ANALISIS DE FRECUENCIA EN HIDROLOGIA

JULIAN DAVID ROJO HERNANDEZ

Probabilidad - Período de retorno y riesgo

La probabilidad de ocurrencia de un fenómeno en hidrología puede citarse de varias Formas:

El evento hidrológico posee una probabilidad de ocurrencia del 2%

El evento hidrológico se presenta una vez cada 10 años.

Existe una relación entre probabilidad de ocurrencia de un evento hidrológico y tiempo de recurrencia el fenómeno. Por ejemplo:

Si un suceso hidrológico se presenta (en promedio) una vez cada 10 años, su probabilidad de ocurrencia será 0.1 (10%).

Si la probabilidad de ocurrencia de un determinado fenómeno hidrológico es de 0.04 (4%), significa que dicho fenómeno se presentará (en promedio) 4 veces en 100 años, es decir, una vez cada 25 años.

Período de retorno (Tr)

- Se define el período de retorno, Tr, de un evento de cierta magnitud como el tiempo promedio que transcurre entre la ocurrencia de ese evento y la próxima ocurrencia de ese evento con la misma magnitud. Se define también como *el tiempo que transcurre para que un evento sea excedido o igualado*, *al menos una vez en promedio*. Si P es la probabilidad de excedencia, entonces:

$$\boxed{P = \frac{1}{Tr}}$$

Riesgo

Supóngase que se calcula un cierto caudal para el periodo de retorno de 50 años, entonces la probabilidad de que se produzca dicho caudal este año será de 0.02 (1/50). Si este año no se produce dicho caudal, La probabilidad de que se produzca el año siguiente sigue siendo 0.02. en cualquier año la probabilidad es 0.02. Surge una pregunta: ¿ **Cual es la probabilidad de que se presente dicho caudal durante los próximos n años**?:

Probabilidad de que un suceso de periodo de retorno T *se presente* este año.................. 1/T

Probabilidad de que un suceso de periodo de retorno T **NO** *se presente* este año.............. 1- 1/T

Probabilidad de que un suceso de periodo de retorno T **NO** *se presente* en dos años.........(1- 1/T) (1- 1/T)

Probabilidad de que un suceso de periodo de retorno T **NO** *se presente* en **n** años..........(1- 1/T)^n

Probabilidad de que un suceso de periodo de retorno T **SI** *se presente* en **n** años.......... 1- (1- 1/T)^n

Entonces la probabilidad de que sí se presente alguna vez un suceso hidrológico con periodo de retorno de T durante los próximos n años se denomina Riesgo.

$$\boxed{\mathrm{R} = 1 - \left(1 - \frac{1}{Tr}\right)^{n}}$$

- **Ejemplo1:** Se va a construir un canal cuya *vida útil* es de 75 años. Si el caudal supera el valor correspondiente al período de retorno de 100 años se desbordará. Cual es la probabilidad de que se produzca algún desbordamiento en los próximos 75 años.

$$\mathrm{R} = 1 - \left(1 - \frac{1}{Tr}\right)^{n} = 1 - \left(1 - \frac{1}{100}\right)^{75} = 0.529 = 52.9\%$$

- **Ejemplo2:** Se esta diseñando una obra cuya vida útil se estima en 50 años y se admite que el riesgo de daño sea del 10%. ¿Cual debe ser el periodo de retorno del caudal de diseño? (Tarea)

Factor de frecuencia

Recordando el proceso de estandarización típico de la distribución normal, la variable aleatoria estandarizada se definía como:

$$z_i = \frac{x_i - \mu}{\sigma}$$

Por tanto se puede inferir que: $x_i = \mu + z_i \sigma$

Y para una muestra poblacional: $x_i = x + z_i s_x$

Es decir, el valor esperado de una variable aleatoria para para un determinado periodo de retorno T (xt) depende de su media y será proporcional a su desviación estándar, en general:

$$\boxed{x_T = x + k_T s_x}$$

Donde la constante de proporcionalidad **kt** se conoce como factor de frecuencia y depende, de la probabilidad de ocurrencia (el periodo de retorno) y su fdp.

Intervalos de confianza

- Cuando se desea hallar cualquier estadístico, por ejemplo la media, generalmente se dispone de una muestra de tamaño limitado.

- Se quiere saber qué tan cercano puede estar ese estimado al verdadero valor desconocido de la población.

- En otras palabras, se quisiera conocer con una cierta certeza (probabilidad) la franja de valores entre los cuales se encontraría el verdadero valor de la población.

- Franja grande: mucha incertidumbre.

α: Nivel de confianza o nivel de probabilidad.

S_T: Error estándar.

$$X_i \pm Z_{1-\alpha/2} S_{ET}$$

El error estándar, ST, es una medida de la desviación estándar de la magnitud de un evento calculado a partir de una muestra respecto a la verdadera magnitud del evento.

$$S_{ET} = \frac{\hat{\sigma}_x}{\sqrt{N}}\left(1 + \frac{K^2}{2}\right)^{\frac{1}{2}}$$

Ejemplo: Caudales de río Escondido

Cual es el caudal correspondiente a un T_R=100 años, si tienen una distribución normal. N=25 años.

$$\mu = 283.5 \ m^3/s$$

$$\sigma = 24.8 \ m^3/s$$

En este caso se puede escribir**:**

$$P = 1/Tr = 0.01$$

$$Fu(K) = 1 - P = 0.99$$

$K = z = F_u^{-1}(0.99)$ **Tabla K=2.326**

$$Q_{100} = \mu + K\sigma = 283.5 + 2.326 \times 24.8$$
$$Q_{100} = 341.18$$

Intervalos de confianza para α=5%:

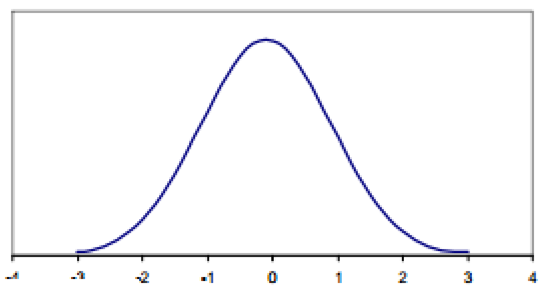
$$S_{ET} = \frac{\hat{\sigma}_x}{\sqrt{N}}\left(1 + \frac{K^2}{2}\right)^{\frac{1}{2}} = 8.71$$

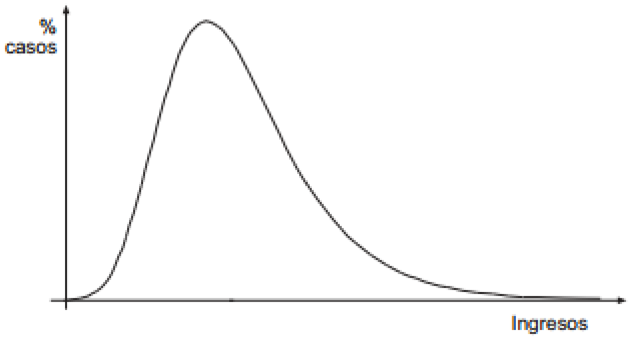
$$z_{1-\alpha/2} = z_{0.975} \Rightarrow Tabla \Rightarrow 1.96$$
$$X_T \pm z_{1-\alpha} S_{ET} \Rightarrow 341.18 \ \pm 1.96 \times 8.71 = 341.19 \pm 11.07 \quad m^3/s$$

DISTRIBUCIONES DE PROBABILIDAD EN EL ANÁLISIS DE FRECUENCIAS

Distribuciones simétricas y asimétricas

Muchas variables naturales se ajustan a la distribución simétrica propuesta por Gaus. Sin embargo existen casos en los que no existe la misma proporción de los pequeños que de los grandes. Dando como resultado una fdp asimétrica.





En hidrología los caudales y las precipitaciones medias anuales suelen ajustarse a la distribución simétrica de Gaus. Pero los caudales y las precipitaciones máximas o mínimas (extremos) no. Estos últimos suelen ajustarse a campanas asimétricas

Distribuciones de probabilidad para análisis de frecuencias.

Los matemáticos han encontrado para nosotros las ecuaciones de muchas de las campanas asimétricas.

- **Distribuciones Asimétricas continuas:**

  - Distribución General de Valor Extremos (Gumbel)
  - Distribución Gamma (Pearson)

  - Distribución Log Normal
  - Distribución Log Gumbel
  - Distribución Log Pearson.

Para cada distribución existen ecuaciones que permiten determinar el factor de frecuencia para un determinado Tr y su respectiva banda de error.

Distribución General de Valor Extremo

Los valores extremos son valores máximos y mínimos seleccionados de un conjuntos de datos.

Las distribuciones de valores extremos seleccionados de conjuntos de muestras de cualquier distribución de probabilidad convergen en una de las tres formas de *distribución de valor extremo*, llamadas:

- Tipo I: Gumbel
- Tipo II: Frechet
- Tipo III: Weibull

- **Función de Distribución de probabilidad para la GEV**

$$F(x) = \exp\left[-\left(1 - \kappa\frac{x-\beta}{\alpha}\right)^{1/\kappa}\right]$$

Donde:

κ, β y α son parámetros que deben ser determinados.

Los tres casos limitantes son:

1. $\kappa = 0 \rightarrow$ Distribución de Valor Extremo Tipo I (Gumbel)

$$f(x) = \frac{1}{\alpha}\exp\left[-\frac{x-\beta}{\alpha}-\exp\left(-\frac{x-\beta}{\alpha}\right)\right]$$

Rango: $-\infty < x < \infty$

Estimación de parámetros: $\beta = \bar{x} - 0.5772\alpha$ $\hat{\alpha} = \frac{\sqrt{6}}{\pi}\hat{\sigma}$

2. $\kappa < 0 \rightarrow$ Distribución de Valor Extremo Tipo II (Frechet)

$$f(x) = \exp\left[-\left(1 - \kappa \frac{x-\beta}{\alpha} \right)^{1/\kappa} \right]$$

Rango: $(\beta + \alpha / \kappa) \leq x \leq \infty$

3. $\kappa > 0 \rightarrow$ Distribución de Valor Extremo Tipo III (Weibull)

$$f(x) = \exp\left[-\left(1 - \kappa \frac{x-\beta}{\alpha} \right)^{1/\kappa} \right]$$

Rango: $-\infty \leq x \leq (\beta + \alpha / \kappa)$

Distribución Gumbel

La fda y el factor de frecuencias es:

$$F(x) = \exp\left[-\exp\left(-\frac{x-\beta}{\alpha}\right)\right] \quad -\infty \le x \le \infty$$

$$K = -\frac{\sqrt{6}}{\pi}\{0.577 + \ln[\ln T_r - \ln(T_r - 1)]\}$$

- **Intervalos de confianza:** Que tan cercano puede estar el estimado al valor verdadero desconocido de la población.

$$X_T \pm u_{1-\alpha/2} S_T$$

α: Nivel de confianza o nivel de probabilidad

S_T: Error estándar

$$S_T = \delta \frac{\sigma}{\sqrt{N}} \qquad \delta = [1 + 1.1396K + 1.1K^2]^{1/2}$$

Como ajustar la fda de la distribución Gumbel a una muestra de datos (Fácil)

Método 1: (rápido, buenos resultados)

$$F(x) = \exp\{-\exp[-a(x-b)]\} \quad -\infty \le x \le \infty$$

Siendo:

$$a = \frac{1.2825}{S_x} \qquad b = \overline{X} - 0.45 S_x$$

Método 2: (EL mejor, mas preciso)

$$F(x) = \exp\{-\exp[-c]\}$$

$$c = a(x-b)$$

$$a = \frac{\sigma_y}{S_x} \qquad b = \overline{X} - \frac{\mu_y}{a}$$

nº datos	μ_y	σ_y
10	0,4952	0,9496
15	0,5128	1,0206
20	0,5236	1,0628
25	0,5309	1,0914
30	0,5362	1,1124
35	0,5403	1,1285
40	0,5436	1,1413
45	0,5463	1,1518
50	0,5485	1,1607
55	0,5504	1,1682
60	0,5521	1,1747
65	0,5535	1,1803
70	0,5548	1,1854
75	0,5559	1,1898
80	0,5569	1,1938
85	0,5578	1,1974
90	0,5586	1,2007
95	0,5593	1,2037
100	0,5600	1,2065

Ejemplo 1: Gumbel

De una serie de 55 caudales extremos se sabe que su media es 21.97 m3/s y la desviación estándar es de 13.22 m3/s.

a) calcular la probabilidad de que se supere un caudal de 60 m3/s.
b) ¿Cuál caudal se superará el 1% de los años?.

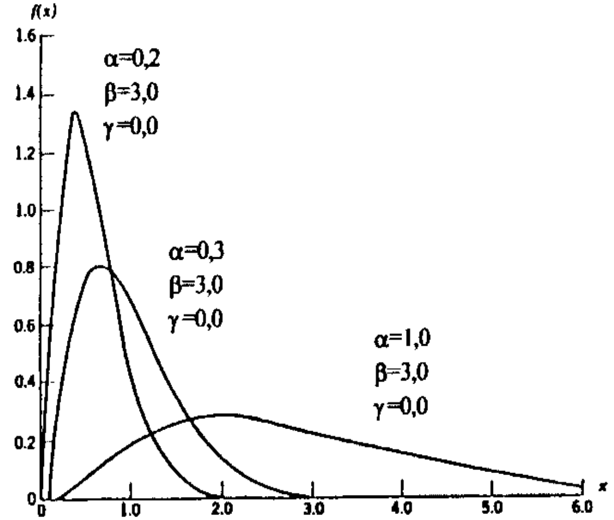
Por el método 1)

Ejemplo 3 : Distribución Gumbel

Los caudales máximos del rio Nare tienen una distribución Gumbel, determinar el caudal con Tr de 100 años si se sabe que el valor medio es de los máximos es de 94.35 y su desviación de 22.45 y sus intervalos de confianza

$$K_{100} = \frac{-\sqrt{6}}{\pi}\left\{0.577 + \ln\left[\ln T_R - \ln(T_R - 1)\right]\right\}$$

$K_{Tr=100} = 3.13$



Ejemplo: Distribución Gumbel

$$Q_{Tr} \pm \mu_{95} S_T \qquad S_T = \delta \frac{\sigma}{\sqrt{N}}$$

$$\delta = \left[1 + 1.1396K + 1.1K^2\right]^{1/2}$$

Intervalos de confianza:

$\delta = 3.91$

$S_T = 14.62$

$Q_{TR} \pm 1.6 * 14.62$

$141.4 \leq 164.77 \pm 188.17$ m^3/s

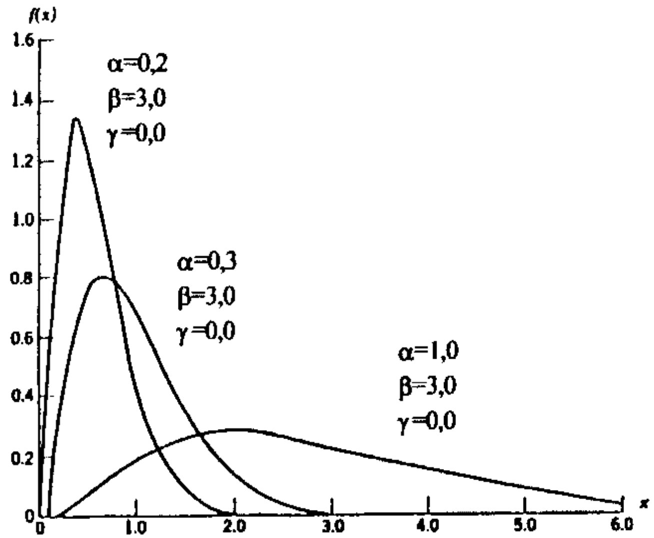
Distribución Gamma (2 Parámetros)

Una de las mas usadas en Hidrología.

$$f(x) = \frac{1}{|\alpha|\Gamma(\beta)}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\frac{x}{\alpha}}$$

- Crecientes máximas anuales
- Caudales mínimos
- Volúmenes de flujo anuales y estacionales
- Valores de precipitaciones extremas
- Volúmenes de lluvia de corta duración

Tiene 2 ó 3 parámetros (Pearson Tipo III).



Parámetros y Factor de frecuencia

- α (Parámetro de escala)
- $\beta > 0$ (Parámetro de forma)
- $\Gamma(\beta)$ es la función Gamma completa

$$\Gamma(\beta) = \int_0^{\infty} z^{\beta-1} e^{-z}\, dz$$

Estimación de parámetros: Método de los momentos

$$\mu = \alpha\beta$$

$$\sigma^2 = \alpha^2\beta$$

Distribución Gamma 3 Parámetros(Pearson Tipo III)

- **Función de distribución de probabilidad**

$$f(x) = \frac{1}{|\alpha|\Gamma(\beta)}\left(\frac{x - x_o}{\alpha}\right)^{\beta\text{-}1} \exp\left(- \frac{x - x_o}{\alpha}\right)$$

- **Función de densidad acumulada**

$$P(X \leq x) = \frac{1}{\alpha\Gamma(\beta)} \int_0^X e^{-\left(\frac{x-x_0}{\alpha}\right)}\left(\frac{x-x_0}{\alpha}\right)^{\beta-1} dx$$

- **Parámetros**

α y β, parámetros de escala y forma respectivamente.
x_o parámetro de localización.

$$\Gamma(\beta) = \int_0^\infty z^{\beta\text{-}1} e^{\text{-}z} dz$$

Parámetros e Intervalos de confianza (Función Gamma)

- **Estimación de Parámetros:** Método de los momentos

$$\hat{\beta} = \left(\frac{2}{\hat{\gamma}}\right)^2 \qquad \hat{\alpha} = \hat{\sigma}\,\frac{\hat{\gamma}}{2} \qquad \hat{x}_0 = \hat{\mu} - \hat{\alpha}\hat{\beta}$$

Que tan cercano puede estar el estimado al verdadero valor desconocido de la población: Conocer con cierta certeza. Franja grande: mucha incertidumbre.

$$X_T \pm u_{1-\alpha/2}\, S_T$$

α: Nivel de confianza o nivel de probabilidad

S_T: Error estándar

$$S_T = \delta \frac{\sigma}{\sqrt{N}}$$

Tabla Factor de frecuencia Pearson tipo III

Coeficiente de Asimetría	Probabilidad de Excedencia						
	0.500	0.200	0.100	0.040	0.020	0.010	0.005
3.0	-0.396	0.420	1.180	2.278	3.152	4.051	4.970
2.9	-0.390	0.440	1.195	2.277	3.134	4.013	4.909
2.8	-0.384	0.460	1.210	2.275	3.114	3.973	4.847
2.7	-0.376	0.479	1.224	2.272	3.093	3.932	4.783
2.6	-0.368	0.499	1.238	2.267	3.071	3.889	4.718
2.5	-0.360	0.518	1.250	2.262	3.048	3.845	4.652
2.4	-0.351	0.537	1.262	2.256	3.023	3.800	4.584
2.3	-0.341	0.555	1.274	2.248	2.997	3.753	4.515
2.2	-0.330	0.574	1.284	2.240	2.970	3.705	4.444
2.1	-0.319	0.592	1.294	2.230	2.942	3.656	4.372
2.0	-0.307	0.609	1.302	2.219	2.912	3.605	4.298
1.9	-0.294	0.627	1.310	2.207	2.881	3.553	4.223
1.8	-0.282	0.643	1.318	2.193	2.848	3.499	4.147
1.7	-0.268	0.660	1.324	2.179	2.815	3.444	4.069
1.6	-0.254	0.675	1.329	2.163	2.780	3.388	3.990
1.5	-0.240	0.690	1.333	2.146	2.743	3.330	3.910

Valores de δ para la Distribución Gamma ó Pearson tipo III

γ	$T_r=2$	$T_r=5$	$T_r=10$	$T_r=20$	$T_r=50$	$T_r=100$
0.0	1.0801	1.1698	1.3748	1.6845	2.1988	2.6363
0.1	1.0808	1.2006	1.4367	1.7810	2.3425	2.8168
0.2	1.0830	1.2309	1.4989	1.8815	2.4986	3.0175
0.3	1.0866	1.2609	1.5610	1.9852	2.6656	3.2365
0.4	1.0913	1.2905	1.6227	2.0915	2.8423	3.4724
0.5	1.0987	1.3199	1.6838	2.1998	3.0277	3.7238
0.6	1.1073	1.3492	1.7441	2.3094	3.2209	3.9895
0.7	1.1179	1.3785	1.8032	2.4198	3.1208	4.2684
0.8	1.1304	1.4082	1.8609	2.5303	3.6266	4.5595
0.9	1.1449	1.4385	1.9170	2.6403	3.8374	4.8618
1.0	1.1614	1.4699	1.9714	2.7492	4.0522	5.1741
1.1	1.1799	1.5030	2.0240	2.8564	4.2699	5.4952
1.2	1.2003	1.5382	2.0747	2.9613	4.4996	5.8240
1.3	1.2223	1.5764	2.1237	3.0631	4.7100	6.1592
1.4	1.2157	1.6181	2.1711	3.1615	4.9301	6.4992
1.5	1.2701	1.6643	2.2173	3.2557	5.1486	6.8427
1.6	1.2952	1.7157	2.2627	3.3455	5.3644	7.1881
1.7	1.3204	1.7732	2.3081	3.4303	5.5761	7.5339
1.8	1.3452	1.8374	2.3541	3.5100	5.7827	7.8783
1.9	1.3690	1.9091	2.4018	3.5844	5.9829	8.2196
2.0	1.3913	1.9888	2.4525	3.6536	6.1755	8.5562

Ejemplo: Distribución Gamma

Hallar el $Q_{TR=100}$. Si la distribución de los caudales de la estación de Nare es Gamma.

$$\mu = 94.35\ m^3/s \text{ y } \sigma = 22.45\ m^3/s,\ \gamma = 0.845$$

$$\mu_Y = 4.52 \text{ y } \sigma_Y = 0.2337,\ \gamma_Y = 0.0069$$

De tabla: $K = 2.32$

$$Q_{TR=100} = 94.35 + 2.32*22.45 = 146.4$$

Intervalos de confianza:

$$\mathbf{X_T \pm u_{1-\alpha/2}\, S_T}$$

De tabla δ=4.7, N= 36 datos.

$$S_T = \delta \frac{\sigma}{\sqrt{N}}$$

S_T = 17.6

De tabla μ_{95}=1.6

146,4$\pm$ 1.6*17.6

146.4 $\pm$ 28.16 m^3/s

Distribución Log Normal

En general, cuando la variable aleatoria X es el producto de un gran número de otras variables aleatorias, la distribución de los logaritmos de X puede aproximarse a la Normal, ya que los logaritmos de X son la suma de los logaritmos de los factores contribuyentes.

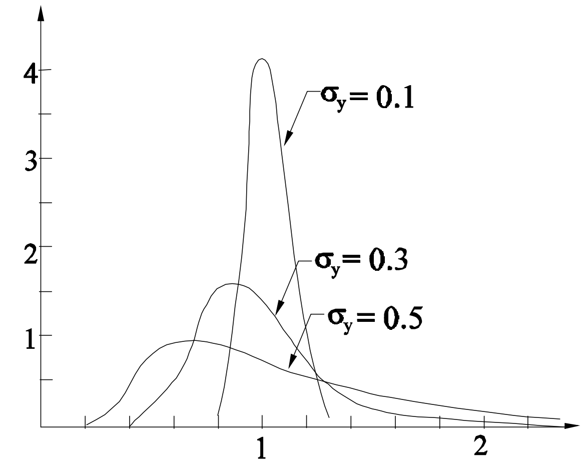
Si se tiene una variable aleatoria X y ln X = Y, se ajusta a una distribución Normal, se dice que la variable aleatoria X es log normalmente distribuida.

- **Función de Distribución de Probabilidad**

Asumiendo $Y = \log_a (X)$

$$f(x) = \frac{1}{\sigma_y x \sqrt{2\pi}} \exp\left[-\frac{1}{2} \frac{(y - \mu_y)^2}{\sigma_y^2} \right]$$

Parámetros y Factor de frecuencia



- Media (Parámetro de escala)
- Desviación estandar (Parámetro de forma)

Estimación de parámetros: Método de los momentos

$$\hat{\mu}_Y = \frac{1}{N}\sum_{i=1}^{N} \log_a(X_i) \qquad \hat{\sigma}_Y = \left\{\left(\frac{1}{N}\right)\sum_{i=1}^{N}\left[\log_a(X_i) - \hat{\mu}_Y\right]^2\right\}^{1/2}$$

$$\ln(X_T) = \mu_y + K\,\sigma_y$$ *K es la misma de la distribución normal*

Si se quiere trabajar con la variable no transformada en el campo logarítmico se tiene que:

$$K = \frac{\exp\left[K_T\left(\ln\left(1+Cv^2\right)\right)^{1/2} - \left[\frac{\ln\left(1+Cv^2\right)}{2}\right]\right] - 1}{Cv}$$

$$K_T = F_u^{-1}\left(1 - \frac{1}{T_r}\right)$$

$F_u^{-1}\left(1-\frac{1}{T}\right) \rightarrow$ Es el inverso de la función de distribución Normal estandarizada acumulada y Cv es el coeficiente de variación

- **Intervalos de confianza**

α: Nivel de confianza o significancia

S_T: Error estándar

$$\ln(X_T) \pm u_{1-\alpha/2} S_T$$

$$S_T = \delta \frac{\sigma_Y}{\sqrt{N}} \qquad \delta = \left(1 + \frac{K_T^2}{2}\right)^{1/2}$$

Ejemplo: Distribución Log Normal

La media y desviación estándar de los Q_{max} anuales de la estación del río Nare son:

$$\mu=94.35 \ m^3/s \ y \ \sigma=22.45 \ m^3/s$$

$$\mu_Y=4.52 \ y \ \sigma_Y=0.2337$$

Hallar el $Q_{TR=100}$ si los Q_{max} tienen una distribución Log Normal.

$$K=2.326$$

$$Q_{Y \ Tr=100}=4.52+2.326*0.237$$

$$Q_{Tr=100}=159 \ m^3/s$$

Intervalos de Confianza: $Ln(Q_{TR=100}) \pm \mu_{95} S_T$

Es un intervalo de dos colas, con una probabilidad en cada una de 5%

$$S_T = \delta \frac{\sigma_Y}{\sqrt{N}} \qquad \delta = \left(1 + \frac{K_T^2}{2}\right)^{1/2}$$

δ=1.92

S_T=0.075

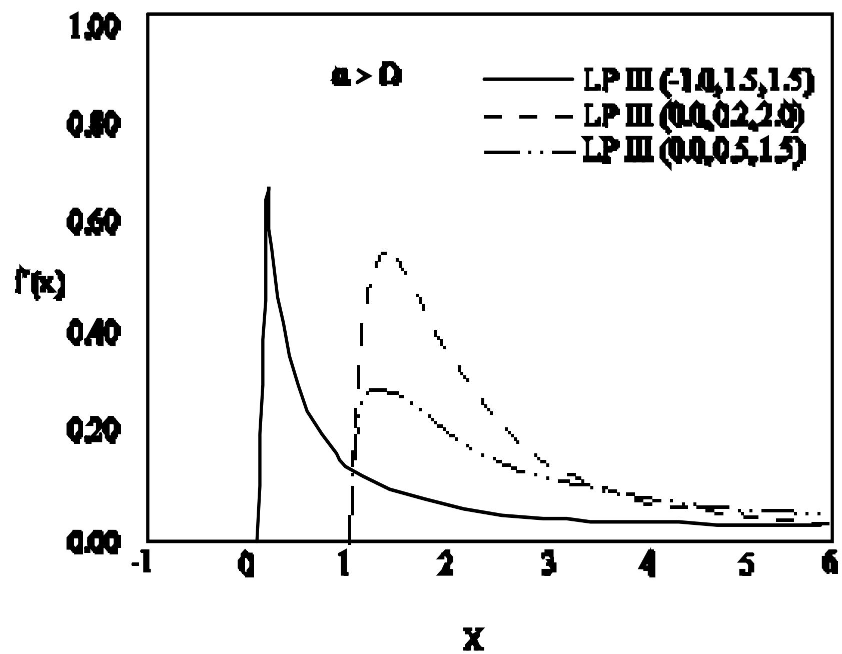
5.071±1.6*0.075

$4.94 \leq Q_Y \leq 5.14$ → $139 \leq 159 \leq 170$ m^3/s

Distribución Log Pearson Tipo III

- **Función de distribución de probabilidad**

$$f_x(x) = \frac{1}{x\,|\alpha|\,\Gamma(\beta)}\left[\frac{\ln(x) - y_o}{\alpha}\right]^{\beta-1} e^{-\left[\frac{\ln(x) - y_o}{\alpha}\right]}$$



- **Parámetros**

α y β, parámetros de escala y forma

y y_o parámetro de localización

- **Estimación de Parámetros**

Método de los momentos

$$\hat{\beta} = \left(\frac{2}{\hat{\gamma}_y}\right)^2 \qquad \hat{\alpha} = \hat{\sigma}_y \frac{\hat{\gamma}_y}{2} \qquad \hat{y}_0 = \hat{\mu}_y - \hat{\alpha}\hat{\beta}$$

- **Factor de Frecuencia:**

$$Y_T = \ln X_T = \hat{\mu}_y + K\,\hat{\sigma}_y$$

- **Intervalos de Confianza:** Que tan cercano puede estar el estimado al verdadero valor desconocido de la población: Conocer con cierta certeza. Franja grande: mucha incertidumbre.

$$X_T \pm u_{1-\alpha/2}\, S_T$$

α: Nivel de confianza o nivel de probabilidad
S_T: Error estándar

$$S_T = \delta \frac{\hat{\sigma}_y}{\sqrt{N}} \qquad\qquad \ln X_T \pm \mu_{1-\alpha/2} S_T$$