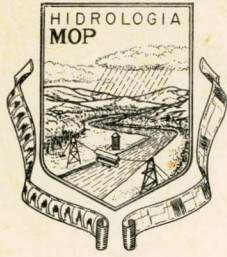


REPUBLICA DE VENEZUELA  
MINISTERIO DE OBRAS PUBLICAS  
DIRECCION DE OBRAS HIDRAULICAS



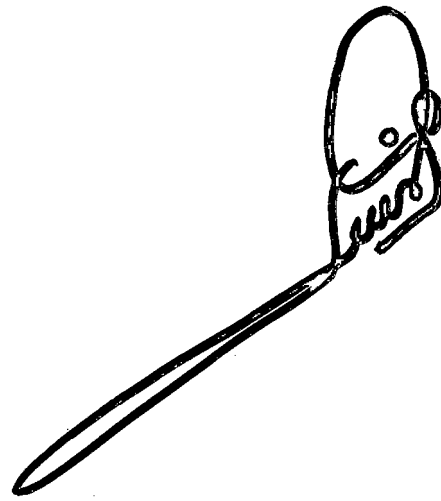
ANALISIS ESTADISTICO  
DE  
VALORES EXTREMOS

G. R. KENDALL

A handwritten signature in blue ink, appearing to be 'G. R. Kendall', written in a cursive style.

DIVISION DE HIDROLOGIA

Adaptado al Español por Hid<sup>o</sup> Gilberto Antonini  
Sección de Estudios - División de Hidrología - Año 1967.



ANALISIS ESTADISTICO  
DE  
VALORES EXTREMOS

G. R. KENDALL

Adaptado al Español por Hdr<sup>o</sup> Gilberto Antonini  
Sección de Estudios - División de Hidrología - Año 1967.

# ANALISIS ESTADISTICO DE VALORES EXTREMOS

G. R. KENDALL

## INTRODUCCION

Suponiendo por ejemplo un río en donde se tienen una serie de mediciones del caudal hechas cada día, se asume que no ha habido cambio en el río y que los datos recopilados son homogéneos. Más adelante nos interesamos en crecientes, o sean las ocasiones cuando el gasto excede a un cierto valor, para estimar por medio de observaciones la probabilidad de ocurrencia de crecientes de cierta magnitud. En general existen dos maneras de proceder:

- 1) Recopilando todas las observaciones en una tabla de frecuencia en la cual se pueda ver en que porcentaje de tiempo hay crecientes más grandes que el del especificado. Hay algunas desventajas en este método, siendo lo principal que las observaciones diarias no son independientes; si el río está alto un día, lo más probable que esté alto al día siguiente, entonces estaremos calculando la misma creciente varias veces. La otra desventaja de este método es que el número de observaciones que deben ser consideradas es de 365 veces el número de años de registro.
- 2) La otra manera de proceder es la de observar la creciente más grande de cada año y entonces tratar con esta observación para cada año. La creciente máxima anual del promedio, o sea un año aparte, será independiente en el sentido

de que un valor alto de un año no tenderá a estar asociado con un valor alto del próximo año. Este tipo de análisis entonces evita las dos serias desventajas del otro. El resultado de analizar valores extremos anuales será también un valor extremo.

### PERIODO DE RETORNO

El concepto del Período de Retorno es discutido haciendo referencia a la relación entre el riesgo de ocurrencia de un evento en un período dado y su Período de retorno.

Tratando con datos de valores extremos, como crecientes máximas anuales, la manera más conveniente de discutir la probabilidad es en forma de período de retorno. Si  $p$  es la probabilidad de ocurrencia de un valor de una creciente máxima anual, igual o mayor que un valor dado  $x$ , entonces el período de retorno  $T$  es definido por:  $(T = \frac{1}{p})$  (1)

En concreto, se habla de crecientes anuales, pero la idea es general y puede ser aplicada a cualquier variable observada en intervalos uniformes de tiempo o espacio.

Se puede notar que en la definición  $p$  es la probabilidad de que un valor de  $x$  sea igualado o excedido. Más adelante cuando tratemos con probabilidades dadas por una función distributiva se habla de la probabilidad de que cualquier varia

ble tenga un valor menor que  $x$ . Esta probabilidad se denomina  $q$ .

Usando las reglas ordinarias para el cálculo de probabilidades se pueden derivar expresiones para la probabilidad de varias combinaciones de eventos. Por conveniencia se hablará de la ocurrencia de  $x$ , pero queriendo decir, la ocurrencia de un valor de la variable aleatoria igual o mayor que  $x$ .

1° La probabilidad que  $x$  ocurra en el próximo año es por definición:

$$p = \frac{1}{T} \quad (2)$$

2° La probabilidad de que  $x$  no ocurra en el próximo año es:

$$q = 1 - p = 1 - \frac{1}{T} \quad (3)$$

3° La probabilidad de que  $x$  ocurra por primera vez en el año  $r$  es:

$$q^{r-1} p = \left(1 - \frac{1}{T}\right)^{r-1} \frac{1}{T} \quad (4)$$

4° La probabilidad de que  $x$  ocurra por lo menos una vez en los próximos  $r$  años es la suma de las probabilidades de su ocurrencia en el 1°, 2° hasta el año  $r$ , y es por lo tanto:

$$\begin{aligned} Pr &= p + pq + pq^2 + \dots + pq^{r-1} \\ &= 1 - q^r = 1 - \left(1 - \frac{1}{T}\right)^r \end{aligned} \quad (5)$$

donde  $Pr$  se introdujo solamente por comodidad. Se puede utilizar la última expresión para calcular la probabilidad de  $x$  de ocurrir dentro de su período

de retorno. Esto será:

$$P_T = 1 - \left(1 - \frac{1}{T}\right)^T \quad (6)$$

Es bién conocido que:

$$\lim_{Y \rightarrow \infty} \left(1 - \frac{1}{Y}\right)^Y = e^{-1} \quad (7)$$

de manera que para grandes valores de T el valor de  $P_T$  se aproxima a:

$$1 - e^{-1} = 0.63$$

TABLA 1

Probabilidad  $P_T$  que un evento con período de retorno T ocurra dentro de los próximos T años.

T	$P_T$	T	$P_T$
2	.7500	100	.6340
5	.6723	200	.6330
10	.6513	500	.6325
20	.6415	1000	.6323
50	.6358	inf	.6321

La tabla muestra que la probabilidad de que x ocurra dentro de su período de retorno es de cerca del 55% . Usando crecientes con un cierto período de retorno calculado como base para diseñar, este factor se debe tener en mente .

Una represa diseñada para resistir una creciente con un período de retorno de 25 años tiene un 65% de chance para que sea excedida antes del término del primer período de 25 años.

Para propósitos de diseño sería mejor especificar alguna probabilidad de que el evento indeseable ocurra dentro del período de diseño y calcular el período de retorno requerido, si U es el riesgo de que el evento ocurra dentro del período de diseño, entonces la ecuación que relaciona estas cantidades es: a partir de (5)

$$U = 1 - q^r$$

$$= 1 - \left(1 - \frac{1}{T}\right)^r \quad (8)$$

La tabla N° 2 da los valores del período de retorno T correspondientes a un número de valores del riesgo U y período de diseño r.

TABLA 2  
Período de retorno requerido T para riesgo U%  
que el evento ocurra por lo menos una vez en  
un período de r años

Riesgo U %	Período de retorno (r)							
	2	5	10	15	20	25	50	100
75	-	4,02	6,69	11,0	14,9	18,0	35,6	72,7
50	3,43	7,74	14,90	22,1	29,4	36,6	72,6	144,8
40	4,44	10,30	20,10	29,9	39,7	49,5	98,4	196,3
30	6,12	14,50	28,50	42,6	56,5	70,6	140,7	281,0
25	7,46	17,90	35,30	52,6	70,0	87,4	174,3	348,0
20	9,47	22,90	45,30	67,7	90,1	122,5	224,6	449,0
15	12,80	31,30	62,00	90,8	123,6	154,3	308,0	616,0
10	19,50	48,10	95,40	142,9	190,3	238,0	475,0	950,0
5	39,50	98,00	195,50	292,9	390,0	488,0	976,0	1949,0
2	99,50	248,00	496,00	743,0	990,0	1.238,0	2.475,0	4.950,0
1	198,00	496,00	996,00	1.492,0	1.992,0	2.488,0	4.975,0	9.953,0

Para ilustrar el uso de la tabla se supone una represa que ha sido diseñada con una vida proyectada de 25 años, pero el diseñador solo desea tener un chance del 10% para que la represa sea sobrepasada dentro de este período. Encontramos

en la tabla que para  $r = 25$  años y  $U = 10\%$  el valor 238, este es el período de retorno de la creciente que debemos utilizar en el diseño.

Una aproximación usada a la relación (8) es también:

$$T = r \left( \frac{1}{U} - \frac{1}{2} \right) \quad (9)$$

Esta regla puede ser expresada por lo siguiente: subtraer  $1/2$  del recíproco del riesgo y multiplicar el resultado por el período de diseño para obtener el período de retorno del diseño. En el ejemplo solo se da que el riesgo es de 10% ó un dé cimo. Su recíproco es por lo tanto 10, restando un medio nos queda  $9 \frac{1}{2}$ , y multi plicando este por 25 de período de diseño nos da  $237 \frac{1}{2}$ . La tabla da 238, lo cual en la práctica es lo mismo. La aproximación es muy cercana para  $r$  igual a 10 ó más y  $U$  igual al 50% ó menos.

#### ESTIMACION DE PROBABILIDADES

En la sección anterior no se ha mencionado la manera como fué estimada la probabilidad, si se tiene por ejemplo treinta crecientes máximas anuales, la mayor en el ejemplo es una estimación de la creciente de retorno de 30 años. En forma si milar, el tercer valor desde arriba dará una estimación de la creciente de retorno de 10 años. Estas estimaciones no dan el uso total de todos los valores del ejemplo y tiene un gran error comparativo.

Para dar una idea sobre la magnitud del error envuelto hay que considerar el

problema relacionado con la estimación de la probabilidad de ocurrencia de algún valor de la creciente máxima anual. Asumiendo que no hay correlación consecutiva entre sucesivos valores, esto es exactamente equivalente que encontrar la probabilidad de sacar una bola blanca de una urna que contenga una mezcla de bolas blancas y negras. Esto es un problema directo en la teoría de probabilidades y depende de la aplicación repetida de la distribución binomial. La figura N° 1 muestra gráficamente los límites del error de estimación de la probabilidad  $p$  para varios ejemplos de tamaños . Suponiendo que se tenga un ejemplo de 10 y 4 de ellos son casos favorables (bolas blancas o crecientes anuales más grandes que algunos valores, como se puede dar el caso) entonces el valor observado de  $p$  es 0.4. Note que solo los puntos del gráfico para  $n=10$  se oponen los valores 1, 2, etc., tienen algún significado. Este valor observado podría provenir de una población cuyo valor de  $p$  se encuentra en cualquier sitio entre 0.12 y 0.74 (con 95% de confianza). Para una discusión más amplia se puede hacer referencia a Fraser (1958) o Pearson y Hartley (1956).

El punto a que se refiere esto, es que, para ejemplos del tamaño que más se usan en hidrología, el 95% de los intervalos seguros para una probabilidad estimada es sorprendentemente largo. Entonces para un valor con un período de retorno observado de 10 años es probablemente 0.10 de un ejemplo de 30 podemos establecer con 95% de confianza, que la probabilidad verdadera se encuentra entre 0.02 y 0.27 que corresponden a períodos de retorno de 50 y 3.7 años respectivamente. Entonces se ve

que las estimaciones de probabilidades debe basarse sobre los métodos más eficientes que estén a mano si el intervalo de confianza no es tan extenso y los resultados no sean inútiles. Estimaciones de probabilidades y períodos de retorno basados en colocar una distribución de frecuencia tiene la ventaja que la estimación está basada en todos los miembros del ejemplo más bien que el máximo valor o tan solo dos de ellos.

### DISTRIBUCION DE VALORES EXTREMOS

Suponiendo que tenemos una variable fortuita  $x$ , con una función de probabilidad  $F(x)$  para que:

$$\text{Prob. } (x \leq x') = F(x') - \int_{-\infty}^{x'} f(x) dx \quad (10)$$

entonces, dado un ejemplo de  $r$  valores independientes de  $x$ , la probabilidad que todos sean menores que  $X_r$  es:

$$\phi(X_r) = F^r(X_r) \quad (11)$$

Esta es la probabilidad de que  $X_r$  es la más grande entre las  $r$  observadas.

Diferenciando, se obtiene la función de probabilidad densidad:

$$\phi'(X_r) = r \cdot F^{r-1}(X_r) \cdot f(X_r) \quad (12)$$

Si la forma de  $F(x)$  es conocida el problema puede ser resuelto, aunque para la mayoría de distribuciones iniciales  $F(x)$ , y para valores largos y razonables de  $r$ , el cálculo de una tabla para  $\phi$  es muy difícil. Generalmente no

se conoce la forma de la distribución inicial, pero teniendo solamente el grupo de los valores extremos podemos continuarlos. El problema entonces viene a ser el de derivar una expresión para  $\phi(x)$  para ejemplos largos, e independientes de la distribución inicial  $F(x)$ .

Fisher y Tippet fueron de los primeros en tratar con este problema, pero E.J. Gumbel ha hecho mucho en poner la teoría en una forma usable por lo cual seguiremos su trabajo.

Si tenemos  $n$  valores extremos independientes  $x_1, x_2, \dots, x_n$  cada uno de los cuales es el miembro mayor del ejemplo de  $r$  observaciones de una variable fortuita, sujeta a ciertas condiciones, entonces la probabilidad para que cualquier  $X_i$  ( $i = 1, 2, \dots, n$ ) es menor que un valor escogido  $x$  aproximándose a  $\phi(x)$  así como  $n$  y  $r$  tienden a infinito, donde:

$$\phi(x) = \exp \left\{ - \exp \left[ - a (x - u) \right] \right\} \quad (13)$$

Las condiciones en la distribución inicial equivale a decir que la variable fortuita  $x$  en su forma original debe ser continua y sin límites, y que su función de frecuencia  $X(x)$  se aproxima a cero con valores grandes de  $x$  por lo menos hasta  $e^{-x}$ , con una condición adicional:

$$\frac{d}{dx} \frac{1 - W(x)}{w(x)} = 0 \quad (14)$$

que se encuentra en casi todas las distribuciones estadísticas. Actualmente la distribución limitativa (13) se aproxima más bien lentamente cuando (n) y (r) aumentan.

Gumbel usa un método de mínimos cuadrados para estimar los parámetros (x) y (u) del ejemplo, por lo que adquiriendo un mejor medio para los valores del parámetro para infinitos ejemplos ha sido usado. Se puede notar aquí que condiciones diferentes en la distribución inicial conduce a diferentes distribuciones limitativas. Por ejemplo, si se asume:

$$\lim_{u \rightarrow \infty} au = d \quad k > 0 \quad (15)$$

por lo que llegamos a la distribución limitativa:

$$F(x) = \exp\left(-\frac{u}{x}\right)^k \quad (16)$$

como fué primeramente demostrado por Frechet.

La distribución de Gumbel puede ser normalizada con la introducción de la variable reducida.

$$y = a(x - u) \quad (17)$$

lo cual reduce la distribución a:

$$\phi(y) = \exp(-e^{-y}) \quad (18)$$

La analogía con la distribución normal standard es muy cercana. Para la última distribución tenemos:

$$z = (x - m) / \sigma \quad (19)$$

donde  $m$  es el valor medio de la variable fortuita  $x$ , y  $y$  es la distribución normal.

La variable normalizada  $z$  tiene la distribución:

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{x^2}{2}} dz \quad (20)$$

la cual es libre de los parámetros  $m$  y  $\sigma$ , y por lo tanto puede ser tabulada para varios valores de  $z$  y  $F$  como, por ejemplo, El Bureau of Standards table (National Bureau of Standards AMS 23, 1953).

En una manera exactamente similar la función  $\exp(-e^{-y})$  puede ser tabulada. (National Bureau of Standards AMS 22, 1953).

Como en el caso normal hay parámetros de situación y escala. En este caso el parámetro de situación  $u$  es el módulo, mientras que el parámetro escalar ( $a$ ) está relacionado estrechamente con el recíproco de la distribución normal.

En términos del ejemplo las estimaciones son:

$$\frac{1}{a} = \frac{S_x}{\sigma n} \quad u = \bar{x} - \bar{y}_n/a \quad (21)$$

donde  $\bar{x}$  y  $S_x$  son la media y la desviación normal de la máxima anual observada,

definida por:

$$\bar{x} = \frac{\sum_i x_i}{n} \quad S_x^2 = \frac{\sum_i (x_i - \bar{x})^2}{n} \quad (22)$$

$\bar{y}_n$  y  $\sigma_n$  son parámetros, llamados la media reducida y la desviación normal que de

penden solamente del tamaño de la muestra. Algunos valores son dados en la tabla 3, adoptados de Gumbel y Court. Los primeros cinco valores fueron calculados por D. W. Boyd de la División de Investigación de Construcción y están hasta ahora sin publicar. El uso de ejemplos de menos de 15 no es recomendado.

TABLA 3  
Media reducida  $\bar{y}_n$

	0	1	2	3	4	5	6	7	8	9
10	.4952	.4996	.5035	.5070	.5100	.5128	.5157	.5181	.5202	.5220
20	.5236	.5252	.5268	.5283	.5296	.5309	.5320	.5332	.5343	.5353
30	.5362	.5371	.5380	.5388	.5396	.5402	.5410	.5418	.5424	.5430
40	.5436	.5442	.5448	.5453	.5458	.5463	.5468	.5473	.5477	.5481
50	.5485	.5489	.5493	.5497	.5501	.5504	.5508	.5511	.5515	.5518
60	.5521	.5524	.5527	.5530	.5533	.5535	.5538	.5540	.5543	.5545
70	.5548	.5550	.5552	.5555	.5557	.5559	.5561	.5563	.5565	.5567
80	.5569	.5570	.5572	.5574	.5576	.5578	.5580	.5581	.5583	.5585
90	.5586	.5587	.5589	.5591	.5592	.5593	.5595	.5596	.5598	.5599
100	.5600									

Desviación típica reducida  $\sigma_n$

	0	1	2	3	4	5	6	7	8	9
10	0.9496	0.9676	0.9833	0.9971	1.0095	1.0206	1.0316	1.0411	1.0493	1.0565
20	1.0628	1.0696	1.0754	1.0811	1.0864	1.0915	1.0961	1.1004	1.1047	1.1086
30	1.1124	1.1159	1.1193	1.1226	1.1255	1.1285	1.1313	1.1339	1.1363	1.1388
40	1.1413	1.1436	1.1458	1.1480	1.1499	1.1519	1.1538	1.1557	1.1574	1.1590
50	1.1607	1.1623	1.1638	1.1658	1.1667	1.1681	1.1696	1.1708	1.1721	1.1734
60	1.1747	1.1759	1.1770	1.1782	1.1793	1.1803	1.1814	1.1824	1.1854	1.1844
70	1.1854	1.1863	1.1873	1.1881	1.1890	1.1898	1.1906	1.1915	1.1921	1.1930
80	1.1938	1.1945	1.1953	1.1959	1.1967	1.1973	1.1980	1.1987	1.1994	1.2001
90	1.2007	1.2013	1.2020	1.2026	1.2032	1.2038	1.2044	1.2049	1.2055	1.2066
100	1.2065									

### PAPEL DE VALORES DE PROBABILIDAD EXTREMA

La manera más sencilla de ajustar valores extremos a la distribución de Gumbel es por medio del papel de valores de probabilidades extrema. Este tiene

coordenadas  $x$  e  $y$ , en donde por conveniencia  $x$  se toma como la ordenada. Si  $x$  está distribuida de acuerdo a la ecuación (13), entonces la ecuación es una línea recta.

$$y = a(x - u) \quad (23)$$

Resolviendo por  $(x)$  é introduciendo los valores de  $(a)$  y  $(u)$  de (21)

$$x = \bar{x} + (y - y_n) S_x / \sigma_n \quad (24)$$

Dado un ejemplo de  $n$  observaciones, estos se acomodan en orden ascendente y se plotean los puntos  $(x_i, y_i)$ , en donde  $(i)$  va de uno a  $(n)$ ,  $x_i$  es la enésima observación a partir desde abajo, e  $Y_i$  es definida por:

$$\phi(y_i) = i/(n + 1) \quad (25)$$

Esto parece difícil, pero si el papel de probabilidades se indica con términos de probabilidades, esto simplemente significa que ploteamos el enésimo valor en la probabilidad  $i/(n + 1)$ . Si el papel está indicado en términos del período de retorno, entonces usamos la relación:

$$T = \frac{1}{1 - \phi} \quad (26)$$

para convertir probabilidades en períodos de retorno.

Cuando se va de probabilidades a período de retorno se presentan algu

nas confusiones cuando la probabilidad es menos de un medio. En la ecuación (26)  $\phi$  puede variar desde cero a uno, mientras que los valores correspondientes a T son uno é infinito. La tabla 4 muestra los valores del período de retorno que corresponden a varias probabilidades.

TABLA 4

Períodos de retornos que corresponden a la probabilidad en porcentaje

	0	1	2	3	4	5	6	7	8	9
00	1.000	1.010	1.020	1.031	1.042	1.053	1.064	1.075	1.087	1.099
10	1.111	1.124	1.136	1.149	1.163	1.177	1.191	1.205	1.220	1.235
20	1.250	1.266	1.282	1.299	1.316	1.333	1.351	1.370	1.389	1.409
30	1.429	1.449	1.471	1.493	1.515	1.539	1.563	1.587	1.613	1.639
40	1.667	1.695	1.724	1.754	1.786	1.818	1.852	1.887	1.923	1.961
50	2.000	2.041	2.083	2.128	2.174	2.222	2.273	2.326	2.381	2.439
60	2.500	2.564	2.632	2.703	2.778	2.857	2.941	3.030	3.125	3.226
70	3.333	3.448	3.571	3.704	3.846	4.000	4.167	4.348	4.546	4.762
80	5.000	5.263	5.556	5.882	6.250	6.667	7.143	7.692	8.333	9.091
90	10.000	11.110	12.500	14.290	16.670	20.000	25.000	33.333	50.000	100.000

Para ilustrar, se supone que se tiene a 19 como ejemplo, entonces la primera observación será ploteada con una probabilidad de  $1/20$ , la segunda con  $2/20$  y así continúa hasta el valor ploteado más alto de  $19/20$ . En términos de período de retorno los puntos deben ser ploteados en 1,05, 1,11....., 20,0 como puede ser verificado con la tabla 4.

Por el uso de este papel los valores pueden ser analizados rápidamente sin

la necesidad de hacer largos cálculos. La posición a plotear ha sido leída en una regla de cálculo y la mejor línea recta fijada a ojo.

El fijar la línea recta a ojo es un procedimiento aproximado, por lo que se pierden algunas informaciones de las observaciones. Para conseguir la mejor fijación, la ecuación de la línea debe ser computada de las ecuaciones 21 y 24.

### FACTORES DE FRECUENCIA

En muchos casos deseamos calcular las estimaciones mejores posibles para uno o más períodos de retornos específicos. Esto puede hacerse sin trazar una línea en el papel de probabilidades haciendo uso de factores de frecuencia.

Expresando la ecuación (24) de la siguiente manera:

$$X_t = \bar{X} + K(T, n) \cdot S_x \quad (27)$$

donde  $X_t$  se refiere al valor de  $X$  con un período de retorno dado, y  $K(T, n)$  es el factor de frecuencia, definido por:

$$K(T, n) = \frac{Y_t - \bar{y}_n}{\sigma_n} \quad (28)$$

$Y_t$  es el valor de  $y$  cuando el período de retorno es  $T$ . Esto puede ser rápidamente visto en el Bureau of Standards Tables (1953).  $\bar{y}_n$  y  $\sigma_n$  son funciones de  $n$  solamente y se pueden sacar de la tabla 3. Así como el factor  $K$  el cuál es independiente de los valores conseguidos y puede ser calculado antes para adelantar. La

determinación de una  $X_f$  dado de una cierta muestra se limita a encontrar la muestra normal, el coeficiente de variación y la desviación normal y sustituirlos en la ecuación (27). Un número de valores del factor de frecuencia K son dados en la tabla 5.

TABLA 5

Tamaño de la muestra n	Factor de Frecuencia K (T, n)										
	5	10	15	20	25	30	50	60	75	100	1.000
15	0.967	1.703	2.117	2.410	2.632	2.823	3.321	3.501	3.721	4.005	6.265
20	0.919	1.625	2.023	2.302	2.517	2.690	3.179	3.352	3.563	3.836	6.006
25	0.888	1.575	1.963	2.235	2.444	2.614	3.088	3.257	3.463	3.729	5.842
30	0.866	1.541	1.922	2.188	2.393	2.560	3.026	3.191	3.393	3.653	5.727
35	0.851	1.516	1.891	2.152	2.354	2.520	2.979	3.142	3.341	3.598	
40	0.838	1.495	1.866	2.126	2.326	2.489	2.943	3.104	3.301	3.554	5.576
45	0.829	1.478	1.847	2.104	2.303	2.464	2.913	3.078	3.268	3.520	
50	0.820	1.466	1.831	2.086	2.283	2.443	2.889	3.048	3.241	3.491	5.478
55	0.813	1.455	1.818	2.071	2.267	2.426	2.869	3.027	3.219	3.467	
60	0.807	1.446	1.806	2.059	2.253	2.411	2.852	3.008	3.200	3.440	
65	0.801	1.437	1.796	2.048	2.241	2.398	2.837	2.992	3.183	3.429	
70	0.797	1.430	1.788	2.038	2.230	2.387	2.824	2.979	3.169	3.413	5.359
75	0.792	1.423	1.780	2.029	2.220	2.377	2.812	2.967	3.155	3.400	
80	0.788	1.417	1.773	2.020	2.212	2.368	2.802	2.956	3.145	3.387	
85	0.785	1.413	1.767	2.013	2.205	2.361	2.793	2.946	3.135	3.376	
90	0.782	1.409	1.762	2.007	2.198	2.353	2.785	2.938	3.125	3.367	
95	0.780	1.405	1.757	2.002	2.193	2.347	2.777	2.930	3.116	3.357	
100	0.779	1.401	1.752	1.998	2.187	2.341	2.770	2.922	3.109	3.349	5.261

Como un ejemplo, supongamos que tenemos un ejemplo de 25 valores extremos anuales con una media en unidades favorables de 42 y desviación típica de

6.5. Para un período de retorno de 25 años la tabla de un valor para K de 2,614.

El valor requerido con 30 años de retorno es entonces:

$$42 + (6,5) (2,614) = 59.0$$

### NOMOGRAMA DE WEISS

Weiss (1955) ideó un nomograma conveniente para encontrar gráficamente la solución de la ecuación (27); dados los valores de la media y la desviación típica del ejemplo o sean  $\bar{x}$  y  $S_x$ . Para tener un concepto más completo se debe hacer referencia al papel original de Weiss.

Para usar el nomograma, primero encontramos el valor requerido de T en el margen de la izquierda. De la intersección de la línea horizontal a través de T con la línea inclinada a través del valor apropiado de n trazamos una línea vertical para intersectar la línea inclinada con el valor requerido de  $S_x$  o desviación normal de la muestra. Una línea horizontal de la intersección encontrada cortará el margen derecho del diagrama en el valor correcto de D, la cantidad a ser añadida a la media para que de el extremo requerido. Si la desviación típica no está dentro del intervalo del diagrama ésta puede ser dividida por un factor conveniente escogido. En este caso el valor de D tiene que ser multiplicado por el mismo factor anteriormente usado.

Como un ejemplo, supongamos que la media era 90.5, la desviación típica

ca 3.6 y n es igual a 30, y deseamos el período de retorno para 10 años. Nos mo vemos horizontalmente a lo largo de la línea de período de retorno para 10 años hasta su intersección con la línea inclinada para  $n = 30$ , trazamos una vertical a la línea para  $S_r = .90$ , usando un cuarto del valor dado de manera de llegar a los límites del diagrama. Moviéndose horizontalmente a partir de esta intersección llegamos al valor de 1.37 para D en el margen derecho del diagrama. Como la desviación típica fué dividida por cuatro, el valor de D que se acaba de encontrar debe ser multiplicado por cuatro, dándonos 5.48. El valor extremo requerido es entonces  $90.5 \text{ más } 5.48 = 95.98 \text{ ó } 96.0$ . El factor K correspondiente en la tabla 5 es 1.54, multiplicando éste por la desviación típica 3.6 resulta 5.55, dando un valor extremo de 96.05, lo cual muestra una razonable y buena relación entre los dos métodos.

### INTERVALOS DE CONFIANZA

Estimando la creciente con un período de retorno dado hay dos fuentes principales de errores en la estimación, además de los errores de los datos usados: 1º podemos estar usando el tipo de curva equivocado, la guía para esto es la colo cación de las observaciones, pero por otro lado no se puede decir mucho acerca de ésto. 2º pueden haber errores debido al muestreo.

En aplicaciones hidrológicas, es frecuentemente necesario usar registros cortos los cuales incrementan la incertidumbre de los valores calculados. Es por

lo tanto necesario asignar límites dentro de los cuales los valores calculados pueden decirse que se encuentran con cierta probabilidad. El siguiente método ha sido desarrollado por Z. Kaczmarek (1957) y tiene la ventaja bastante cómoda de usarse, dadas las tablas básicas. Sin embargo, a solo la distribución de Gumbel se le hace referencia hasta aquí, el papel de Kaczmarek también trata la distribución normal logarítmica y una forma incompleta de distribución gamma usada por Foster. Nosotros señalamos con  $X_t$  la creciente máxima anual con un período de retorno  $T$ . Entonces de acuerdo al método de Gumbel tenemos:

$$X_T = x + K S_x \quad (27a)$$

$$K(T, n) = \frac{Y_T - \bar{Y}_n}{\sigma_n} \quad (28a)$$

$$Y_T = \ln \left[ -\left( \ln \left( 1 - \frac{1}{T} \right) \right) \right] \quad (29)$$

Para  $T$  y  $n$  fijos,  $X_t$  es una función de los valores observados. Si consideramos todos los posibles muestreos de  $n$  el valor de  $X_t$  dado por (27) variará de muestra en muestra. Esta será por lo tanto una variable fortuita.

Esto es una teorema de matemática estadísticas (Cramer 1946) que la distribución de una función de momentos como muestra es asintóticamente normal, esto es que se aproxima más a la normal cuanto más aumenta el tamaño de la muestra.

Para valores suficientemente largos de  $n$ , Kaczmarek ha demostrado que  $X_t$  está distribuido normalmente con la desviación típica.

$$\sigma_H = \int (T) \frac{S_x}{\sqrt{n}} \quad (30)$$

donde  $\int (T)^2 = (1 + 1,140 K + 1,100 K^2)$  (31)

y  $K$  ha sido previamente definido en (26). La tabla de  $\int (T)$  ha sido recalculada y los valores difieren ligeramente de aquellos dados por Kaczmarek. Las diferencias entre los dos juegos de valores son del orden del 2%, siendo los más pequeños los dados en la tabla 6. Esto es de poca importancia práctica de su aplicación.

TABLA 6  
Valores de  $\int (T, n)$

$n/T$	10	20	25	30	50	75	100
15	2.476	3.233	3.409	3.604	4.113	4.525	4.818
20	2.400	3.075	3.292	3.468	3.968	4.362	4.643
25	2.350	3.007	3.218	3.391	3.874	4.259	4.533
30	2.317	2.960	3.166	3.336	3.811	4.187	4.455
40	2.272	2.898	3.099	3.264	3.725	4.093	4.353
50	2.244	2.857	3.056	3.217	3.671	4.031	4.288
60	2.224	2.830	3.025	3.185	3.633	3.989	4.242
75	2.201	2.800	2.976	3.150	3.592	3.943	4.194
100	2.181	2.769	2.959	3.114	3.549	3.896	4.142

Al usar esta teoría para conseguir una región en la que se tenga una

confianza aproximada para la estimación de  $X_t$  nos referimos primero a las tablas para la distribución típica. De esto se concluye, por ejemplo, que el 95% de los valores de un universo estadístico distribuido normalmente cae alrededor de 1.96 de la desviación típica del universo estadístico medio. Por lo tanto una variable dada, escogida al azar en un universo estadístico normal caerá en el intervalo de  $(\text{media} - 1,96 \sigma_H)$  a  $(\text{media} + 1.96 \sigma_H)$ , 95 veces fuera de 100. Este argumento puede ser acomodado diciendo que si un valor es escogido al azar en tal universo estadístico, entonces el 95% de las veces se convertirá alrededor de 1.96 de la desviación típica de la media. Para el valor estimado  $X_t$ , se puede decir entonces que la probabilidad es 95% del verdadero valor medio de  $X_t$  denominada con  $E(X_t)$ , el cual cae alrededor de 1.96  $\sigma_H$  por encima o por debajo. Obviamente una exposición similar podría hacerse para cualquier otro nivel de confianza, además de 95%.

La tabla 7 nos muestra valores del factor  $t$  para varios niveles de confianza. Dados estos valores podemos asegurar que:

$$\text{Probabilidad } (-t \sigma_H < X_t - E(X_t) < t \sigma_H) = \alpha \quad (32)$$

TABLA 7

Valores de la variación típica normal para la probabilidad  $\alpha$   
(ver a (32))

$\alpha$	0.500	0.683	0.800	0.900	0.950	0.980	0.990
$t$	0.674	1.000	1.282	1.645	1.960	2.326	2.576

## "OTROS METODOS DE ANALISIS"

Debido a la falta de espacio no se pudo detallar las discusiones de los muchos diferentes métodos que han sido usados para analizar valores extremos, pero algunos de los más importantes deben ser mencionados.

Se notó rápidamente que la distribución de crecientes anuales no se acoplaba con la curva normal debido a la imprecisión presentada. Introduciendo una transformación logarítmica, buena parte de esta imprecisión puede ser eliminada. El uso de una transformación normalizada es por cierto un legítimo y útil artificio de estadística. Una buena exposición de la resultante distribución log-normal ha sido dada por Ven Te Chow (1955). Una dificultad que se presenta es que la distribución log-normal de Ven Te Chow tiene tres parámetros. Ya que la ecuación de una línea recta tiene básicamente solo dos parámetros, una distribución con tres parámetros no puede ser ploteada como una línea recta en el papel de probabilidades excepto para valores especiales de los parámetros. Debido a que una línea recta es la más fácil de todas las curvas en tratar, y como el ploteo en el papel especial de probabilidades es tan conveniente como instrumento, este solo debe ser abandonado por buenas razones.

Actualmente, no parece haber ninguna razón teórica por lo que los valores extremos deban seguir una ley logarítmica normal. En realidad el trabajo de Fisher y Tippet deduce que para muestras muy largas, la distribución de los extremos debe ser

uno de los tres tipos:

$$Y = e^{-e^{-x}} \quad (33)$$

$$Y = 1 - e^{-x^{-k}} \quad (34)$$

$$Y = 1 - e^{-x^k} \quad (35)$$

siendo los dos últimos iguales excepto por los signos de  $(x)$  y de  $(K)$ . Para que la ecuación (34) sea una función de frecuencia debemos tener  $X$  y  $K$  mayores que cero. Así que esta variable está ligada debajo. Similarmente en (35),  $x$  debe ser mayor que cero, para que esta variable se encuentre en el segundo exponente.

La segunda forma contiene tres parámetros, lo que significa que no puede ser rápidamente representada por una gráfica de línea recta en el papel de probabilidades excepto para un valor ya fijado de uno de los parámetros.

El método de Foster (Foster 1924) depende en la colocación de los tres parámetros por el método de los momentos. Esto tiene la desventaja que envuelve tener que usar el tercer momento de la muestra el cuál está sujeto a grandes fluctuaciones de muestreo. De acuerdo con el análisis de Kaczmarek para una cantidad de imprecisión, la eficiencia es del mismo orden de magnitud como en el método de Gumbel, así como la imprecisión aumente la desviación típica de la estimación de los valores con período de retorno dado llega a ser rápidamente más grande pero esto es debido a la distribución más bien que al método. En estos casos presumiblemente la distri-

bución de Gumbel no podría dar un buen acomodo a las observaciones.

H.C.S. Thom ha introducido un método muy interesante en el cuál él asume que el límite más bajo de la distribución (34) es cero (Thom 1954). Esto deja solo dos parámetros, y por eso un papel de probabilidad puede ser usado con los valores observados situados a lo largo de una línea recta. El estipula un método para estimar los valores de los parámetros basado en semejanza máxima, lo cual aunque en vuelve gran tiempo en cálculos y extrae la máxima cantidad de información de las observaciones.

El método ideado por Jenkinson tiene la ventaja que todos los tres parámetros son estimados de los valores de la muestra. Sin embargo, se cree que el método de estimación es probablemente ineficiente ya que no está basado en semejanza máxi ma, aunque esto no ha sido demostrado.

Un método de análisis basado en el modelo de tres parámetros y semejanza máxima dará la mejor estimación en el sentido de dar la menor variación de muestreo. Sin embargo, la complejidad de la ecuación de la máxima semejanza haría que el método sea muy laborioso para su aplicación.

Mientras tanto el método de Gumbel el cuál tiene alguna base teórica parece encajar las observaciones de datos de crecientes razonablemente bien y es facil para usar.

Como se señaló para Boyd y Kendall (1956) si el análisis de valores extremos

es el trazado de una curva a través de los puntos ploteados en el papel de los valores extremos de Gumbel, la distribución de Gumbel aparece como una línea recta a través de los puntos, la distribución de (34) y (35) aparece como líneas con la curvatura hacia arriba o hacia abajo dependiendo en el valor del parámetro  $K$ . Así como estamos relacionados con períodos de retorno dentro del intervalo de los valores observados, podemos tener bastante confianza en que nuestros análisis darán resultados razonables. Como extrapolamos más allá de los puntos observados los tipos diferentes de curvas divergen más y más rápidamente y los resultados llegan a ser más dudosos. Allí no parece haber alguna manera de estimar el tamaño de los errores que salen de usar la curva equivocada en nuestro análisis, excepto por supuesto en períodos de retorno dentro de los valores ploteados donde la colocación de los puntos es una guía.

## APENDICE A

### DEDUCCION DE LA DISTRIBUCION DE GUMBEL

La moda de la distribución de  $x_n$  dada en (12) puede ser encontrada diferenciando (12) y haciendo el resultado igual a cero. El valor de  $x_n$  que satisface la ecuación resultante es la abscisa del punto más alto de la curva de distribución, por lo tanto la moda, o el valor más probable de  $x_n$ .

Eliminando la  $n$  mencionada por conveniencia, la ecuación nos conduce

a:

$$\frac{n-1}{F(x)} f(x) + \frac{f'(x)}{f(x)} = 0 \quad (36)$$

Definimos dos parámetros con las ecuaciones:

$$F(u) = 1 - \frac{1}{n} \quad (37)$$

$$\alpha = n f(u) \quad (38)$$

haciendo notar que  $(u)$  y  $(\alpha)$  son funciones de  $n$  y los parámetros de la función de distribución.

Si  $F(x)$  es del tipo exponencial entonces esta converge hacia la unidad con incremento de la  $x$ , por lo menos tan rápido como  $e^{-(x)}$ . Para grandes valores de  $x$  podemos decir, de acuerdo a la regla de 1ª Hospital

$$\frac{f(x)}{1 - F(x)} \sim - \frac{f'(x)}{f(x)} \quad (39)$$

en (36). Dividiendo por  $f(x)$ , el cuál es diferente de cero encontramos que:

$$n = \frac{1}{1 - F(x)} \quad (40)$$

que se satisface con  $x = u$  debido a (37), por lo que  $u$  es el modo.

Ahora si expandemos  $F(x)$  en la vecindad de  $u$  en las series de Taylor tenemos:

$$\begin{aligned} F(x) &= F(u) + (x - u) f(u) + \frac{(x - u)^2}{2} f'(u) + \dots \\ &= 1 - \frac{1}{u} + (x - u) \frac{\alpha}{n} + \frac{(x - u)^2}{2} \frac{\alpha^2}{n} + \dots \end{aligned} \quad (41)$$

Las sustituciones para  $f'(u)$  y derivadas más amplias, se obtienen por aplicaciones sucesivas de la regla de 'l Hospital. La expansión se simplifica en:

$$F(x) = 1 - \frac{1}{n} e^{-(x - u)\alpha} \quad (42)$$

entonces

$$\vartheta_n(x) = F^n(x) = \left[ 1 - \frac{1}{n} e^{-(x - u)\alpha} \right] \quad (43)$$

$$\vartheta(x) = \lim_{n \rightarrow \infty} \vartheta_n(x) = \exp \left[ -e^{-\alpha(x - u)} \right] \quad (44)$$

Esta es la forma limitante, y es independiente de la distribución inicial.

## APENDICE B

### Estimación de los Parámetros para la Distribución de Gumbel

El sistema ideado por Gumbel para estimar los parámetros de la distribución, depende del ajuste de una línea recta de los datos a analizar en un gráfico cuyas coordenadas son  $x$  e  $y$ , usando el método de mínimos cuadrados (Gumbel 1954). Los puntos en particular que se deben colocar tienen coordenadas  $x_m$   $y_m$ , donde:

$$\phi(y_m) = \frac{m}{n+1} \quad \text{o} \quad y_m = l_n \cdot \left( -l_n \frac{m}{n+1} \right) \quad (45)$$

siendo  $x_m$  el  $m$ ésimo valor observado, cuando los valores analizados son ordenados en orden creciente de magnitud.

Hay numerosas maneras de ajustar una mejor línea recta. Gumbel asumió una línea haciendo que la suma de los cuadrados de las desviaciones verticales fuera mínima como se muestra en la figura 4. Luego ajustó otra línea considerando que la suma de los cuadrados de las desviaciones horizontales sea mínima, según se muestra en la (figura 5). Las ecuaciones normales para los parámetros de estas dos líneas pueden ser escritas:

$$n \mu_1 + \frac{1}{\alpha_1} \sum y_m = \sum x_m \quad (46)$$

$$\mu_1 \sum y_m + \frac{1}{\alpha_1} \sum y_m^2 = \sum x_m y_m \quad (47)$$

y

$$\alpha_2 \sum x_m - n \alpha_2 \mu_2 = \sum y_m \quad (48)$$

$$\alpha_2 \sum x_m^2 - \alpha_2 \mu_2 \sum x_m = \sum x_m y_m \quad (49)$$

Estos coeficientes pueden ser escritos en términos de las medias y las desviaciones típicas de las variables. Escribimos:

$$\bar{x} = \frac{1}{n} \sum x_m \quad (50)$$

$$s_x^2 = \frac{1}{n} \sum x_m^2 - \bar{x}^2 \quad (51)$$

$$\bar{y} = \frac{1}{n} \sum y_m \quad (52)$$

$$\sigma^2 = \frac{1}{n} \sum y_m^2 - \bar{y}^2 \quad (53)$$

$$\overline{xy} = \frac{1}{n} \sum x_m y_m \quad (54)$$

Entonces las ecuaciones normales se convierten en:

$$\mu_1 + \frac{\bar{y}}{\alpha_1} = \bar{x} \quad (46a)$$

$$\mu_1 \bar{y} + \sum y_m^2 / \alpha_1 = \overline{xy} \quad (47a)$$

$$\alpha_2 \bar{x} - \alpha_2 \mu_2 = \bar{y} \quad (48a)$$

$$\alpha_2 \sum x_m^2 - \alpha_2 \mu_2 \bar{x} = \overline{xy} \quad (49a)$$

Las soluciones de estos conjuntos de ecuaciones son:

$$\frac{1}{\alpha_1} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_n^2} \quad (55)$$

$$x - \mu_1 = \frac{\bar{y}}{\alpha_1} \quad (56)$$

$$\frac{1}{\alpha_2} = \frac{S_x^2}{\overline{xy} - \bar{x} \cdot \bar{y}} \quad (57)$$

$$x = \mu_2 = \frac{\bar{y}}{\alpha_2} \quad (58)$$

Notamos en estas ecuaciones que  $(\mu_1)$  y  $(\mu_2)$  son las mismas funciones de  $(\alpha_1)$  y  $(\alpha_2)$ . En otras palabras, dado un valor de  $\alpha$  podemos derivar un valor de  $\mu$  sin conocer cuál juego de ecuaciones resulta.

En las ecuaciones para las  $\alpha$  vemos que las cantidades resultantes son de tres tipos: Primero,  $(\bar{x})$  y  $(S_x^2)$  son funciones tan solo de la observación  $y$  pueden ser fácilmente calculados de los valores de la muestra. Entonces  $(y)$  y  $(\sigma_n)$  dependen solo de las ordenadas, por lo tanto pueden ser calculadas sin hacer referencia a la muestra: si conocemos a  $n$ , la cantidad de datos a analizar y una tabla adecuada para  $y$ . Finalmente tenemos un término  $\overline{xy}$  el que envuelve a ambos  $(x)$  e  $(y)$ , el cuál debe por lo tanto ser calculado para cada muestra. Para obviar la ne

cesidad de este tedioso cálculo, Gumbel tomó como  $\alpha$  la media geométrica de  $\alpha_1$

y  $\alpha_2$  consiguiendo la forma simplificada:

$$\frac{1}{\alpha} = \frac{S_x}{\sigma_n} \tag{59}$$

y, de (56) ó (58)

$$\mu = \bar{x} - \bar{y} \frac{S_x}{\sigma_n} \tag{60}$$

La ecuación requerida de la línea que mejor encaja puede ahora ser escrita

$$x = \bar{x} - \bar{y}_n \frac{S_x}{\sigma_n} + y \frac{S_x}{\sigma_n} \tag{61}$$

Puesto que  $-\bar{y}$  es tan solo una función de la cantidad de datos a analizar, nosotros añadimos el subíndice  $-n$

### APENDICE C

#### Distribución de las muestras de Estimaciones:

Se demostró (ver a Cramer 1946) por ejemplo, que la distribución de una función tal como  $x_T$  de la ecuación (27) es normalmente asintótica. Precisando más, si tenemos una función  $H(\bar{x}, m_2)$  su distribución es normalmente asintótica, ó sea, se aproxima normalmente tanto como  $n$  aumenta con un valor medio, y con la variación dada por:

$$E(H) = H[E(x), \mu_2] \quad (62)$$

$$\sigma_H^2 = \mu_2(\bar{x}) \left(\frac{\partial H}{\partial \bar{x}}\right)^2 + \mu_2(m_2) \left(\frac{\partial H}{\partial m_2}\right)^2 + 2\mu_{11}(\bar{x}, m_2) \frac{\partial H}{\partial \bar{x}} \cdot \frac{\partial H}{\partial m_2} \quad (63)$$

en donde las derivadas parciales son evaluadas en  $\bar{x} = E(x)$ ,  $m_2 = \mu_2$ . Es de hacer notar aquí, que  $E(x)$  y  $\mu_2$  son parámetros característicos del universo estadístico del cuál han sido extractadas las muestras, mientras que  $\bar{x}$  y  $m_2$  son variables dependientes de las posibles muestras, siendo  $m_2$  lo mismo que  $S_x$  de (27).

Aplicando esto al presente caso, tenemos (Cramer 1946)

$$\mu_2(\bar{x}) = \frac{\mu_2}{n} \quad (64)$$

$$\mu_2(m_2) = \frac{\mu_4 - \mu_2^2}{n} \quad (65)$$

$$\mu_{11}(\bar{x}, m_2) = \frac{\mu_3}{n} \quad (66)$$

en donde los momentos en el lado derecho se refieren a la distribución del cuál ha sido extractada la muestra. De (27) tenemos:

$$\frac{\partial H}{\partial \bar{x}} = 1 ; \quad \frac{\partial H}{\partial m_2} = \frac{K}{2\mu_2^{1/2}} \quad (67)$$

$$\sigma_H^2 = \frac{\mu_2}{n} \left( 1 + K \frac{\mu_3}{\mu_2^{3/2}} + \frac{K^2}{4} \frac{\mu_4 - \mu_2^2}{\mu_2^2} \right) \quad (68)$$

$$E(H) = E(\bar{x}) + K \sqrt{\mu_2} \quad (69)$$

Para la distribución de la variable reducida

$$\phi(y) = \exp(-y - e^{-y})$$

$$\mu_2' = 1.6449$$

$$\mu_3' = 2.4041$$

$$\mu_4' = 14.6110$$

La relación entre los momentos de (y) é (x) continúan debido a que la relación entre ellos es lineal, de la forma:

$$x_T = a + by_T \quad (70)$$

por lo que  $\mu_r = b^r \mu_r'$ . Llevando estas relaciones a la ecuación (68)

$$\sigma_H^2 = \frac{\mu_2}{n} (1 + 1.140 K + 1.100 K^2) \quad (71)$$

Reemplazando  $\mu_2$  por la variación de la muestra  $S_x^2$ .

y poniendo  $\zeta(T) = (1 + 1.140 K + 1.100 K^2)^{1/2}$

$$\text{nos da } \sigma_H = \zeta(T) \frac{S_x}{n} \quad (72)$$

Del teorema dado arriba se expresa que la variable

$$t = \frac{X_t - E(X_t)}{\sigma_H} \quad (73)$$

esta distribuida normalmente con una media cero y variabilidad unitaria.

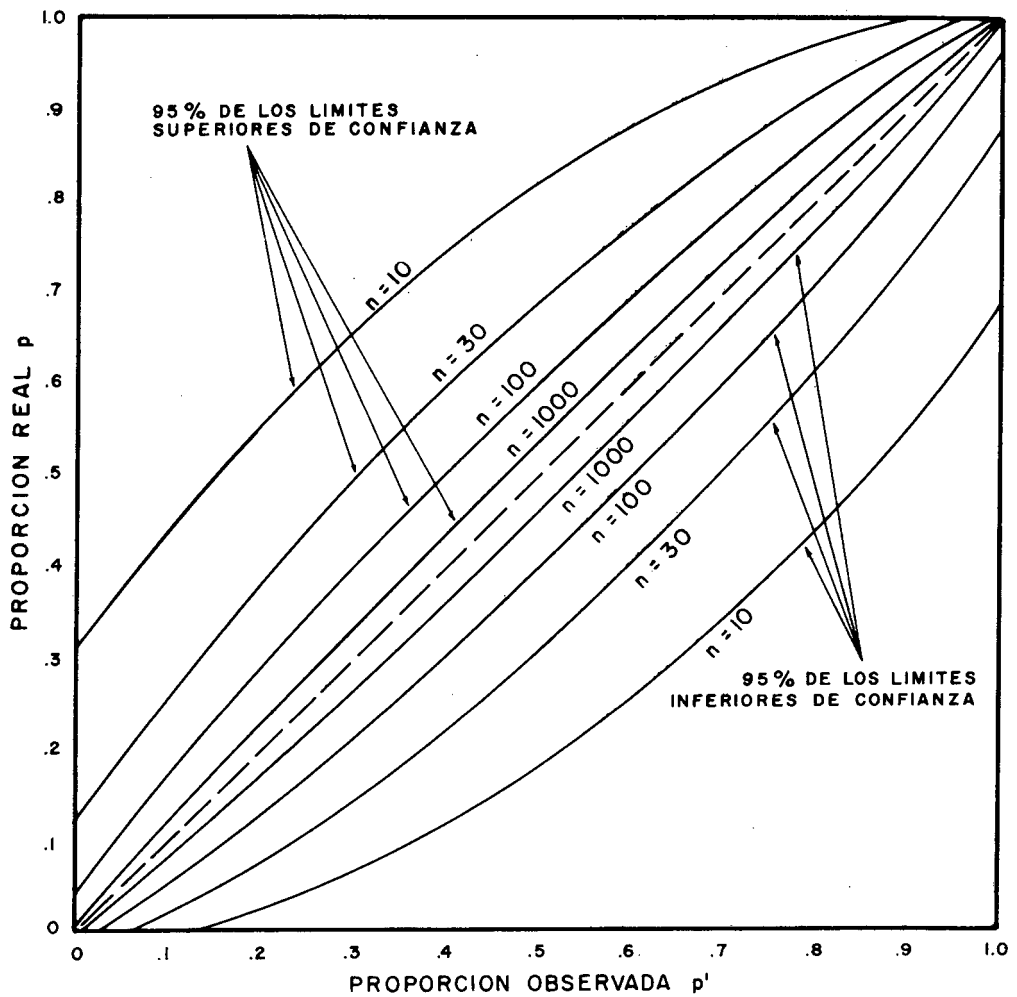


Figura 1.- 95% de los límites de confianza superior e inferior del valor de la proporción real  $p$  después que se ha estimado una  $p'$  derivada de una muestra de  $n$ .

PAPEL DE PROBABILIDADES DE GUMBEL

PERIODO \_\_\_\_\_

1.01

1.1

1.2

1.3

1.4

1.5

2

3

4

5

6

7

8

9

10

20

30

40

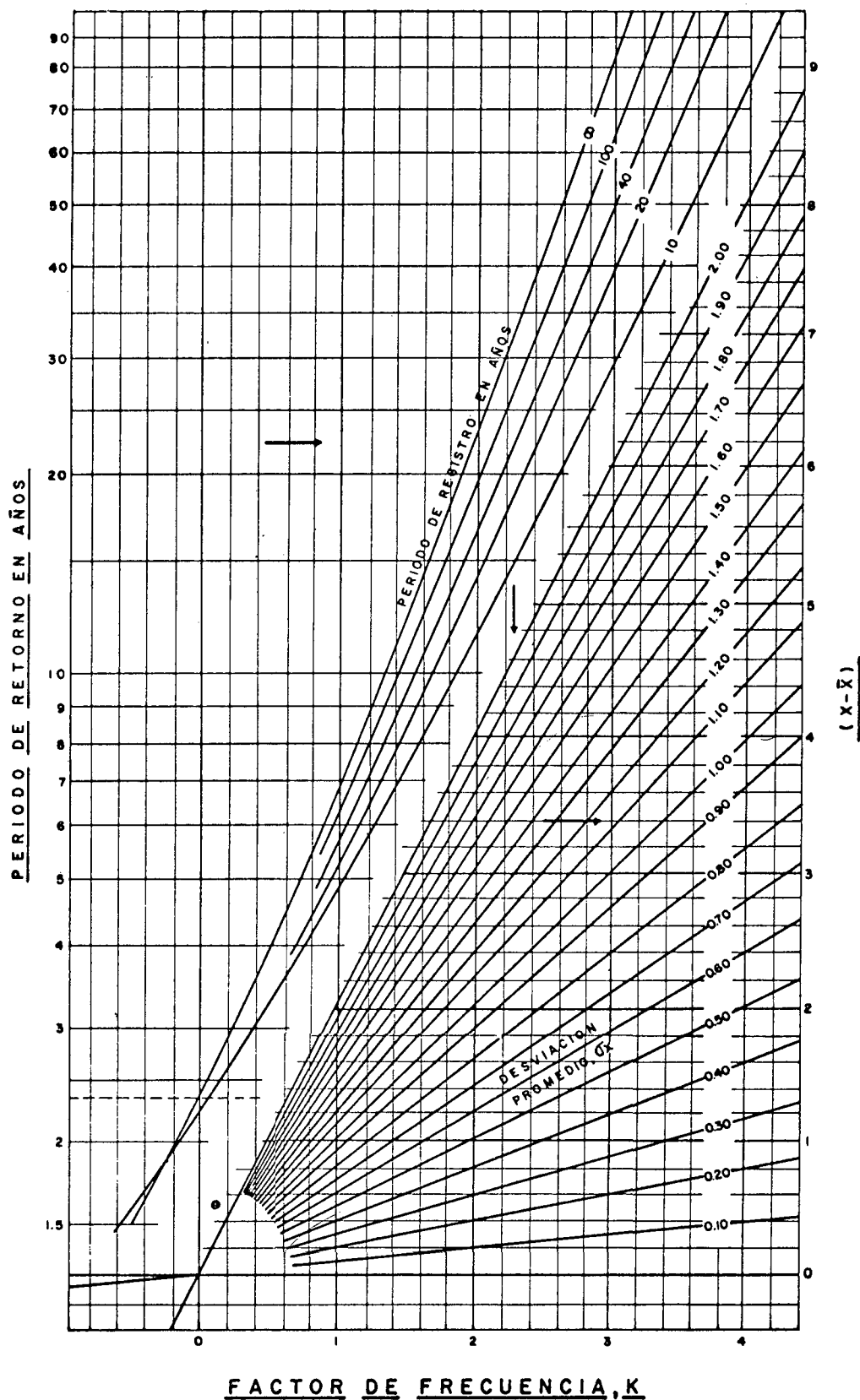
50

100

200

PERIODO DE RETORNO (AÑOS)

# NOMOGRAMA DE WEISS



NOTA: SI EL VALOR DE  $\sigma_x$  CAE FUERA DEL RANGO INDICADO, SE MULTIPLICAN TANTO  $\sigma_x$  COMO  $x - \bar{x}$  POR UN FACTOR APROPIADO

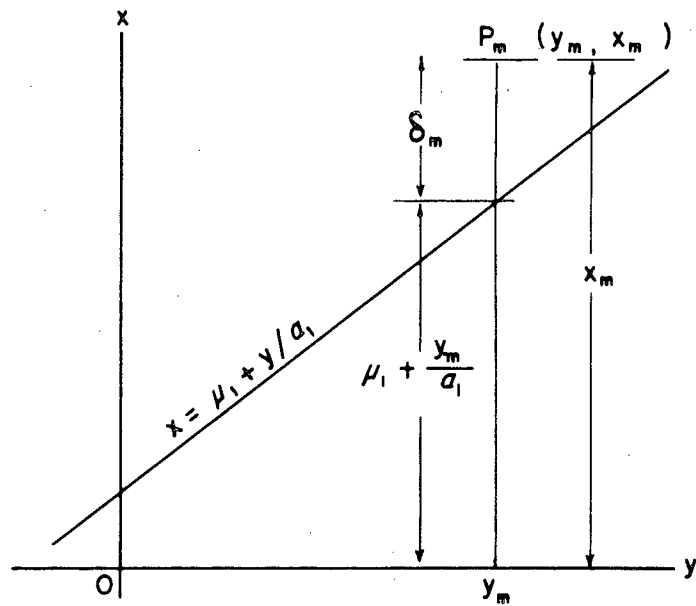


Figura 4.-Ajuste de la línea a través de los mínimos cuadrados de las desviaciones verticales.

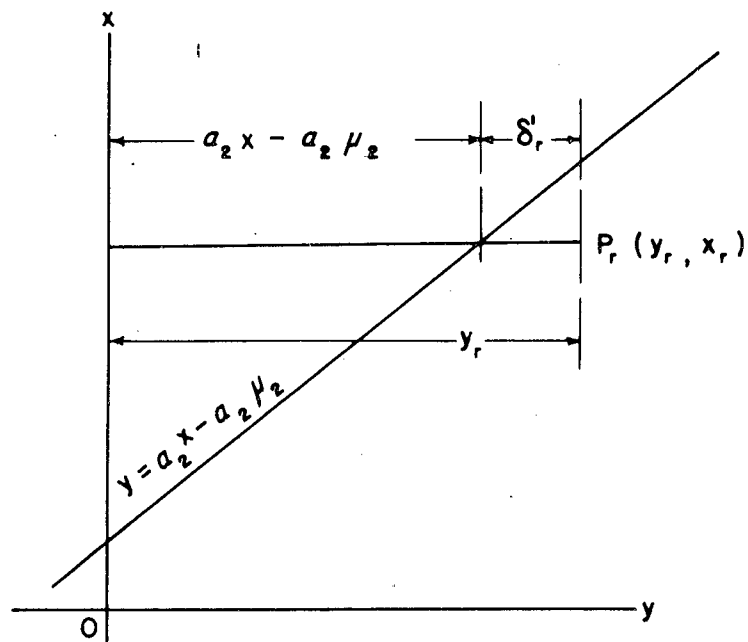


Figura 5.-Ajuste de la línea a través de los mínimos cuadrados de las desviaciones horizontales.

