DEDUCCIÓN DE LA DISTRIBUCIÓN DE GUMBEL (TIPO I DE MÁXIMOS)

Sean n variables aleatorias, $Z_1, Z_2, ..., Z_n$, independientes e igualmente distribuidas, con rango $R_Z = \mathbb{R}$, y sea Y su máximo. Para n suficientemente grande, Y tomará valores (fundamentalmente) en la cola superior de Z. Supongamos que la distribución de Z cae asintóticamente en su cola superior, de forma que, para valores grandes de z,

$$F_Z(z) = 1 - e^{-g(z)}$$

siendo $g(z)$ una función creciente para valores crecientes de z. (Obsérvese que muchas distribuciones se comportan de este modo: Normal, Poisson con media grande, Binomial con parámetro n grande, Logarítmico normal, Pascal, Gamma, Exponencial, etc.)

En nuestro caso, como $Y = \max\limits_{i=1,n} \{Z_i\}$,

$$F_Y(y) = \big[F_Z(y)\big]^n = \big[1 - e^{-g(y)}\big]^n$$

Se denomina valor característico, u_n, de la variable aleatoria Y_n, a aquel tal que

$$F_Z(u_n) = 1 - \frac{1}{n}$$

Por ejemplo, si n=2, u_n es la mediana de las X_i. Obsérvese que para n grande, $1 - \frac{1}{n}$ tiende a 1, y por lo tanto u_n está en la cola superior de la distribución de Z.

En nuestro caso

$$F_Z(u_n) = 1 - e^{-g(u_n)} = 1 - \frac{1}{n}; \qquad e^{-g(u_n)} = \frac{1}{n}; \qquad 1 = \frac{1}{n} e^{g(u_n)}$$

y podemos escribir la expresión de la función de distribución acumulada de Y en la cola superior de la distribución como

$$F_Y(y) = \big[1 - e^{-g(y)}\big]^n = \left[1 - \frac{1}{n} e^{g(u_n)} e^{-g(y)}\right]^n = \left[1 - \frac{1}{n} e^{-\big(g(y) - g(u_n)\big)}\right]^n$$

Desarrollando ahora $g(y)$ en serie de Taylor alrededor de u_n, y tomando únicamente hasta el término lineal

$$g(y) = g(u_n) + (y - u_n) \left.\frac{dg(y)}{dy}\right|_{u_n} = g(u_n) + \alpha(y - u_n)$$

En esta expresión, α no depende de y, ya que la derivada está particularizada en u_n. Por tanto

$$F_Y(y) = \left[1 - \frac{1}{n} e^{-\alpha(y - u_n)}\right]^n$$

Tomando límites cuando $n \to \infty$,

$$\boxed{F_Y(y) = e^{-e^{-\alpha(y-u)}}, \quad R_Y = \mathbb{R}}$$

que es la denominada distribución de Gumbel, o Tipo I de máximos. Si derivamos para calcular la función de densidad obtenemos

$$\boxed{f_Y(y) = \alpha e^{-\alpha(y-u)-e^{-\alpha(y-u)}}, \quad R_Y = \mathbb{R}}$$

Es fácil demostrar que u es la moda de Y, luego la aproximación lineal que hemos realizado anteriormente es suficientemente precisa. El parámetro α es el inverso de una cierta medida de la dispersión, exactamente $\alpha = \frac{\pi}{\sigma\sqrt{6}}$