidamente se enuncia un teorema que da una clasificación más general de las funciones de distribución según su rama derecha.

TEOREMA III.— Sea el conjunto de las funciones de distribución absolutamente continuas unidimensionales. La relación, R_3 , definida mediante

$$F_1(x) R_3 F_2(x) \Leftrightarrow \begin{cases} \lim_{x \rightarrow \infty} \frac{1 - F_1(x)}{[1 - F_2(x)]^n} = c; a, c \in \mathbb{R}^+ \text{ si } \begin{cases} F_1(x) < 1; \forall x < \infty \\ F_2(x) < 1; \forall x < \infty \end{cases} & (a) \\ \exists K \in \mathbb{R} / F_1(x) = F_2(x) = 1 \quad ; \quad \forall x > K & (b) \end{cases} \quad (9)$$

es una relación de equivalencia.

Igual que para las relaciones R_1 y R_2 , el conjunto de las distribuciones limitadas en la rama derecha constituyen una clase de \mathbf{F}/R_3 que pertenece asimismo a \mathbf{F}/R_1 y \mathbf{F}/R_2 .

Dos distribuciones pertenecientes a la misma clase de \mathbf{F}/R_3 se dice que son del mismo tipo en su rama derecha. Cada clase de \mathbf{F}/R_3 constituida por distribuciones no limitadas en su rama derecha contiene al menos una clase de \mathbf{F}/R_2 , precisamente la de las distribuciones tales que $\lim_{x \rightarrow \infty} \frac{1 - F_1(x)}{[1 - F_2(x)]^1} = C$.

Existirán tantos tipos de distribuciones respecto de su rama derecha como infinitésimos, es decir, infinitos. No obstante, una gran parte de las distribuciones más usuales, podrán aproximarse en dicha rama mediante uno de los cinco tipos de infinitésimos que se indican en la Tabla I, y que se corresponden con los cinco tipos fundamentales de infinitésimos, definidos para $x \rightarrow \infty$. Para cada tipo se da una familia de representantes de su correspondiente clase de \mathbf{F}/R_3 . (Para fijar uno de ellos basta fijar las constantes a y c).

TABLA I

Distribución con la rama derecha potencial exponencial	$F(x) \simeq 1 - c \cdot x^{-ax} ; c, a > 0$
Distribución con la rama derecha exponencial	$F(x) \simeq 1 - c \cdot a^{-x} ; c > 0, a > 1$
Distribución con la rama derecha potencial	$F(x) \simeq 1 - c \cdot x^{-a} ; c, a > 0$
Distribución con la rama derecha logarítmico-potencial	$F(x) \simeq 1 - c \cdot \text{Ln}^{-a} x ; c, a > 0$
Distribución con la rama derecha limitada	$F(x) = 1$

La familia de representantes de cada tipo, dada en la Tabla I, es tal que todos los representantes con un mismo valor de "a" pertenecen a la misma clase de \mathbf{F}/R_2 , es decir, son del mismo orden, y sólo hay un representante de cada clase de \mathbf{F}/R_1 . Por tanto dada una cierta distribución $F_1(x)$ de uno de los cinco tipos anteriores existe sólo un representante de la familia que es equivalente en la rama derecha a $F_1(x)$.

Con objeto de relacionar los representantes dados en la Tabla anterior con las tres distribuciones clásicas, se procede a continuación a sustituirlos por los que siguen

$$1. F_{X_{m\acute{a}x}}(x) = 1 - c x^{-ax} = \left[\left(1 - \frac{1}{\frac{x^{ax}}{c}} \right)^{\frac{x^{ax}}{c}} \right]^{cx^{-ax}} \simeq \exp(-c \cdot x^{-ax});$$

$$a, c > 0 \quad (10)$$

$$2. F_{X_{m\acute{a}x}}(x) = 1 - c \beta^{-x} = \left[\left(1 - \frac{1}{\frac{\beta^x}{c}} \right)^{\frac{\beta^x}{c}} \right]^{c \beta^{-x}} \simeq \exp(-c \beta^{-x}) =$$

$$= \exp(-\beta^{\alpha-x}) = \exp[-e^{(\alpha-x) \text{Ln} \beta}] =$$

$$= \exp[-e^{-a(x-b)}]; \quad a > 0 \quad (11)$$

$$3. F_{X_{m\acute{a}x}}(x) = 1 - c x^{-a} = \left[\left(1 - \frac{1}{\frac{x^a}{c}} \right)^{\frac{x^a}{c}} \right]^{cx^{-a}} \simeq \exp(-c x^{-a}) =$$

$$= \exp[-(b/x)^a]; \quad a, b > 0 \quad (12)$$

$$4. F_{X_{m\acute{a}x}}(x) = 1 - c \text{Ln}^{-a} x = \left[\left(1 - \frac{1}{\frac{\text{Ln}^a x}{c}} \right)^{\frac{\text{Ln}^a x}{c}} \right]^{c \text{Ln}^{-a} x} \simeq$$

$$\simeq \exp(-c \text{Ln}^{-a} x) = \exp[-c/\text{Ln}^a x] \quad a, c > 0 \quad (13)$$

Como puede comprobarse, las distribuciones 2 y 3 resultantes son precisamente las clásicas tipo I y II para máximos, también llamadas de Gumbel y Fréchet respectivamente. Ello indica que para valores grandes de x , ambos tipos I y II son insuficientes y que para poder ajustar cualquier distribución es necesario efectuar una extensión de las familias clásicas.

Las expresiones (10), (11), (12) y (13) pueden adoptar la siguiente forma doble exponencial

$$1. F(x) = \exp[-c x^{-ax}] = \exp[-\exp(-a x \operatorname{Ln} x + \operatorname{Ln} c)];$$

$$a > 0; c > 0 \quad (14)$$

$$2. F(x) = \exp[-\exp[-a(x-b)]]; \quad a > 0; b > 0 \quad (15)$$

$$3. F(x) = \exp[-(b/x)^a] = \exp[-\exp(-a \operatorname{Ln} x + a \operatorname{Ln} b)];$$

$$a > 0; b > 0 \quad (16)$$

$$4. F(x) = \exp[-c/(\operatorname{Ln} x)^a] = \exp[-\exp(-(a \operatorname{Ln}(\operatorname{Ln} x) - \operatorname{Ln} c))];$$

$$a > 0; c > 0 \quad (17)$$

Las cuatro expresiones anteriores pertenecen a la familia tetra-paramétrica dada por

$$G(x; a, b, \beta, \gamma) = \exp[-\exp(-a \operatorname{Ln}^{(\beta)}(x) x^\gamma + b)]$$

$$a > 0, \beta \in \mathbb{N}, \gamma \geq 0, \gamma + \beta > 0, x > x_0; \quad (18)$$

donde $x_0 > 0$ depende de β y γ , siendo $\operatorname{Ln}^{(\beta)}(x) = \operatorname{Ln}(\operatorname{Ln}(\dots^{(\beta)} \operatorname{Ln}(x)\dots))$.

A continuación se enuncian varios resultados de interés para la interpretación y evaluación del significado físico de los parámetros a, b, γ y β . Las demostraciones se pueden encontrar en Moreno (13).

El siguiente teorema caracteriza el orden parcial estricto definido en el conjunto \mathbf{F}/R_2 mediante (7), permitiendo distinguir la ordenación relativa en la cola derecha de dos distribuciones $G(x; a_1, b_1, \beta_1, \gamma_2)$ y $G(x; a_2, b_2, \beta_2, \gamma_2)$ de la familia (18).

TEOREMA IV.— La distribución $G(x; a_1, b_1, \beta_1, \gamma_1)$ es de orden inferior en la rama derecha a la $G(x; a_2, b_2, \beta_2, \gamma_2)$ si se satisface cualquiera de las condiciones siguientes:

$$\gamma_2 > \gamma_1 \quad (a)$$

$$\text{ó } \gamma_1 = \gamma_2, \beta_2 < \beta_1, \quad (b) \quad (19)$$

$$\text{ó } \gamma_1 = \gamma_2, \beta_1 = \beta_2, a_2 > a_1 \quad (c)$$

Para relacionar la familia (18) con el conjunto \mathbf{F}/R_1 , e identificar los miembros de dicha familia que son equivalentes en la rama derecha, se pueden utilizar el siguiente teorema y corolario.

TEOREMA V.— La familia definida por (18) sólo posee un representante en cada clase de \mathbf{F}/R_1 , si bien no posee representantes de todas las clases.

COROLARIO II.— Dos distribuciones de la familia (18) son equivalentes en la rama derecha si y sólo si las cuaternas de parámetros asociadas son idénticas.

Con objeto de caracterizar las distribuciones de la familia (18) atendiendo a su orden y tipo en la rama derecha, se dan los dos siguientes teoremas:

TEOREMA VI.— Dos distribuciones de la familia (18) son del mismo orden en la rama derecha si y sólo si son idénticos el primero y los dos últimos elementos de sus respectivas cuaternas.

TEOREMA VII.— Dos distribuciones de la familia (18) son del mismo tipo en la rama derecha si y sólo si son idénticos los dos últimos elementos de sus respectivas cuaternas.

La familia (18) se puede utilizar para aproximar distribuciones en la rama derecha. No obstante pueden darse algunos problemas, como no verificarse las condiciones de función de distribución para $x < x_0$ o la posible indefinición de (18) para $x < 0$. Por ello, es conveniente considerar, en lugar de la familia (18) la siguiente:

$$U(x; a, b, \gamma) = \exp[-\exp(-a|x|^\gamma \operatorname{sig}(x) + b)];$$

$$a > 0, \gamma > 0, x \in \mathbb{R} \quad (20)$$

Si la variable aleatoria considerada sólo está definida para $x > x_0 \geq 0$, también tiene interés considerar la familia

$$V(x; a, b, \beta) = \exp[-\exp(-a \operatorname{Ln}^{(\beta)}(x) + b)];$$

$$a > 0, \beta \in \mathbb{N}, x > x_0 \quad (21)$$

donde x_0 viene dado por

$$\operatorname{Ln}(\operatorname{Ln}^{(\beta)} \dots (\operatorname{Ln} x_0) \dots) = -\infty$$

es decir

$$x_0 = \exp [\exp^{(\beta)} \dots \exp (-\infty)] \dots] \quad (22)$$

Ambas familias $U(x; a, b, \gamma)$ y $V(x; a, b, \beta)$ son subfamilias de la familia (18) y, además, satisfacen todas las condiciones de función de distribución para cualquier x .

3.2. Aproximación en la rama derecha de distribuciones limitadas

Para obtener una familia válida para aproximar una distribución limitada en la rama derecha en $x = w_2$ en un entorno de su límite, a partir de las familias dadas, se ha de utilizar una transformación que haga corresponder al entorno del infinito dicho entorno. Si se hace uso de la familia (20), válida para todo x perteneciente a \mathbb{R} , en el caso de que la distribución inicial esté limitada por la izquierda en $x = w_1$, habrá que utilizar una transformación que haga corresponder el entorno de w_1 el entorno de $-\infty$. Además para obtener una función de distribución, la derivada de dicha transformación debe ser positiva.

La Tabla II muestra varias transformaciones que permiten resolver el problema de aproximación en la rama derecha de distribuciones limitadas, para las diversas situaciones en los límites y para las familias (20) y (21).

Considerando el caso particular de la familia $V(x; a, b, \beta)$ con $\beta = 1$, se tiene $x_0 = 0$, y aplicando la transformación (1) de la Tabla II, se obtiene la siguiente distribución:

$$\begin{aligned} V(y; a, b, 1) &= \exp [-\exp (-a \operatorname{Ln} (y - w_1) + b)] = \\ &= \exp \left[-\left(\frac{k}{y - w_1} \right)^a \right] \end{aligned} \quad (23)$$

que es la distribución de Fréchet con límite inferior w_1 mayor o igual que cero.

Utilizando la misma familia y condiciones, y aplicando la transformación (2) de la Tabla II, resulta

$$V(y; a, b, 1) = \exp \left[-\exp \left(-a \operatorname{Ln} \frac{1}{w_2 - y} + b \right) \right] = \exp \left[-\left(\frac{w_2 - y}{k} \right)^a \right] \quad (24)$$

TABLA II

$$U(x; a, b, \gamma) = \exp[-\exp[-a|x|^\gamma \operatorname{sig}(x) + b]]$$

$w_1 \leq x \leq w_2$	$(w_1, w_2) \rightarrow (-\infty, \infty)$	$y = \frac{1}{w_1 - x} + \frac{1}{w_2 - x}$
$-w_2 \leq x \leq w_2$	$(-w_2, w_2) \rightarrow (-\infty, \infty)$	$y = \frac{1}{w_2 - x} - \frac{1}{w_2 + x}$
		$y = \operatorname{tg} \frac{\pi \cdot x}{2 \cdot w_2}$
$-\infty \leq x \leq w_2$	$(-\infty, w_2) \rightarrow (-\infty, \infty)$	$y = x + \frac{1}{w_2 - x}$
$w_1 \leq x \leq \infty$	$(w_1, \infty) \rightarrow (-\infty, \infty)$	$y = x + \frac{1}{w_1 - x}$

$$V(x; a, b, \beta) = \exp[-\exp[-a \operatorname{Ln}^\beta(x) + b]]$$

$x_0 + w_1 \leq x \leq \infty$	$(x_0 + w_1, \infty) \rightarrow (x_0, \infty)$	$y = x - w_1$ (1)
$-\infty < x \leq w_2$	$(-\infty, w_2) \rightarrow (x_0, \infty)$	$y = x_0 + \frac{1}{w_2 - x}$ (2)

que es la expresión de la tercera asíntota.

De este modo queda demostrado que las tres distribuciones clásicas de extremos son casos particulares de las mismas distribuciones aproximantes en la rama derecha aquí presentadas.

Si se considera la familia (20) con las transformaciones de la Tabla II, se obtiene la Tabla III.1, que se recomienda para ajuste de distribuciones en la rama derecha.

TABLA III.1

Funciones de distribución de valores extremos en la cola derecha

$-\infty < x < \infty$	$\exp [-\exp (-a x ^\gamma \operatorname{sig}(x) + b)]$
$w_1 \leq x < \infty$	$\exp \left[-\exp \left(-a \left x + \frac{1}{w_1 - x} \right ^\gamma \operatorname{sig} \left(x + \frac{1}{w_1 - x} \right) + b \right) \right]$
$-\infty < x \leq w_2$	$\exp \left[-\exp \left(-a \left x + \frac{1}{w_2 - x} \right ^\gamma \operatorname{sig} \left(x + \frac{1}{w_2 - x} \right) + b \right) \right]$
$-w \leq x \leq w$	$\exp \left[-\exp \left(-a \left \frac{x}{w^2 - x^2} \right ^\gamma \operatorname{sig} \left(\frac{x}{w^2 - x^2} \right) + b \right) \right]$
	$\exp \left[-\exp \left(-a \left \operatorname{tg} \frac{\pi x}{2w} \right ^\gamma \operatorname{sig} \left(\operatorname{tg} \frac{\pi x}{2w} \right) + b \right) \right]$
$w_1 \leq x \leq w_2$	$\exp \left[-\exp \left(-a \left \frac{1}{w_2 - x} + \frac{1}{w_1 - x} \right ^\gamma \operatorname{sig} \left(\frac{1}{w_2 - x} + \frac{1}{w_1 - x} \right) + b \right) \right]$

4. APROXIMACION DE DISTRIBUCIONES EN LA RAMA IZQUIERDA

Para resolver el problema de aproximar funciones de distribución en la rama izquierda, basta efectuar las siguientes transformaciones

$$\begin{aligned}
 X &= -x \\
 W_i &= -w_i \\
 F(X) &= 1 - F(x)
 \end{aligned}
 \tag{25}$$

en las correspondientes distribuciones para aproximar en la rama derecha. Resulta así la Tabla III.2, que se recomienda para ajustar funciones de distribución en la rama izquierda.

TABLA III.2

Funciones de distribución de valores extremos en la rama izquierda

$-\infty < x < \infty$	$1 - \exp [-\exp (a x ^\gamma \operatorname{sig}(x) + b)]$
$-\infty < x \leq w_2$	$1 - \exp \left[-\exp \left(a \left x + \frac{1}{w_2 - x} \right ^\gamma \operatorname{sig} \left(x + \frac{1}{w_2 - x} \right) + b \right) \right]$
$w_1 \leq x < \infty$	$1 - \exp \left[-\exp \left(a \left x + \frac{1}{w_1 - x} \right ^\gamma \operatorname{sig} \left(x + \frac{1}{w_1 - x} \right) + b \right) \right]$
$-w \leq x \leq w$	$1 - \exp \left[-\exp \left(a \left \frac{x}{w^2 - x^2} \right ^\gamma \operatorname{sig} \left(\frac{x}{w^2 - x^2} \right) + b \right) \right]$
	$1 - \exp \left[-\exp \left(a \left \operatorname{tg} \frac{\pi x}{2w} \right ^\gamma \operatorname{sig} \left(\operatorname{tg} \frac{\pi x}{2w} \right) + b \right) \right]$
$w_1 \leq x \leq w_2$	$1 - \exp \left[-\exp \left(a \left \frac{1}{w_2 - x} + \frac{1}{w_1 - x} \right ^\gamma \operatorname{sig} \left(\frac{1}{w_2 - x} + \frac{1}{w_1 - x} \right) + b \right) \right]$

5. EJEMPLO DE APLICACION

Seguidamente se presenta una aplicación de la familia (20) para ajustar datos experimentales de altura de ola significativa máxima, comparándose el resultado con el obtenido por otros investigadores.

Houmb y otros (12), publicaron las medidas de altura de ola significativa máxima obtenidas en el período 1949-1976 en Myken-Skomvaer (Noruega). Efectuaron el ajuste de los datos mediante la distribución de Gumbel, resultando la curva a trazos de la fig. 1. En dicha figura y en el eje de abscisas se han especificado simultáneamente los valores de la probabilidad acumulada F y del período de retorno p , que como es sabido están ligados por la relación $p = 1/F$.

El ajuste es bastante deficiente en la rama, resultando estimaciones del período de retorno menores que las reales, esto es, alturas de ola de cálculo mayores que las necesarias en la realidad, con el consiguiente considerable incremento en el costo de construcción de las obras marítimas.

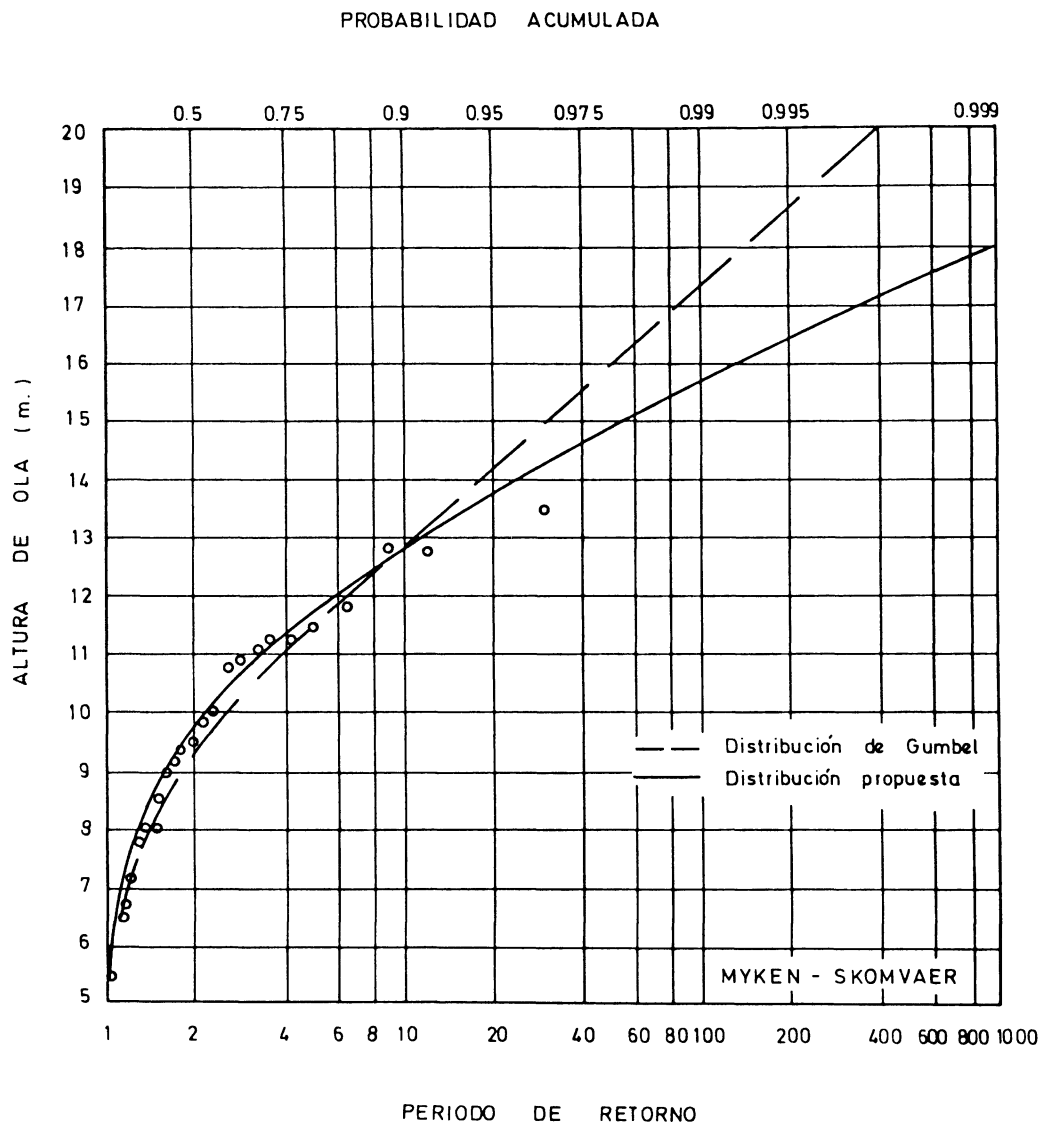


Fig. 1

Ajustando los mismos datos con una función de distribución de la familia (20) resulta la curva continua de la fig. 1. El criterio empleado en el ajuste está basado en el concepto de equivalencia en la rama derecha (Teorema I). Consiste en minimizar la expresión

$$\sum_{i=1}^n \left[\frac{1 - z_i}{1 - F(x_i)} - 1 \right]^2 = \sum_{i=1}^n \frac{1}{[1 - F(x_i)]^2} [z_i - F(x_i)]^2 \quad (26)$$

donde n es el tamaño muestral, $z_i = i/(n + 1)$ y $F(x_i)$ son los valores de la función de distribución muestral y teórica que corresponden a los datos muestrales ordenados.

Se observa claramente que esta última solución se ajusta a los datos mejor que la de Houmb, especialmente en la rama derecha, que es la zona que realmente tiene interés desde el punto de vista de proyecto. La figura muestra cómo la solución de Houmb conduce a una altura de ola de cálculo que es casi dos metros mayor que la resultante del modelo propuesto, para un período de retorno de cien años.

6. OBSERVACIONES FINALES

En el corolario I del teorema II, se ha demostrado que para una muestra finita de tamaño n , la distribución inicial y la del máximo son del mismo orden y por tanto del mismo tipo. Por ello, pensando en las aplicaciones, para aproximar la distribución del máximo de una muestra asociada a cierta distribución, será suficiente identificar el tipo en la rama derecha de la distribución inicial, para entonces ajustar parámetros adecuados de la distribución equivalente en la rama derecha a $F^n(x)$, que es la distribución del máximo.

7. CONCLUSIONES

Las conclusiones más importantes de este trabajo son:

1. La teoría expuesta clarifica el problema de la aproximación de una función de distribución en sus ramas, introduciendo los

conceptos de equivalencia, orden y tipo en una de las ramas o límites, lo que implica tres diferentes clasificaciones en el conjunto de las funciones de distribución absolutamente continuas.

2. Esta teoría muestra que la distribución del máximo de una muestra aleatoria y la distribución inicial son del mismo orden. Ello justifica el uso de todos los valores muestrales para identificar el orden de una función de distribución.
3. La misma teoría es útil para señalar que un problema práctico de extremos es con frecuencia, más un problema de ajuste en la rama o límite de interés que un problema de ajuste de una distribución de extremos, siendo por ello insuficientes las tres asíntotas I, II y III clásicas para ajustar todos los casos prácticos.
4. La familia de distribuciones presentada en el trabajo contiene a las tres familias clásicas, generalizando, por ello, la aproximación clásica. Además contiene solamente una distribución de alguna de las clases dadas del conjunto cociente \mathbf{F}/R_1 , aunque no contiene representantes de todas las clases.
5. Para el caso de distribuciones limitadas en la rama, se dan varias transformaciones que, manteniendo el rango de la variable aleatoria, transforman un entorno del límite en uno de $\infty(-\infty)$, permitiendo el análisis de este tipo de distribuciones.

8. REFERENCIAS

- (1) Battjes, J.A. (1972) - Longterm Wave Height Distribution at Seven Stations Around the British Isles. *Deutschen Hydrographischen Zeitschrift*, 25, No. 4, 179-189.
- (2) Borthiewicz, L. (1922) - Variations breite und mittlerer Fehler, Sitzungsberichte. *Berliner Math. Ges.*, 21, 2 y siguientes.
- (3) Castillo, E. (1978) - *Introducción a la Estadística Aplicada*. Ed. E. Castillo.
- (4) Copeiro, E. (1978) - *Análisis Extremal de Variables Geofísicas*. Tesis Doctoral. E.T.S. de Ingenieros de Caminos. Univ. de Santander.
- (5) Dodd, E.L. (1923) - The Greatest and the Least Variate under General Laws of Error. *Trans. Am. Math. Soc.*, 25, 525 - 539.

- (6) Fisher, R.A. & Tippett, L.H.C. (1928) - Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.*, 24, 180-190.
- (7) Fréchet, M. (1927) - Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Polonoise Math. (Cracow)* 6, 93 - 116.
- (8) Fuller, W.E. (1914) - Flood Flows. *Trans. Am. Soc. Civ. Eng.*, 77, 564 and following.
- (9) Gnedenko, B.V. (1943) - Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.*, 44, 423 - 453.
- (10) Green, R.F. (1976) - Partial Attraction of Maxima. *J. Appl. Prob.*, 13, 159-163.
- (11) Gumbel, E.J. (1958) - *Statistics of Extremes*. Columbia Univ. Press., New York.
- (12) Houmb, O.G., Mo, K. and Overvik, T. (1978) - Reliability Tests of Visual Wave Data and Estimation of Extreme Sea States. Rapport No. 5 of Port and Ocean Engineering. The University of Trondheim. Norway.
- (13) Moreno, E. (1980) - *Nuevos modelos de ajuste en las colas. Distribución de valores extremos. Aplicación a obras marítimas*. Tesis Doctoral. E.T.S. de Ingenieros de Caminos. Univ. de Santander.
- (14) Resnick, S.I. (1971) - Tail Equivalence and its Applications. *J. Appl. Prob.* 8, 135-156.
- (15) Von Mises, R. (1936) - La distribution la plus grande de n valeurs. *Rev. Math. de l'Union Interbalkanique (Athens)*, 1, 1-20.
- (16) Weibull, W. (1939) - A statistical theory of strength of materials. *Ing. Vet. Ak. Handl., No. 151*, Stockholm.
- (17) Weinstein, S.B. (1973) - Theory and Application of some Classical and Generalized Asymptotic Distributions of Extreme Values. *I.E.E. E. Transact. on Inform. Theory. II - 19*, No. 2, 148 - 154.