

2. FUNDAMENTOS TEÓRICOS PARA EL CÁLCULO DE CAUDALES MÁXIMOS

2.1 Introducción

Los fundamentos teóricos de las metodologías utilizadas para la determinación de caudales máximos en este estudio, las cuales son algunas de las utilizadas en El Salvador; fueron obtenidos de variada bibliografía, detallada en el apartado bibliográfico de este documento.

El objetivo de este capítulo es dar a conocer los fundamentos teóricos de las metodologías utilizadas, que posteriormente se validarán, para la obtención de caudales máximos en El Salvador, mejorando así su comprensión; además de servir como guía para posteriores estudios relacionados con la temática abordada. Existen metodologías tanto Hidrometeorológicas como Estadísticas para la determinación de caudales máximos. Las metodologías Hidrometeorológicas tomadas en cuenta en este trabajo son: Método Racional, Hidrogramas Unitarios Complejo y Sintéticos de Snyder, Triangular y SCS; y las metodologías Estadísticas utilizadas son Puntuales y Regionales.

2.2 Fundamentos teóricos sobre Hidrometeorología

La hidrometeorología es el estudio de la meteorología aplicada a los parámetros hídricos. La teoría hidrometeorológica en general, comprende la observación, procesamiento y análisis del comportamiento de los elementos hídricos, fundamentalmente las descargas de los ríos y los volúmenes almacenados en reservorios y lagunas; y de los elementos meteorológicos, fundamentalmente la precipitación pluvial. [Segoviano, 1974: p.257]

Las teorías hidrometeorológicas para el cálculo de caudales máximos son las siguientes: Método Racional, el cual comprende determinación de coeficiente de escurrimiento C , Curvas de Intensidad-Frecuencia-Duración y cálculos de tiempos de concentración; Hidrogramas Unitarios, los cuales se dividen en Sintéticos (Snyder, Triangular y SCS) y Complejos.

2.2.1 Cálculo de caudales máximos utilizando el método de la Fórmula Racional

Este método, que la literatura inglesa atribuye a Lloyd-George en 1906, si bien los principios del mismo fueron establecidos por Mulvaney en 1850, permite determinar el caudal máximo que escurrirá por una determinada sección, bajo el supuesto que éste acontecerá para una lluvia de intensidad máxima constante y uniforme en la cuenca correspondiente a una duración D igual al tiempo de concentración de la sección.

$$Q_{m\acute{a}x} = CiA \quad (\text{Ec. 2.1})$$

En donde: $Q_{m\acute{a}x}$: Caudal máximo en la sección de cálculo, C : Coeficiente de escorrentía medio ponderado de la cuenca, A : Área total de la cuenca vertiente en la sección de cálculo, i : Intensidad media máxima para una duración igual al tiempo de concentración, de la sección de cálculo. A continuación se detallan los fundamentos teóricos para determinar cada una de las variables mencionadas anteriormente. [Schmidth, 1986: p.356]

a) Determinación del Coeficiente de Escurrimiento C

El coeficiente de escurrimiento C representa la fracción de la lluvia que escurre en forma directa y toma valores entre cero y uno, y varía apreciablemente entre una cuenca y otra, y de una tormenta a otra, debido a las condiciones de humedad iniciales. Sin embargo, es común tomar valores de C representativos de acuerdo con ciertas características de las cuencas como la vegetación, pendientes del terreno y uso de suelos. [German Monsalve, 1999: p.179].

b) Cálculo de curvas de Intensidad-Frecuencia-Duración. [Aparicio, 2001: p.167]

El método utilizado relaciona simultáneamente las tres variables Intensidad-frecuencia-Duración, en una familia de curvas cuya ecuación es:

$$i = \frac{kT^m}{(d+c)^n} \quad (\text{Ec. 2.2})$$

Donde k , m , n y c son constantes que se calculan mediante un análisis de correlación lineal múltiple; T : Período de retorno en años, d : Duración en minutos e i : Intensidad en mm/h.

Si se toman logaritmos de la ecuación 2.2, se obtiene:

$$\text{Log } i = \log k + m \log T - n \log (d + c) \quad (\text{Ec. 2.3})$$

O bien:

$$y = a_0 + a_1 x_1 + a_2 x_2 \quad (\text{Ec. 2.4})$$

Donde: $y = \log i$, $a_0 = \log k$, $a_1 = m$, $x_1 = \log T$, $a_2 = -n$, $x_2 = \log (d + c)$.

La ecuación 2.4 es la de una familia de líneas rectas de pendientes a_2 , ordenada al origen a_0 y espaciamiento a_1 .

Si los datos registrados de i , d y T se dibujan en papel logarítmico, usualmente se agrupan en torno a líneas rectas. A veces las curvas resultan ligeramente curvas, lo que se puede corregir agregando a las duraciones un valor constante c , o bien, en algunos casos, cuando la pendiente de las líneas varía mucho, dividiendo la línea para cada período de retorno en dos rectas. Si los datos se agrupan lo suficiente en torno a líneas rectas, el valor de c puede tomarse como cero.

Al hacer un ajuste de correlación lineal múltiple de una serie de tres tipos de datos, se obtiene un sistema de ecuaciones como el siguiente:

$$? y = N a_0 + a_1 ? x_1 + a_2 ? x_2 \quad (\text{Ec. 2.5})$$

$$y = a_0 + a_1 x_1 + a_2 (x_1 x_2) \quad (\text{Ec. 2.6})$$

$$y = a_0 + a_1 x_2 + a_2 (x_2)^2 \quad (\text{Ec. 2.7})$$

Donde N es el número de datos y las incógnitas son a_0 , a_1 y a_2 ; x_1 es el logaritmo del período de retorno, x_2 es la duración (con el valor de c agregado de ser necesario) y y es la intensidad, obtenidos de un registro de precipitación. El período de retorno se obtiene de la siguiente ecuación:

$$T = (n + 1) / m \quad (\text{Ec. 2.8})$$

Donde m = número de orden en una lista de mayor a menor de los datos y n = número de datos. (El inverso del período es la frecuencia).

Una vez calculados los coeficientes a_0 , a_1 y a_2 es posible evaluar los parámetros k , m y n de la ecuación 2.1 y con ella, dibujar las curvas Intensidad – Frecuencia – Duración.

c) Determinación de Tiempos de Concentración t_c .

El tiempo de concentración no es más que el tiempo que tardaría una gota de agua en recorrer la longitud desde el punto más distante de la corriente de agua de una cuenca hasta el lugar de medición. Los tiempos de concentración son calculados a partir de las características físicas de la cuenca, las cuales son: las pendientes, longitudes, elevaciones medias y el área de la cuenca. Es de notar que todas las fórmulas tienen factores de corrección que aplican según la cobertura de la cuenca. [German Monsalve, 1999: p.180]

Tiempo de concentración a partir de la fórmula empírica de Kirpich

Desarrollada a partir de la información del SCS en siete cuencas rurales en Tennessee con canales bien definidos y pendientes empinadas (de 3% a 10%)

$$T_c = 0.0003245 L^{0.77} S^{-0.385} \text{ (En horas)} \quad (\text{Ec. 2.9})$$

L: Longitud máxima del canal o río desde aguas arriba hasta la salida, (en metros).

S: Pendiente del cauce o H/L (m/m) donde H es la diferencia de elevación entre el punto más elevado y el punto de interés. [Environmental Modeling System http://www.emsi.com/wmshelp/HydrologicModels/Calculators/Computing_Travel_Times/Using_Basin_Data/Equations/Time_of_Concentration/Kirpich_Tc_Equation.htm. Abril de 2004].

Tiempo de concentración a partir de la fórmula de la Federal Aviation Administration (FAA)

Desarrollada a partir de la información sobre el drenaje de aeropuertos, recopilada por el cuerpo de ingenieros de Estados Unidos; el método tiene como finalidad el ser usado en problema de drenaje de aeropuertos, pero ha sido frecuentemente usado para flujo superficial en cuencas urbanas. [José Llamas, 1993: p. 96]

$$T_c = 0.7024 (1.1 - C)L^{0.5} S^{-0.33} \text{ (En horas)} \quad (\text{Ec. 2.10})$$

C: Coeficiente de escorrentía del Método Racional (adimensional).

L: Longitud de flujo superficial (en metros).

S: Pendiente de la superficie.

Tiempo de concentración a partir de la fórmula empírica de Giandotti

$$T_c = [4\sqrt{A} + 1.5L] / [0.8\sqrt{E}] \text{ (En horas)} \quad (\text{Ec. 2.11})$$

A: Área de la cuenca (en kilómetros cuadrados).

L: Longitud promedio de flujo superficial (en kilómetros).

E: Elevación media de la cuenca. (en metros).

Tiempo de concentración a partir de la fórmula empírica del U.S. Bureau of Reclamation of California

Conocida también como La fórmula de California Culverts Practice. Esencialmente es la ecuación de Kirpich, desarrollada para pequeñas cuencas montañosas en California.[Ven Te Chow, 1994: p. 96]

$$T_c = 0.94788 \left(L^3 / H \right)^{0.385} \text{ (En horas)} \quad \text{(Ec. 2.12)}$$

L: Longitud de flujo superficial (en kilómetros).

H: Diferencia de nivel entre la divisoria de aguas y la salida (en metros).

d) Validación de fórmulas empíricas para la estimación de Tiempos de Concentración.

Cuando la lluvia cae sobre una cuenca y los niveles de infiltración y de evaporación son iguales o inferiores a la intensidad del aguacero, comienza el fenómeno de escorrentía sobre toda la superficie de la cuenca afectada por dicho aguacero.

El agua resbala, antes de alcanzar un río principal o secundario, bajo la forma de capas de agua, de una cierta altura (según la intensidad de las precipitaciones y la pendiente de la superficie de flujo) y una cierta velocidad, la cual se denomina *velocidad de la escorrentía*. [José Llamas, 1993: p. 375]

La velocidad de la escorrentía se utiliza para validar las fórmulas con que se obtienen los tiempos de concentración, convirtiendo, estos últimos, a velocidades y comparándolas con las velocidades de escorrentía. Ella depende en primer lugar, de la pendiente de la superficie de flujo y después, de las características del suelo. A continuación se presenta una fórmula empírica desarrollada para el cálculo de esta velocidad. En la ecuación, *a* es el ángulo del terreno. [José Llamas, 1993: p. 375]

$$v = 20sen^{3/5}a \text{ (m/s)} \quad (\text{Ec. 2.13})$$

v : Velocidad de esorrentía (m/s).

Anteriormente se detalló la metodología para la determinación de tiempos de concentración de cuatro maneras distintas. Los tiempos de concentración se transforman a velocidad dividiendo la longitud del cauce máximo, desde el nacimiento del cauce hasta el punto de interés, de cada una de las subcuencas en estudio, entre los tiempos de concentración obtenidos por cada metodología.

Existe un rango de mediciones de velocidades obtenidas por la Ing. Adriana Erazo del Servicio Hidrológico Nacional en la cuenca del río Paz entre las estaciones hidrológicas de La Hachadura y El Jobo, el cual fue utilizado como base de comparación (1-3m/s).

También existe un estudio realizado en cuencas de Zaragoza, España, el cual relacionó los tiempos de concentración con las velocidades de esorrentía y obtuvo la siguiente ecuación:

$$T_c = 0.3 \left[\left(\frac{1/3L}{S^{1/4}} \right)^{0.76} \right] + \frac{2/3L}{3.6v} \quad (\text{Ec. 2.14})$$

Donde: L : Longitud del cauce principal (km) desde el punto de inicio del cauce hasta el punto de medición, S : Pendiente media del cauce, T_c : Tiempo de concentración (en horas)

En una de las conclusiones del estudio anterior, presentaron rangos de velocidades de esorrentía en función de la pendiente media del cauce. Concluyeron que para pendientes de 0 a 10% las velocidades de esorrentía estaban en el rango de 1 a 2 m/s como se muestra en la tabla 2.1. [Documento sobre Condicionantes Físicos para el Ordenamiento de la Orla Sudoeste de Suelo Urbanizable].

Pendiente media de la cuenca (%)	Velocidad de recorrido inicial (m/s)
Menor del 5 %	1 m/s
Del 5 a 10 %	1 - 2 m/s
Mayor del 10 %	2 m/s

Tabla 2.1. Velocidades de recorrido de tanteo, en función de la pendiente media de la cuenca.

Existe además, un estudio de "Sistematización por medio de terrazas de predios destinados a la agricultura"; J.C. Molinelli; Dirección de Agronomía, Publ. N° 93, 1948; Montevideo. Citado por Ghiggia, R. 1981. En el que obtiene rangos de velocidades de escorrentía específicos para la determinación de tiempos de concentración en función del tipo de superficie o suelo en la que el agua escurrió. Estos rangos se presentan en la siguiente tabla.

Promedio de velocidades de escurrimiento para calcular el tiempo de concentración (m/s)				
Condiciones de la superficie	0-3 %	4-7 %	8-11 %	12-15 %
Aguas no concentradas				
Montes	0.3	0.61	0.9	1.07
Pasturas	0.45	0.91	1.22	1.37
Tierras cultivadas	0.61	1.22	1.52	1.83
Pavimentos	1.52	3.65	4.72	5.49

Nota: Las condiciones de la superficie se representan en porcentajes de pendientes.

Tabla 2.2. Promedio de velocidades de escurrimiento para calcular el tiempo de concentración.

[www.fagro.edu.uy, septiembre 2004].

Los estudios anteriores sirvieron como base de sustentación para los distintos tiempos de concentración obtenidos en este documento [www.fagro.edu.uy, septiembre 2004].

Ven Te Chow también propone una tabla similar a la anterior, obtenida de datos representativos al estado de Texas, Estados Unidos. .[Ven Te Chow, 1994: p. 169].

**Velocidades promedio aproximadas en m / s del flujo de escorrentía para
calcular el tiempo de concentración.**

Descripción del curso del agua	Pendiente en porcentaje			
	0-3	4-7	8-11	> 12
No concentrado				
Bosques	0-0.46	0.46-0.76	0.76-1.00	> 1.00
Pastizales	0-0.76	0.76-1.07	1.07-1.30	> 1.30
Cultivos	0-0.91	0.91-1.37	1.37-1.68	> 1.68
Pavimentos	0-2.59	2.59-4.11	4.11-5.18	> 5.18

Esta condición usualmente ocurre en las partes superiores de la cuenca antes de que el flujo superficial se acumule en un canal.

Tabla 2.3. Velocidades promedio aproximadas de escorrentía para calcular el tiempo de concentración

2.2.2 Estimación de Hidrogramas Unitarios.

Para la estimación de caudales máximos utilizando hidrogramas unitarios, siempre es necesario contar con al menos un hidrograma medido a la salida de la cuenca y con los registros de precipitación que originaron el hidrograma.

La mayor parte de las cuencas, no solo en El Salvador, sino en todo el mundo, no cuentan con una estación hidrométrica o bien con los registros pluviográficos necesarios. Por ello, es conveniente contar con métodos con los que puedan obtenerse hidrogramas unitarios usando únicamente datos de características generales de la cuenca. Los hidrogramas unitarios así obtenidos se denominan *sintéticos*. Los hidrogramas unitarios sintéticos a analizar son los siguientes: Snyder, Triangular y SCS. [Francisco Aparicio, 2001: p.228]

a) Estimación de Hidrogramas Unitarios Sintéticos

Fundamentos teóricos para la estimación del Hidrograma Unitario Sintético de Snyder.

[German Monsalve, 1999: p.219].

Este procedimiento tiene utilidad cuando no se cuenta con los datos necesarios conjuntos de caudal y precipitación históricos para la deducción del hidrograma unitario de una cuenca.

La deducción de los parámetros para definir los hidrogramas unitarios sintéticos se basan en las características geométricas y morfológicas de la cuenca hidrográfica. En la región de los Montes Apalaches, en Los Estados Unidos, Snyder estableció que, para cuencas de 16 a 16,100 kilómetros cuadrados:

$$t_p = 0.7517C_t(LxL_t)^{0.3} \quad (\text{Ec. 2.15})$$

Donde:

t_p : Tiempo de retardo de la cuenca (en horas)

C_t : Coeficiente adimensional variando entre 1.8 y 2.2, tomando los valores menores para cuencas con grandes inclinaciones.

L : Longitud del río principal desde la divisoria de aguas hasta el punto en consideración (en kilómetros).

L_t : Longitud desde el punto del río principal más próximo al centro geométrico de la cuenca hasta el punto en consideración (en kilómetros).

$$t_r = t_p / 5.5 \quad (\text{Ec. 2.16})$$

Donde:

t_r : Duración de la lluvia neta (en horas).

$$q_p = \frac{0.275C_p A}{t_p} \quad (\text{Ec. 2.17})$$

Donde:

q_p : Caudal pico del HU por milímetro de lluvia neta ($\text{m}^3/\text{s}/(\text{mm})$).

A : Área de drenaje de la cuenca (en km^2).

C_p : Coeficiente adimensional variable entre 0.56 y 0.69 tomando valores mayores para cuencas con grandes inclinaciones.

El hidrograma unitario sintético obtenido corresponde a un milímetro de precipitación neta sobre toda la cuenca.

$$T = 3 + t_p / 8 \quad (\text{Ec. 2.18})$$

Donde:

T : Tiempo base de la escorrentía.

En la figura 2.1 se muestran los gráficos de los hidrogramas sintéticos de Snyder con la simbología utilizada.

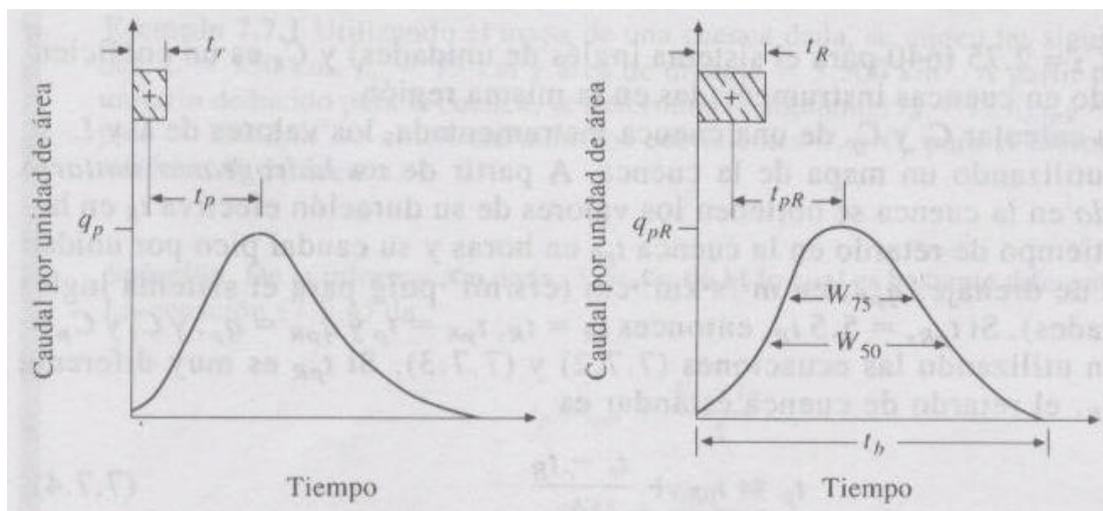


Figura 2.1. Hidrogramas Sintéticos de Snyder

Para otra lluvia neta de duración t_R diferente a t_p el tiempo de retardo correspondiente a t_{pR} el caudal pico q_{pR} y el caudal base T_R según Linsley son:

$$t_{pR} = t_p + \frac{T_R - t_r}{4} \quad (\text{Ec. 2.19})$$

$$T_R = 3 + \frac{t_{pR}}{8} \quad (\text{Ec. 2.20})$$

$$q_{pR} = \frac{0.275C_p A}{t_{pR}} \quad (\text{Ec. 2.21})$$

Adicionalmente:

$$W_{75} = \frac{0.10192}{q_p^{1.08}} \quad \text{ó} \quad W_{75R} = \frac{0.10192}{q_{pR}^{1.08}} \quad (\text{Ec. 2.22})$$

$$W_{50} = \frac{0.17836}{q_p^{1.08}} \quad \text{ó} \quad W_{50R} = \frac{0.17835}{q_{pR}^{1.08}} \quad (\text{Ec. 2.23})$$

Sea que se trabaje con una lluvia neta de duración t_r o con una duración t_R en donde:

W_{75} : Ancho del HU , horas, correspondiente a un valor de las ordenadas igual al 75% del caudal pico q_p o q_{pR} .

W_{50} : Ancho del HU , horas, correspondiente a un valor de las ordenadas igual al 50% del caudal pico q_p o q_{pR} .

$$q_p = \frac{q_p}{A} \left(\frac{(m^3/s)/(mm)}{km^2} \right) \quad (\text{Ec. 2.24})$$

$$q_{pR} = \frac{q_{pR}}{A} \left(\frac{(m^3/s)/(mm)}{km^2} \right) \quad (\text{Ec. 2.25})$$

El valor de t_p , tiempo de retardo de la cuenca es igual al tiempo entre el centro geométrico de la duración de la lluvia neta t_r , y el pico del hidrograma unitario (en horas) con base en las características t_p , t_r , q_p , W_{75} , W_{50} y T (ó t_{pR} , t_R , q_{pR} , W_{75R} , W_{50R} , y T_R). Se traza el hidrograma unitario de escorrentía superficial. Se debe comprobar que:

$$\int Q_{ES} dt = P_n xA = 1mmxA \quad (\text{Ec. 2.26})$$

En caso contrario se debe ajustar el hidrograma (usualmente el tiempo base T o T_R) hasta lograr que se cumpla la anterior ecuación.

Fundamentos teóricos para la estimación del Hidrograma Unitario Sintético Triangular.
[Francisco Aparicio, 2001: p. 232]

Desarrollado por Mockus en 1957. Se busca un gasto pico Q_p , en función del área de la cuenca en km^2 y el tiempo base en horas:

$$Q_p = \frac{0.555A}{t_b} (m^3 / s / mm) \quad (\text{Ec. 2.27})$$

Del análisis de varios hidrogramas, Mockus concluye que el tiempo base y el tiempo pico se relacionan mediante la expresión:

$$t_b = 2.67 t_p \quad (\text{Ec. 2.28})$$

El tiempo pico se expresa:

$$t_p = d_e / 2 + t_r \quad (\text{Ec. 2.29})$$

En donde:

d_e : Duración en exceso.

t_r : Tiempo de retraso.

y:

$$t_r = 0.6 t_c \text{ (en horas)} \quad (\text{Ec. 2.30})$$

Donde:

t_c : Tiempo de concentración.

Además la duración:

$$d_e = 2\sqrt{t_c} \text{ Para cuencas grandes (en horas).} \quad (\text{Ec. 2.31})$$

$$d_e = t_c \text{ Para cuencas pequeñas (en horas).} \quad (\text{Ec. 2.32})$$

Y se calculan las características del *HU* triangular.

Fundamentos teóricos para la estimación del Hidrograma Adimensional SCS. [Ven Te Chow, 1994: p.236]

El hidrograma adimensional *SCS* es un hidrograma unitario sintético en el cual el caudal se expresa por la relación del caudal q con respecto al caudal pico q_p y el tiempo por la relación del tiempo t con respecto al tiempo de ocurrencia del pico en el hidrograma unitario T_p . Dados el caudal pico y el tiempo de retardo para la duración de exceso de precipitación, el hidrograma unitario puede estimarse a partir del hidrograma sintético adimensional para la cuenca dada. La figura 2.2 muestra el hidrograma adimensional, preparado utilizando los hidrogramas unitarios para una variedad de cuencas. Los valores de q_p y T_p pueden estimarse usando un modelo simplificado de un hidrograma unitario triangular, en donde el tiempo está dado en horas y el caudal en $\text{m}^3/\text{s}\cdot\text{cm}$.

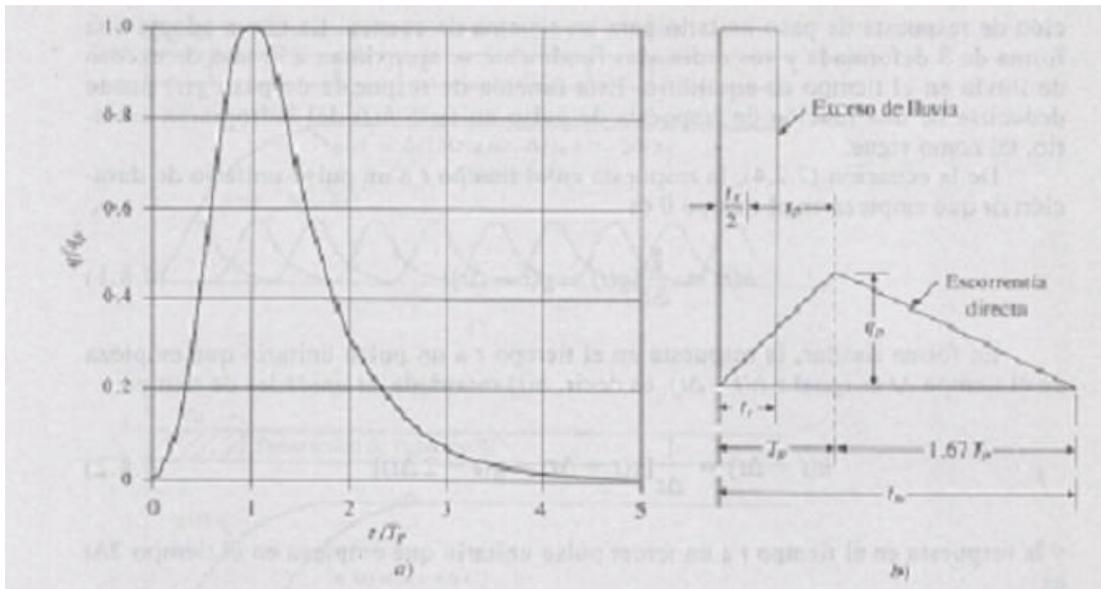


Figura 2.2. Hidrogramas Sintéticos SCS y Triangular

Con base en la revisión de un gran número de hidrogramas unitarios, el Soil Conservation Service sugiere que el tiempo de recesión puede aproximarse como $1.67T_p$. Como el área bajo el hidrograma unitario debería ser igual a una escorrentía directa de 1 cm, puede demostrarse que:

$$q_p = CA/T_p \quad (\text{Ec. 2.33})$$

Donde $C = 2.08$ y A es el área de drenaje en kilómetros cuadrados.

Adicionalmente, un estudio de los hidrogramas unitarios de muchas cuencas rurales grandes y pequeñas indica que el tiempo de retardo $t_p @ 0.6T_c$, donde T_c es el tiempo de concentración de la cuenca. Como se muestra en la figura 2.2, el tiempo de ocurrencia del pico T_p puede expresarse en términos del tiempo de retardo t_p y de la duración de la lluvia efectiva t_r .

$$T_p = t_r/2 + t_p \quad (\text{Ec. 2.34})$$

b) Estimación de Hidrogramas Unitarios Complejos

Construcción del Hidrograma de Caudal

Un hidrograma de caudal es una gráfica o una tabla que muestra la tasa de flujo como función del tiempo en un lugar dado de la corriente. Se grafican los valores de caudales máximos horarios para la tormenta a la que le corresponde el caudal máximo anual para cada estación en estudio [Ven Te chow, 1994: p.135].

Separación del Caudal Base y el Caudal de Escurrimiento Directo mediante el método de la Línea Recta

El método consiste en ubicar el punto donde empieza a crecer la gráfica (comienzo de incremento de caudal) y luego se traza una línea hasta interceptar el comienzo de la curva de agotamiento. El caudal por debajo de esa línea corresponde al aporte del agua subterránea y el resto a la escorrentía superficial total. [German Monsalve, 1999: p.186]. El cálculo del hidrograma de escorrentía directa consiste en restar el caudal base del caudal observado para la tormenta en particular.

Cálculo de Hietograma de Exceso de Lluvia [Ven Te Chow, 1994: p.138-142]

Para calcular el hietograma de exceso de lluvia se debe calcular el volumen de escurrimiento directo V_d y la profundidad de escorrentía directa r_d de la siguiente manera:

$$V_d = \sum_{n=1}^N Q_n \Delta t \quad (\text{Ec. 2.35})$$

Donde Δt es la longitud del intervalo de tiempo seleccionado para el análisis.

$$r_d = V_d / Ac \quad (\text{Ec. 2.36})$$

Donde Ac es el área de la cuenca.

Como siguiente paso se procede a la estimación de la tasa de abstracciones de lluvia (f) que se originan por infiltración y almacenamiento superficial en la cuenca, así como el número de pulsos o intervalos M de lluvia diferentes de cero de escorrentía en exceso, los cuales se encuentran por método de ensayo y error de la siguiente manera; el valor de f se calcula seleccionando un intervalo de tiempo de longitud Δt , estimando el número de intervalos M de lluvia que realmente contribuyen a la escorrentía directa, restando $f\Delta t$ de la precipitación que se observa en cada intervalo (R_m) y ajustando los valores de f y M tantas veces como sea necesario para que las profundidades de escorrentía directa y de exceso de precipitación sean iguales, así:

$$r_d = \sum_{m=1}^M (R_m - f\Delta t) \quad (\text{Ec. 2.37})$$

Las ordenadas del hietograma de exceso de precipitación se calculan sustrayendo $f\Delta t$ de las ordenadas del hietograma de precipitación observada, despreciando todos los intervalos en las cuales la profundidad de precipitación observada es menor que $f\Delta t$.

Cálculo del Hidrograma Unitario Complejo [Ven Te Chow, 1994: p.233, 234]

La ecuación de convolución discreta que es la función respuesta de pulso discreto para un sistema lineal, permite el cálculo de la escorrentía directa Q_n dado un exceso de lluvia P_m y el hidrograma unitario U_{n-m+1} , así:

$$Q_n = \sum_{m=1}^{n \leq M} P_m U_{n-m+1} \quad (\text{Ec. 2.38})$$

El hidrograma unitario se deduce utilizando un proceso inverso llamado *deconvolución*, dada una información de P_m y Q_n . Supóngase que existen M pulsos de exceso de lluvia y N pulsos de escorrentía directa en la tormenta considerada; luego pueden escribirse N ecuaciones para Q_n , $n = 1, 2, \dots, N$, en términos de $N-M+1$ valores desconocidos del hidrograma unitario.

2.2.3 Cálculo de caudales máximos a través de los Hidrogramas Unitarios (Sintéticos y Complejo)

a) Cálculo de Lluvia Neta a través del método SCS.

El procedimiento inicia estimando el número de curva CN de tablas, según el tipo de suelo y utilizando factores de corrección según el caso. Una vez estimado el CN , se procede a calcular la retención potencial máxima (S):

$$S = \left(\frac{1000}{CN} - 10 \right) (2.54) (cm) \quad (\text{Ec. 2.39})$$

Una vez estimada la retención potencial máxima, se procede a estimar la precipitación efectiva o profundidad de escorrentía con los valores de precipitación total (P) y retención potencial máxima (S):

$$Pe = \frac{(P - 0.2S)^2}{P + 0.8S} (2.54) (cm) \quad (\text{Ec. 2.40})$$

Siendo esta ecuación válida únicamente para valores de $P > 0.2S$. Cuando $P < 0.2S$

$$Pe = 0.$$

b) Cálculo de caudales máximos a través del Hidrograma Sintético de Snyder [http://www.geocities.com/awesome_quad/cap3/eyr3w.htm, septiembre de 2004]

De acuerdo con las investigaciones de F. Snyder en 1938 para la determinación de caudales máximos a través del Hidrograma Sintético de Snyder para cuencas de 16 a 16,000 km², la fórmula empírica necesaria para la determinación de estos caudales es la siguiente:

$$Q_{m\acute{a}x} = q_p \times I \times t_c \quad (\text{en m}^3/\text{s}) \quad (\text{Ec. 2.41})$$

Donde t_c es el tiempo de concentración en horas, I es la intensidad máxima en mm/h, y q_p es el caudal pico del Hidrograma unitario de Snyder en $m^3/s/mm$.

c) Cálculo de caudales máximos a través de los Hidrogramas Sintéticos Triangular, SCS e Hidrograma Unitario Complejo

Para la determinación de caudales máximos por medio de Hidrogramas Unitarios Complejos y Sintéticos se necesita de un hietograma de lluvia. Cuando no existe registro de precipitación sobre la cuenca, es posible obtener dicho hietograma utilizando el método del Bloque Alterno.

Método del Bloque Alterno para el cálculo de Hietogramas de precipitaciones de diseño

El método del bloque alternativo es una forma simple para desarrollar un hietograma de diseño utilizando una curva I-D-F. El hietograma de diseño producido por este método especifica la profundidad de precipitación que ocurre en n intervalos de tiempo sucesivos de duración Δt sobre una duración total de $T_d = n\Delta t$. Después de seleccionar el período de retorno de diseño, la intensidad es medida en una curva I-D-F para cada una de las duraciones $\Delta t, 2\Delta t, 3\Delta t, \dots$, y la profundidad de precipitación correspondiente se encuentra al multiplicar la intensidad y la duración. Tomando diferencias entre valores sucesivos de profundidad de precipitación, se encuentra la cantidad de precipitación que debe añadirse por cada unidad adicional de tiempo Δt . Estos incrementos o bloques se reordenan en una secuencia temporal de modo que la intensidad máxima ocurra en el centro de la duración requerida T_d y que los demás bloques queden en orden descendente alternativamente hacia la derecha y hacia la izquierda del bloque central para formar el hietograma de diseño.

2.3 Cálculo de caudales máximos utilizando el método Regional de Índice de Creciente [Documento de Regionalización de Caudales Máximos y Medios en El Salvador SNET, 2004: apartado 3.5]

Este método se basa en el uso simultáneo de los datos registrados en todas las estaciones ubicadas dentro de una zona considerada hidrológicamente homogénea, con lo que la serie resultante es mucho más larga que la de una estación en particular, además de que permite el cálculo de caudales máximos en cualquier cuenca que no posea estación hidrométrica, ya que establece relaciones entre las características fisiográficas de las cuencas y los caudales máximos.

La regionalización de caudales máximos que se llevó a cabo en El Salvador, fue a través del método del índice de creciente (Flood Index) el cual supone que los máximos anuales dentro de la región hidrológicamente homogénea siguen una misma función de distribución y lo que varía es un factor de escala de acuerdo a cada cuenca ubicada en dicha región. Para cada región, se establece una serie que consiste en la suma de los caudales divididos por su media en las estaciones que conforman la región.

Para aplicar el método Regional de Índice de Creciente, para la estimación de caudales máximos en cualquier punto de El salvador, se hace lo siguiente:

Se ubica la región Hidrológicamente homogénea a la cual pertenece la cuenca a la cual se le va a estimar los caudales máximos para diferentes períodos de retorno (mapa de regiones Hidrológicamente homogéneas, (anexo A).

Se determina el caudal correspondiente al promedio de los caudales máximos ($Q_{2.33}$) con base en el área de la cuenca y a la región Hidrológicamente homogénea a la que pertenece (tabla 2.4).

REGIÓN	ECUACIÓN	R ²	RANGO DE ÁREA (km ²)
1	$Q_{2.33}=0.6839*A+72.986$	0.9925	100-1991
2	$Q_{2.33}=2.1408*A-71.75$	0.9946	55-110
2b	$Q_{2.33}=0.9257*A-172.78$	0.9275	187-430
3	$Q_{2.33}=0.5871*A+198.91$	0.9931	100-1930
3b	$Q_{2.33}=0.0701*A+122.32$	0.7167	1640-2240
4	$Q_{2.33}=0.6758*A+53.357$	0.9197	25-200
5	$Q_{2.33}=-0.0008*A^2+1.6108*A+4.2165$	0.991	45-120
6	$Q_{2.33}=0.3519*A+53.544$	0.6362	45-845
7	$Q_{2.33}=0.4868*A^{1.107}$	0.9882	13-560
8	$Q_{2.33}=-5E-06*A^2+0.3154*A+205.28$	0.9702	915-18200

Tabla 2.4. Ecuaciones de relación entre el valor medio de los caudales máximos $Q_{2.33}$ y el área de la cuenca

[Documento de Regionalización de Caudales Máximos y Medios en El Salvador SNET, 2004: p. 3.5]

El $Q_{2.33}$ se multiplica por los factores de ajuste para los diferentes períodos de retorno (tabla 2.5), obteniendo así, el valor de caudal máximo para diferentes períodos de retorno.

PERÍODO DE RETORNO	REGIÓN									
	1	2	2b	3	3b	4	5	6	7	8
5	1.64	1.50	1.39	1.40	1.54	1.50	1.51	1.42	1.38	1.40
10	2.28	1.96	1.73	1.74	2.05	1.96	1.99	1.79	1.71	1.75
15	2.68	2.24	1.93	1.94	2.36	2.24	2.28	2.01	1.90	1.96
20	2.98	2.45	2.07	2.09	2.60	2.44	2.49	2.17	2.04	2.11
25	3.23	2.61	2.18	2.20	2.79	2.61	2.66	2.30	2.15	2.22
50	4.05	3.14	2.54	2.57	3.41	3.13	3.22	2.71	2.49	2.59
100	4.96	3.71	2.90	2.94	4.08	3.70	3.81	3.14	2.84	2.98

Tabla 2.5. Factores de ajuste para el cálculo de caudales máximos [Documento de Regionalización de

Caudales Máximos y Medios en El Salvador SNET, 2004: p. 3.5].

2.4 Fundamentos teóricos Estadísticos para la determinación de caudales máximos y determinación de Función de mejor ajuste

El estudio de un fenómeno hidrológico requiere del análisis de datos o muestras históricas recopiladas para poder comprender el comportamiento del mismo, así como para tomar decisiones relativas a un proyecto de ingeniería que dependa en gran medida del fenómeno en cuestión.

La determinación de caudales máximos a partir de datos históricos de estaciones hidrométricas localizadas en las zonas hidrográficas *B, C, D, E, F, G, H, I, J* de El Salvador, se pretende realizar mediante un análisis estadístico, por lo que es de vital importancia el conocer de una manera formal las técnicas estadísticas más apropiadas para así obtener la mejor información posible y para poder cuantificar el riesgo que representa la generalización a partir de informaciones parciales. [Llamas, 1993: p. 87].

Algunos conceptos de probabilidad y estadística

Para estudiar las técnicas y recursos estadísticos es necesario tomar en cuenta algunos conceptos básicos sobre probabilidad, debido a que ésta es el nexo entre Hidrología y Estadística.

Cuando se va a realizar un experimento, sólo se presentan dos situaciones: se sabe con certeza lo que va a ocurrir o no se sabe. En Hidrología esta semejanza es válida, y ocurre lo que en el segundo caso: no se sabe con certeza lo que ocurrirá, aún bajo las mismas condiciones, los fenómenos hidrológicos se "predicen" en base a "probabilidades".

Visto desde este punto, se dice que estamos en este caso frente a lo que se denomina "experiencia aleatoria".

La teoría de la probabilidad estudia las experiencias aleatorias y las posibles formas en que ocurren, llamadas *sucesos* o *eventos*.

La probabilidad es un número real, entre *ceros* y *uno*, que se le adjudica a los sucesos como una medida de su posible ocurrencia. Si un suceso no puede ocurrir jamás, como resultado de la experiencia, se llama "suceso imposible" y tiene una probabilidad de *ceros*. Si, por el contrario, es evidente que siempre ocurrirá, se le llama "suceso seguro" y tiene una probabilidad de *uno*.

Cualquier otro suceso diferente de los señalados tomará su valor de probabilidad en el intervalo real abierto $]0, 1[$. [Hernández, 2002: p. 11 y 12].

En este intervalo, como se dijo, cada suceso es relacionado con un número real llamado "variable aleatoria", de un modo tal que la asignación de la probabilidad se realiza desde un conjunto de números reales hacia el intervalo $[0, 1]$, donde la suma de todos los valores de probabilidad es igual a 1.

Las variables aleatorias numéricas se clasifican en *discretas* y *continuas*. Se llaman discretas cuando toman valores puntuales; por lo general, números enteros. Se llaman continuas cuando, real o teóricamente, pueden tomar cualquier valor dentro de un conjunto bien definido de números reales. [Hernández, 2002, p.72].

Esta asignación de la probabilidad obliga a analizar la forma en que dicha probabilidad se distribuye en cada una de las variables aleatorias. Cuando el número de observaciones se incrementa, el tamaño de los intervalos decrece. Si los intervalos en los que ocurren los sucesos u observaciones, es tan pequeño que tienden a cero, la distribución de probabilidades se hace mediante una determinada función.

Funciones de probabilidad

Existen dos tipos de funciones de probabilidad:

Funciones discretas: cuando el número de valores x que puede tomar una variable aleatoria X es finito, se dice que la variable aleatoria X es discreta (por ejemplo, el arrojar un dado).

Funciones continuas: cuando el número de valores x que puede tomar una variable aleatoria X es infinito se dice que dicha variable aleatoria es continua.

Por su naturaleza, a la hidrología le interesa estudiar a las funciones continuas de probabilidad, ya que los eventos ocurrentes no tienen un rango finito. (Por ejemplo los volúmenes de escurrimiento mensual de un río).

Estimar a qué función de distribución pertenece un fenómeno hidrológico es un problema complejo, puesto que sólo se tiene la muestra histórica disponible como herramienta, y un análisis previo de la misma dará la pauta para elegir la función más adecuada. En el análisis de dicha muestra pueden considerarse los siguientes criterios:

El tipo de fenómeno hidrológico: Se pueden elegir las funciones cuyos límites matemáticos sean compatibles con los límites impuestos por la naturaleza del fenómeno. [Llamas, 1993: p. 96].

Las características estadísticas de la muestra: La búsqueda de la función de distribución debe limitarse a aquellas familias de funciones de simetría equivalente.

La longitud de la muestra: Cuando se cuenta con muestras cortas o de reducido número de valores (50 o menos), sólo deben analizarse las funciones definidas por un número reducido de parámetros (de 1 a 3). Si la muestra es larga o lo suficientemente grande se deben utilizar funciones con mayor número de parámetros para dar mayor precisión debido a que el riesgo de la extrapolación a partir de una muestra limitada, es inversamente proporcional a la longitud de la muestra. [Llamas, 1993, p. 96].

Distribuciones de probabilidad

El comportamiento de las variables aleatorias discretas o continuas se describe con la ayuda de Distribuciones de Probabilidad. La variable se designa por mayúscula y un valor específico de ella por minúscula.

Por $P(x = a)$ se denota la probabilidad de que un evento asuma el valor a ; similarmente $P(a \leq x \leq b)$ denota la probabilidad de que un evento se encuentre en el intervalo (a,b) . Si conocemos la probabilidad $P(a \leq x \leq b)$ para todos los valores de a y b , se dice que conocemos la Distribución de Probabilidades de la variable x .

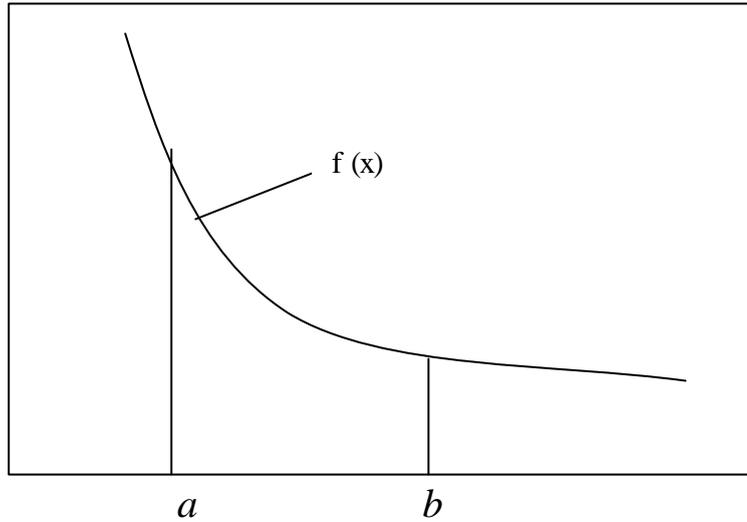


Figura 2.3. Distribución de la función de probabilidad de la variable x

$$P(a \leq x \leq b) = \int_a^b f(x)dx = A \quad (\text{Ec. 2.42})$$

Si x es un número dado y consideramos la probabilidad $P(X \leq x)$:

$$F(x) = P(X \leq x) \quad (\text{Probabilidad de no excedencia}). \quad (\text{Ec. 2.43})$$

y llamamos $F(x)$ la *función de distribución acumulada*.

Cuando el número de observaciones se incrementa, el tamaño de los intervalos decrece. $f(x)$ es la llamada *función de densidad de probabilidades* y tiene las siguientes características:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \quad (\text{Ec. 2.44})$$

$$P(a \leq x \leq b) = \int_a^b f(x)dx \quad (\text{Ec. 2.45})$$

$$\int_a^b f(x)dx = 0 \quad (\text{Ec. 2.46})$$

Lo que implica que las probabilidades se definen sólo como *áreas* bajo la función de densidad de probabilidad (FDP) entre límites finitos.

Momentos de las distribuciones

Las propiedades de las distribuciones pueden ser definidas completamente en términos de los momentos. Los momentos en estadística son similares a los momentos en física (rotación respecto al o

$$M^g = \int_{-\infty}^{\infty} x^g f(x) dx \text{ para la variable continua} \quad (\text{Ec. 2.47})$$

o respecto a la media (eje de rotación diferente al origen)

$$M^g = \int_{-\infty}^{\infty} (x - \mathbf{m})^g f(x) dx \text{ para la variable continua} \quad (\text{Ec. 2.48})$$

Parámetros estadísticos

Los estadísticos extraen información de una muestra, indicando las características de la población. Los principales estadísticos son los momentos de primer, segundo y tercer orden correspondiente a la media, varianza, y asimetría respectivamente.

Media μ :

Es el valor esperado de la variable misma. Primer momento respecto al origen. Muestra la tendencia central de la distribución

$$\mathbf{m} = \int_{-\infty}^{\infty} xf(x) dx \quad (\text{Ec. 2.49})$$

El valor estimado de la media a partir de la muestra es:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{Ec. 2.50})$$

Varianza σ^2 :

Mide la variabilidad de los datos. Es el segundo momento respecto a la media.

$$s^2 = \int_{-\infty}^{\infty} (x - m)^2 f(x) dx \quad (\text{Ec. 2.51})$$

El valor estimado de la varianza a partir de la muestra es:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (\text{Ec. 2.52})$$

En el cual el divisor es $n-1$ en lugar de n para asegurar que la estadística de la muestra *no sea sesgada*, es decir, que no tenga una tendencia, en promedio, a ser mayor o menor que el valor verdadero. Las unidades de la varianza son la media al cuadrado, la desviación estándar s es una medida de la variabilidad que tiene las mismas dimensiones que la media y simplemente es la raíz cuadrada de la varianza, se estima por s .

Coefficiente de variación

$$C_v = \frac{\sigma}{\mu}$$

es una medida adimensional de la variabilidad. Su estimado es:

$$C_n = \frac{s}{\bar{x}} \quad (\text{Ec. 2.53})$$

Coefficiente de asimetría γ

La distribución de los valores de una distribución alrededor de la media se mide por la asimetría. Se obtiene a partir del tercer momento alrededor de la media, dividiéndolo por el cubo de la desviación estándar para que sea adimensional.

$$E[(x - \mathbf{m})^3] = \int_{-\infty}^{\infty} (x - \mathbf{m})^3 f(x) dx \quad \text{tercer momento respecto a la media} \quad (\text{Ec. 2.54})$$

$$\mathbf{g} = \frac{1}{s^3} E[(x - \mathbf{m})^3] \quad (\text{Ec. 2.55})$$

Un estimativo del coeficiente de asimetría está dado por:

$$Cs = \frac{n \sum_{i=1}^n (x - \bar{x})^3}{(n-1)(n-2)s^3} \quad (\text{Ec. 2.56})$$

Período de Retorno

Cada espacio muestral tiene su propia función de distribución o de densidad de probabilidad, que normalmente no se conoce a priori. Cuando de ese espacio se extrae un grupo de datos (muestra) al azar, es razonable esperar que su función de distribución de probabilidad sea similar a la del espacio completo, en particular si la muestra es grande. Además, lo más razonable que se puede suponer en cuanto a la frecuencia de cada dato del grupo es que ésta sea, dentro del espacio muestral, igual a la observada.

El período de retorno o intervalo de recurrencia, T , se define como el tiempo promedio en el cual un evento de cierta magnitud va a ser igualado o superado por lo menos en una ocasión. Por ejemplo, si se tiene un evento de diseño con una probabilidad de ocurrencia $P = 0.25$ y tomando un proyecto con una vida útil de 100 años se tendría entonces que de los 100 años que dure el proyecto, en 25 se excedería el evento de diseño. Por lo tanto, el

tiempo promedio de recurrencia del evento es de $100/25 = 4$ años. Si P es la probabilidad, entonces:

$$T = 1/P \quad (\text{Ec. 2.57})$$

La probabilidad de que el evento no ocurra sería entonces:

$$F(x) = (1-P) = 1-1/T \quad (\text{Ec. 2.58})$$

La probabilidad de que no ocurra en n años de la vida útil del proyecto sería:

$$F(x) = (1 - 1/T)^n \quad (\text{Ec. 2.59})$$

y la probabilidad de que ocurra el evento en n años o el riesgo de que se presente es:

$$F(x) = [1 - (1 - 1/T)]^n \quad (\text{Ec. 2.60})$$

Cuando se tiene una lista de datos ordenados en forma decreciente con respecto al evento, el período de retorno se puede calcular como:

$$T = (n+1)/m \quad (\text{Ec. 2.61})$$

Donde:

n = número de datos.

m = número consecutivo de la lista para dicho evento.

Una vez que se asigna un período de retorno al gasto de diseño de la obra que se está analizando, es necesario hacer extrapolaciones a partir de los gastos máximos anuales que se hayan registrado para conocer el gasto de diseño que correspondería al período de retorno seleccionado. Esas extrapolaciones se logran mediante el uso de funciones de distribución de probabilidad, aquellas que más se ajusten a los registros con los que se cuenta.

De las las funciones de distribución de probabilidad más usadas en hidrología, estas son algunas de ellas para la determinación de caudales máximos:

1. Normal
2. Log-normal
3. Pearson III
4. log-Pearson III
5. Gumbel

Para la estimación de caudales máximos en este documento se seleccionaron la función Log-Normal porque es apropiada para variables aleatorias que cubren todo el rango de valores de los resultados posibles del experimento bajo análisis (volúmenes de escurrimiento mensual, por ejemplo); la función Gumbel debido que se desarrolló para el análisis de los valores extremos de dichos resultados, como los gastos máximos o mínimos anuales; y la función Log-Pearson III ya que ocupa un lugar intermedio. [Francisco Aparicio, 2001: p. 253]

2.4.1 Distribuciones de Probabilidad más usadas en hidrología para variables continuas.

Distribución normal

La distribución normal es una distribución simétrica en forma de campana, también conocida como Campana de Gauss. Aunque muchas veces no se ajusta a los datos

hidrológicos tiene amplia aplicación por ejemplo a los datos transformados que siguen la distribución normal.

La función de densidad de probabilidad normal se define como:

$$f(x) = \frac{1}{\sqrt{2\Pi s}} e^{-\frac{1}{2}\left(\frac{x-m}{s}\right)^2} \quad (\text{Ec. 2.62})$$

y la función de distribución de probabilidad normal como:

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\Pi s}} e^{-\frac{1}{2}\left(\frac{x-m}{s}\right)^2} dx \quad (\text{Ec. 2.63})$$

Donde:

x = variable aleatoria.

m = media de la población.

s = desviación estándar de la población.

Para resolver esta función se recurren a métodos numéricos para evaluarla, y para hacer esto más sencillo se ha asignado como variable estandarizada:

$$z = \frac{x-m}{s} \quad (\text{Ec. 2.64})$$

Que está normalmente distribuida con media cero y desviación estándar unitaria. Así la función principal queda como:

$$f(x) = f(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\Pi}} e^{-z^2/2} dz \quad (\text{Ec. 2.65})$$

Hay dos maneras de estimar $F(z)$, una es mediante una tabla que se ha publicado de dicha ecuación. Y la segunda es mediante fórmulas aproximadas, la función de densidad $f(z)$ se aproxima como:

$$f(z) = (a_0 + a_1 z^2 + a_2 z^4 + a_3 z^6)^{-1} \quad (\text{Ec. 2.66})$$

Donde:

$$a_0 = 2.490895$$

$$a_1 = 1.466003$$

$$a_2 = -0.024393$$

$$a_3 = 0.178257$$

Y la función de distribución sería la siguiente:

$$F(z) = H(z), \quad z > 0 \quad (\text{Ec. 2.67})$$

$$F(z) = 1 - H(z), \quad z < 0 \quad (\text{Ec. 2.68})$$

Donde:

$$H(z) = 1 - \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} (b_1 q + b_2 q^2 + b_3 q^3) \quad (\text{Ec. 2.69})$$

Siendo:

$$q = \frac{1}{1 + b_0 |z|} \quad (\text{Ec. 2.70})$$

$$b_0 = 0.33267$$

$$b_1 = 0.43618$$

$$b_2 = -0.12017$$

$$b_3 = 0.93730$$

Distribución Log-Normal

En esta función los logaritmos naturales de la variable aleatoria se distribuyen normalmente.

$$f(x) = \frac{1}{\sqrt{2\Pi}} \frac{1}{xb} e^{-\frac{1}{2}\left(\frac{\ln x - a}{b}\right)^2} \quad (\text{Ec. 2.71})$$

donde α y β son parámetros de la distribución, y por lo tanto son la media y la desviación estándar de los logaritmos de la variable aleatoria. La función de distribución de probabilidad es:

$$a = \sum_{i=1}^n \frac{\ln x_i}{n} \quad (\text{Ec. 2.72})$$

$$b = \left[\sum_{i=1}^n \frac{(\ln x_i - a)^2}{n} \right]^{1/2} \quad (\text{Ec. 2.73})$$

La distribución de probabilidad:

$$F(x) = \int_0^x \frac{1}{\sqrt{2\Pi}} \frac{1}{xb} e^{-\frac{1}{2}\left(\frac{\ln x - a}{b}\right)^2} dx \quad (\text{Ec. 2.74})$$

Al igual que en la distribución normal, se le asigna a "z" los siguientes valores:

$$z = \frac{\ln x - a}{b} \quad (\text{Ec. 2.75})$$

A causa de su asimetría pronunciada, la función Log- Normal se usa bastante en hidrología, sobre todo en el estudio de valores extremos.

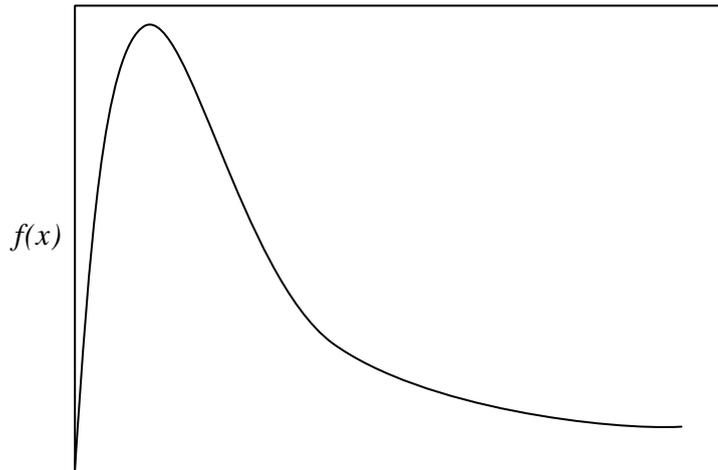


Figura 2.4. Función de densidad Log-Normal

Distribución Pearson III

Esta distribución ha sido una de las más utilizadas en hidrología. Como la mayoría de las variables hidrológicas son sesgadas, la función Gamma se utiliza para ajustar la distribución de frecuencia de variables tales como crecientes máximas anuales, Caudales mínimos, Volúmenes de flujo anuales y estacionales, valores de precipitaciones extremas y volúmenes de lluvia de corta duración. La función de distribución Gamma tiene dos o tres parámetros.

La función de densidad de probabilidad Pearson de tres parámetros (tipo III) se define como:

$$f(x) = \frac{1}{a_1 \Gamma(b_1)} \left\{ \frac{x-d_1}{a_1} \right\}^{b_1-1} e^{-\frac{x-d_1}{a_1}} \quad (\text{Ec. 2.76})$$

donde $\alpha_1, \beta_1, \delta_1$ son los parámetros de la función y $\Gamma(\beta_1)$ es la función de Gamma. Los parámetros $\alpha_1, \beta_1, \delta_1$ se evalúan a partir de n datos medidos, mediante el siguiente sistema de ecuaciones:

$$\bar{x} = \alpha_1 \beta_1 + \delta_1 \quad (\text{Ec. 2.77})$$

$$S^2 = \alpha_1^2 \beta_1 \quad (\text{Ec. 2.78})$$

$$\gamma = 2 / (\beta_1)^{0.5} \quad (\text{Ec. 2.79})$$

donde \bar{x} es la media de los datos, S^2 su variancia y γ su coeficiente de sesgo, que se define como:

$$g = \frac{\sum_{i=1}^n (x_i - \bar{x})^3 / n}{S^3} \quad (\text{Ec. 2.80})$$

La función de distribución de probabilidad es:

$$F(x) = \frac{1}{a_1 \Gamma(b_1)} \int_0^x e^{-\left(\frac{x-d_1}{d_1}\right)} \left(\frac{x-d_1}{d_1}\right)^{b_1-1} dx \quad (\text{Ec. 2.81})$$

y sustituyendo

$$y = \frac{x-d_1}{a_1} \quad (\text{Ec. 2.82})$$

La ecuación quedaría

$$F(y) = \frac{1}{\Gamma(b_1)} \int_0^y y^{b_1-1} e^{-y} dy \quad (\text{Ec. 2.83})$$

Siendo la anterior una función ji cuadrada con $2b_1$ grados de libertad y $x^2 = 2y$:

$$F(y) = F(x^2|n) = F_{x^2}(2y|2b_1) \quad (\text{Ec. 2.84})$$

Esta función se usa solo cuando $\beta_1 = n/2$, donde n es un entero positivo cualquiera. Como 2β es no entero, puede tomarse como el entero más próximo o bien interpolar en la tabla del anexo B. Cuando $\beta < 0.3$, se tendrá que acudir a tablas de la función de distribución de Gamma de un parámetro. Cuando δ da valores extraños, como negativos o muy grandes, es recomendable fijar dicho valor de d a ojo, estimándolo, como la ordenada al origen en una gráfica de gasto contra periodo de retorno.

Distribución Log-Pearson III

Si los logaritmos Y de una variable aleatoria X se ajustan a una distribución Pearson tipo III, se dice que la variable aleatoria X se ajusta a una distribución Log Pearson Tipo III. Esta distribución es ampliamente usada en el mundo para el análisis de frecuencia de caudales máximos. Esta se trabaja igual que para la Pearson Tipo III incluso el uso de la distribución ji cuadrada, pero con X_y y S_y como la media y desviación estándar de los logaritmos de la variable original X .

Distribución Gumbel

Supóngase que se tienen N muestras cada una con n eventos. Si se selecciona el máximo x de los n eventos de cada muestra, es posible demostrar que, a medida que aumenta n , la función de probabilidad de x (probabilidad de no excedencia) tiende a:

$$F(x) = e^{-e^{-a(x-b)}} = P(X=x) \quad (\text{Ec. 2.85})$$

La función de densidad de probabilidad es entonces:

$$f(x) = ae^{-a(x-b)-e^{-a(x-b)}} \quad (\text{Ec. 2.86})$$

y la probabilidad de excedencia es:

$$Pt = [1 - F(x)] \quad (\text{Ec. 2.87})$$

Donde α y β son los parámetros de la función.

Los valores de α y β :

$$a = \frac{p}{s\sqrt{6}} \quad (\text{Ec. 2.88})$$

$$b = m - g/a \quad (\text{Ec. 2.89})$$

Donde γ es la constante de Euler y es igual a:

$$\gamma = 0.577215 \quad (\text{Ec. 2.90})$$

2.4.2 Pruebas de bondad de ajuste

Para determinar que tan adecuado es el ajuste de los datos a una distribución de probabilidades se han propuesto una serie de pruebas estadísticas que determinan si es adecuado el ajuste. Estos son análisis estadísticos y como tal se deben entender, es decir, no se puede ignorar el significado físico de los mismos.

En la teoría estadística, las pruebas de bondad del ajuste más conocidas son la Smirnov - Kolmogorov y la Chi - Cuadrado, las cuales son las que se utilizarán en nuestro análisis y que se describen a continuación:

Prueba Smirnov-Kolmogorov

La prueba de bondad de ajuste estadístico Smirnov - Kolmogorov considera la desviación de la función de distribución de probabilidades de la muestra $P(x)$ de la función de probabilidades teórica, escogida $Po(x)$ tal que:

$$Dn = \max(P(x) - Po(x)) \quad (\text{Ec. 2.91})$$

La prueba requiere que el valor D calculado con la expresión anterior sea menor que el valor tabulado d_{crit} para un nivel de probabilidad (significancia) requerido.

Esta prueba es fácil de realizar y comprende las siguientes etapas:

El valor estadístico D es la máxima diferencia entre la función de distribución acumulada de la muestra y la función de distribución acumulada teórica escogida.

Se fija el nivel de probabilidad (nivel de significancia) α , valores de 0.05 y 0.01 son los más usuales.

El valor crítico d_{crit} de la prueba debe ser obtenido de la tabla 2.6, el cual está en función de α y n , pues depende del nivel de significancia y del número de datos.

Si el valor calculado D es menor que el d_{crit} , la distribución escogida se debe aceptar.

Por el contrario, si el valor calculado D es mayor que el d_{crit} , la distribución escogida se debe rechazar.

Tamaño de la muestra	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
5	0.51	0.56	0.67
10	0.37	0.41	0.49
15	0.30	0.34	0.40
20	0.26	0.29	0.35
25	0.24	0.26	0.32
30	0.22	0.24	0.29
40	0.19	0.21	0.25
n grande	$1.22 / (n)^{1/2}$	$1.36 / (n)^{1/2}$	$1.63 / (n)^{1/2}$

Tabla 2.6. Valores críticos d_{crit} para la prueba Smirnov-Kolmogorov de Bondad de ajuste. [Aparicio, 1993, p. 289].

Prueba ji-Cuadrado

Una medida de la discrepancia entre las frecuencias observadas (f_o) y las frecuencias calculadas (f_c) por medio de una distribución teórica esta dada por el estadístico χ^2

$$\chi^2 = \sum_{i=1}^k \frac{(f_o - f_c)^2}{f_c} \quad (\text{Ec. 2.92})$$

en donde $\sum f_o = \sum f_c$

Si el estadístico $\chi^2=0$ significa que las distribuciones teórica y empírica ajustan exactamente, mientras que si el estadístico $\chi^2>0$, ellas difieren. La distribución del estadístico χ^2 se puede asimilar a una distribución Chi-cuadrado con $(k-1-m)$ grados de libertad, donde k es el número de intervalos y m es el número de los parámetros de la distribución teórica. La función χ^2 se encuentra tabulada tal y como se mostró en la figura 2.6. Supóngase que una hipótesis H_o es aceptar que una distribución empírica se ajusta a una distribución Normal. Si el valor calculado de χ^2 por la ecuación anterior es mayor que algún valor crítico de χ^2 , con niveles de significancia α de 0.05 y 0.01 (el nivel de confianza es $1-\alpha$) se puede decir que las frecuencias observadas difieren significativamente de las frecuencias esperadas (o calculadas) y entonces la hipótesis H_o se rechaza, si ocurre lo contrario entonces se acepta.

2.4.3 Selección de la Función de Distribución de mejor ajuste.

Al aplicar cada una de las funciones de distribución a los datos de los registros de cada estación es posible que se observen que existen diferencias que son apreciables entre las mismas. Debido a esto es necesario seleccionar con cuidado la función a utilizar, ya que una mala selección de cualquiera de las funciones puede traducirse, para el ingeniero o entidad que utilizará en el futuro esta información, en una estructura sobrediseñada y costosa o subdiseñada y peligrosa [Aparicio, 2001, p. 270].

La selección de la función de mejor ajuste a utilizar, con el propósito de calcular los valores de Caudales Máximos en cada una de las estaciones analizadas, se realiza de la siguiente manera:

En primera instancia se aplican ambas pruebas de bondad de ajuste (ji-cuadrado y Smirnov-Kolmogorov) a las distribuciones obtenidas de cada función para los registros de cada estación.

Luego se calcula el valor del parámetro estadístico χ^2 para la prueba ji-cuadrado, y el parámetro D para Smirnov-Kolmogorov, para así compararlos con los valores estadísticos críticos $\chi^2_{1-\alpha, k-1-m}$ y d_{crit} respectivamente, los cuales se encuentran en función del nivel de significancia a usado. Para este último se seleccionó el valor de $\alpha = 0.05$, es decir se acepta una incertidumbre en el modelo de un 5 por ciento, o lo que es lo mismo se requiere de un 95 por ciento de certeza para aceptar el modelo o la hipótesis.

Ya comparados los parámetros calculados con los valores críticos se procede a concluir si se acepta o se rechaza la función aplicada, y a escoger cual sería la preferible según las pruebas aplicadas.

Para seleccionar finalmente la función a utilizar se tabulan los resultados de las pruebas aplicadas, calificando las funciones según el orden de preferencia, otorgando un valor de 1 a la de "mejor" ajuste y creciendo este valor hasta la de "peor" ajuste, o definitivamente rechazando la función.