



# HidroCalc 2.1

The screenshot displays the HidroCalc 2.1 software interface with several panels:

- Block table (SCS):** A table with columns for 'PP', 'Intensidad (mm/h)', and 'PP.acum. (mm)'. It lists various precipitation intensity values and their cumulative totals.
- Soils table:** A table with columns for 'Intensidad (mm/h)', 'Factor SCS', 'Potencial (mm/h)', 'Potencial parcial (mm/h)', and 'basal (mm/h)'. It provides soil-related parameters for different intensity levels.
- Gráfico XY (Curva IDF - SCS):** A line graph showing the relationship between precipitation intensity (mm/h) on the y-axis and duration (min) on the x-axis. The curve shows a sharp decrease in intensity as duration increases.
- Gráfico de series (Histograma - SCS):** A histogram showing the frequency distribution of precipitation intensity (mm/h) on the x-axis and frequency on the y-axis.

## Manual de Referencias

Unu Ingenieros

Protegido por la ley de derechos de autor ©



# **HidroCalc 2**

## **Manual de Referencias**

**2021**

## **MANUAL DE REFERENCIAS DE HIDROCALC 2.1**

**©Unu Ingenieros**

☎(+51) 961304698

[contacto@unuingenieros.com](mailto:contacto@unuingenieros.com)

[unu.ingenieros@gmail.com](mailto:unu.ingenieros@gmail.com)

<http://www.unuingenieros.com/hidrocalc2.html>

Primera edición: Junio 2021

Autor: Luis Félix Vásquez Paredes

Prohibida la reproducción parcial o total, por cualquier medio o método de este manual sin previa autorización de la empresa Unu Ingenieros E.I.R.L.

**Copyright ©2021**

Impreso en Perú

## Términos y condiciones

Al hacer uso del presente Manual de Referencias, acepta los términos y condiciones aclarados por Unu Ingenieros.

- Unu Ingenieros aclara que los métodos utilizados en la programación del software HidroCalc 2, no son producto del desarrollo de investigaciones propias de la empresa y que han sido producto de la recopilación bibliográfica de investigaciones hechas por los diferentes autores citados en este Manual de Referencias a quienes Unu Ingenieros les agradece sus notables contribuciones a la ciencia y la ingeniería hidrológica.
- Unu Ingenieros ha utilizado artículos de divulgación científica, libros e información disponible en fuentes de uso libre para el desarrollo del programa HidroCalc 2, aclarando que no se ha infringido ningún tipo de violación a la Ley de Derechos de autor o los tratados internacionales del ©Copyright.
- Los autores, las investigaciones y los métodos empleados por HidroCalc 2 para resolver los problemas hidrológicos, para los cuales se realizó la programación del software, son citados en sus fuentes respectivas en el presente Manual de Referencias.
- Los usuarios que hagan uso de este manual en investigaciones u otras actividades académicas deben citar los métodos, ecuaciones, imágenes y otros, de sus fuentes respectivas ya que Unu Ingenieros no es autor de ninguno que no se especifique explícitamente.
- Unu Ingenieros se reserva el derecho de compartir sus algoritmos de cálculo y el código computacional empleado en HidroCalc 2 en los casos en que estos sí sean de autoría de los creadores y de los cuales no exista difusión previa al lanzamiento del software.

# Índice general

<b>1. Introducción</b>	<b>6</b>
1.1. Aplicaciones de HidroCalc 2 . . . . .	7
1.2. Ventajas de trabajar con HidroCalc 2 . . . . .	7
1.3. Desarrollo futuro . . . . .	8
<b>2. Cuencas hidrográficas</b>	<b>9</b>
2.1. Cuencas hidrográficas . . . . .	9
2.1.1. Delimitación de cuencas hidrográficas . . . . .	9
2.2. Parámetros morfológicos . . . . .	11
2.2.1. Área ( $A$ ) . . . . .	11
2.2.2. Perímetro ( $P$ ) . . . . .	11
2.2.3. Longitud del cauce principal ( $L$ ) . . . . .	12
2.2.4. Coeficiente o índice de compacidad ( $Kc$ ) . . . . .	12
2.2.5. Factor de forma ( $R_f$ ) . . . . .	13
2.2.6. Pendiente de la cuenca ( $S_c$ ) . . . . .	13
2.2.7. Altitud media ( $H$ ) . . . . .	14
2.2.8. Pendiente del cauce principal ( $S$ ) . . . . .	15
2.2.9. Ancho promedio ( $B$ ) . . . . .	16
2.2.10. Altitud más frecuente . . . . .	16
2.2.11. Coeficiente de masividad ( $C_m$ ) . . . . .	16
2.2.12. Coeficiente orográfico ( $C_o$ ) . . . . .	17
2.2.13. Rectángulo equivalente . . . . .	17
2.2.14. Orden de la cuenca . . . . .	17
2.2.15. Relación de confluencias ( $r_c$ ) . . . . .	19
2.2.16. Relación de longitudes ( $r_L$ ) . . . . .	20
2.2.17. Densidad de drenaje ( $D_d$ ) . . . . .	20
2.2.18. Frecuencia de ríos ( $F_r$ ) . . . . .	21
2.2.19. Tiempo de concentración ( $T_c$ ) . . . . .	21
2.2.20. Curva hipsométrica . . . . .	22
2.2.21. Frecuencia de altitudes . . . . .	22

2.2.22. Orientación de la cuenca . . . . .	23
2.2.23. Centro de gravedad ( <i>C.G</i> ) . . . . .	23
<b>3. Hidrología estadística</b>	<b>25</b>
3.1. Conceptos básicos de estadística . . . . .	25
3.1.1. Términos básicos . . . . .	26
3.1.2. Variable aleatoria . . . . .	26
3.1.3. Probabilidad . . . . .	26
3.1.4. Periodo de retorno . . . . .	26
3.1.5. Probabilidad empírica . . . . .	27
3.2. Análisis de frecuencias . . . . .	28
3.2.1. Distribuciones teóricas de probabilidad . . . . .	28
3.2.2. Estimación de parámetros de distribuciones . . . . .	35
3.2.3. Bondad de ajuste: Smirnov-Kolmogorov . . . . .	39
3.3. Análisis de consistencia y tendencia . . . . .	42
3.3.1. Valores atípicos . . . . .	42
3.3.2. Estadística paramétrica y no paramétrica . . . . .	43
3.3.3. Homogeneidad en la media . . . . .	44
3.3.4. Homogeneidad en la varianza . . . . .	47
3.3.5. Tendencias . . . . .	49
3.4. Análisis de regresión . . . . .	50
3.4.1. Regresión lineal simple . . . . .	50
3.4.2. Regresión múltiple . . . . .	51
<b>4. Precipitación</b>	<b>52</b>
4.1. Curvas IDF . . . . .	53
4.2. Hietogramas . . . . .	55
<b>5. Procesos precipitación - escorrentía</b>	<b>56</b>
5.1. Caudales máximos . . . . .	56
5.1.1. Fórmulas empíricas . . . . .	56
5.1.2. Hidrogramas sintéticos . . . . .	60
5.2. Caudales mensuales . . . . .	64
5.2.1. Modelo de Témez . . . . .	64
5.2.2. Modelo de Lutz Scholz . . . . .	66
<b>6. Redes Neuronales Artificiales</b>	<b>73</b>

# Capítulo 1

## Introducción

El proyecto **HidroCalc**, propiedad de **Unu Ingenieros E.I.R.L.**, ha sido desarrollado con la finalidad de brindar un software de apoyo en el análisis de procesos hidrológicos. **HidroCalc** pretende dar al hidrólogo una herramienta que le permita enfocar su análisis en el criterio y no en las limitaciones de los cálculos complejos.

El software se ha desarrollado en lenguaje **Visual Basic.Net**. Cada parte de la herramienta de proceso o motor de cálculo ha sido escrita por Unu Ingenieros mediante las funciones básicas del lenguaje y está basada en diferentes métodos numéricos, algebraicos y estadísticos, necesarios para la resolución de problemas hidrológicos.

Los módulos del programa han sido añadidos en base a la experiencia de Unu Ingenieros en proyectos de consultoría en hidrología superficial. En esta segunda versión, no se ha contemplado la programación de módulos basados en las investigaciones de la empresa, reservando esta tarea para la creación de la tercera versión.

Este manual, señala los **fundamentos teóricos y los conceptos básicos de hidrología superficial** utilizados en el desarrollo del programa, se citan las fuentes de información consultada.

## 1.1. Aplicaciones de HydroCalc 2

HydroCalc 2 puede ser utilizado en los siguientes procesos:

- Determinación de parámetros morfológicos de cuencas
- Resumen de datos hidrometeorológicos
- Análisis de frecuencias de variables hidrológicas
- Análisis de valores atípicos
- Análisis de doble masa
- Homogeneidad de información hidrometeorológica
- Análisis de tendencias
- Estimación de caudales máximos en cuencas pequeñas y medianas
- Generación de series sintéticas de caudales mensuales
- Estudio de precipitaciones para elaboración de curvas IDF e hietogramas de diseño
- Entrenamiento de perceptrón multicapa

## 1.2. Ventajas de trabajar con HydroCalc 2

Las ventajas de trabajar con HydroCalc 2 son:

- Interfaz de fácil uso y rápida curva de aprendizaje
- Simplificación de cálculos engorrosos que pueden conducir a errores
- Procesamiento de fondo de cálculos extensos
- Reporte detallado de cálculos
- Gráficos personalizables de series de tiempo y dispersiones
- Compatibilidad con hojas de cálculo de MS Excel
- Grabado y carga de archivos de configuración
- Grabado de cuencas en base de datos
- Grabado de variables meteorológicas en base de datos

### **1.3. Desarrollo futuro**

El desarrollo futuro de HidroCalc contempla la incorporación de hallazgos científicos de las investigaciones en curso de Unu Ingenieros en el campo de la hidrología superficial. A corto (versión 2.2) y largo plazo (versión 3), el programa incluirá:

- Herramienta para extensión de variables
- Herramienta para visualización de variables, tipo “dashboard”
- Herramienta de completación automática
- Herramienta para análisis regional de frecuencias
- Módulo de SIG para visualización
- Licencia portable

## Capítulo 2

# Cuencas hidrográficas

Este capítulo aborda los conceptos elementales de delimitación de cuencas hidrográficas y obtención de las características morfológicas del terreno (parámetros morfológicos).

## 2.1. Cuencas hidrográficas

### 2.1.1. Delimitación de cuencas hidrográficas

Partiendo de la definición de cuenca hidrográfica, como la región geográfica en la que toda el agua precipitada drena a un punto en común (Ordoñez, 2011), se pueden establecer algunos criterios para realizar la delimitación manual de cuencas hidrográficas partiendo de un modelo digital de elevación, tales como:

- Definir la red de drenaje, partiendo del cauce principal.
- Ubicar los puntos altos que están definidos por las curvas de nivel.
- Unir los puntos altos definidos anteriormente cortando perpendicularmente las curvas de nivel.
- En cualquier punto del terreno la línea divisoria debe ser el punto de mayor altitud excepto cerros o puntos altos que se encuentran dentro de la cuenca
- La línea divisoria nunca debe cortar un río, quebrada o arroyo.

En la actualidad el uso de sistemas de información geográfica simplifican el proceso de delimitación de cuencas hidrográficas mediante el uso de

herramientas o algoritmos especialmente creados para tal fin. En líneas generales existe un procedimiento en común empleado por estos sistemas para realizar la delimitación de cuencas hidrográficas, tal como lo muestra el flujograma de la figura 2.1. La calidad de resultados de este proceso de delimitación dependerá siempre en mayor parte de la información de entrada, es decir, del modelo digital de elevación, recalcando que entre mejor resolución posea el modelo, mayor precisión tendrá la delimitación.

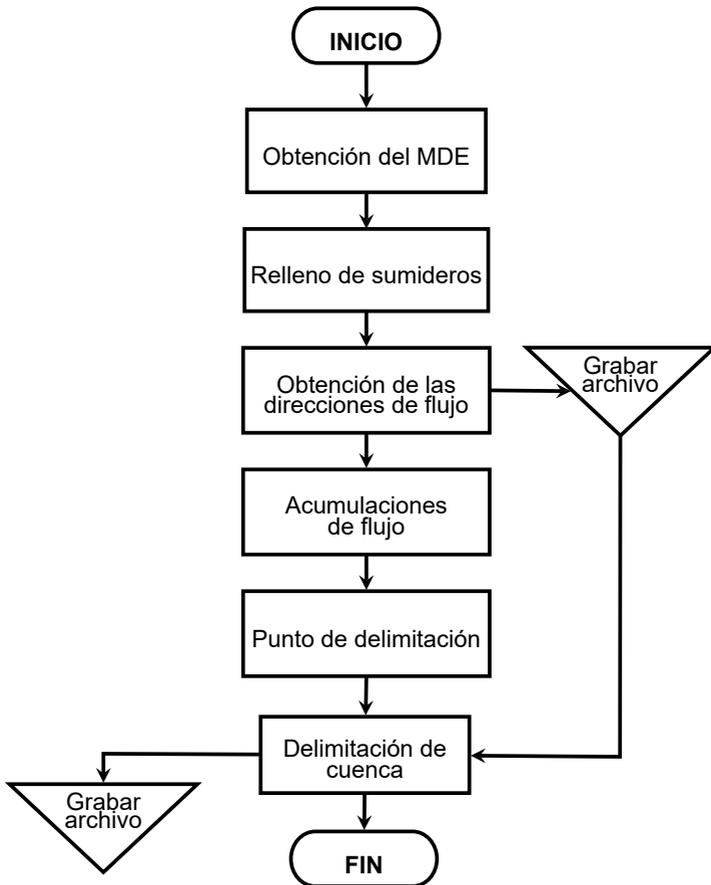


Figura 2.1: Diagrama de flujo: Delimitación de cuencas con SIG

Fuente: (ESRI, 2018)

Elaboración: Propia.

## 2.2. Parámetros morfológicos

Denominados por algunos autores como morfométricos, geomorfológicos o fisiográficos. Son aquellos que sirven para caracterizar una cuenca hidrográfica, es decir, su estructura física y territorial (o en esencia la *forma de la cuenca*) con el fin de establecer las posibilidades y limitaciones de sus Recursos Naturales pero también para identificar los problemas presentes y potenciales (Ordoñez, 2011).

A continuación se describen los principales.

### 2.2.1. Área ( $A$ )

Representa la proyección horizontal del área de drenaje en la que todo lo precipitado fluye hacia un mismo punto. El área de la cuenca está definida por el espacio delimitado por la curva del perímetro. El área es probablemente la característica morfométrica e hidrológica más importante; el tamaño relativo de estos espacios define o determina el nombre que adoptara el lugar según su área. Puede medirse en *ha*,  $m^2$ ,  $km^2$  o cualquier unidad de área.

Cuadro 2.1: Clasificación de cuencas según el área de drenaje

Área	Clasificación
< 5	Unidad
5 – 20	Sector
20 – 100	Microcuenca
100 – 300	Subcuenca
> 300	Cuenca

Fuente: (Moreno Grande et al., 2015)

### 2.2.2. Perímetro ( $P$ )

Este parámetro mide el contorno de la cuenca como una línea continua, es un parámetro importante ya que al igual que el área permite inferir sobre la forma de la cuenca. Su forma incide en el tiempo de respuesta de la cuenca, es decir, en el tiempo de recorrido de las aguas a través de la red de drenaje, y por consiguiente, en la forma del hidrograma resultante de una lluvia dada. Lleva unidades de longitud, tal como *km* o *m*.

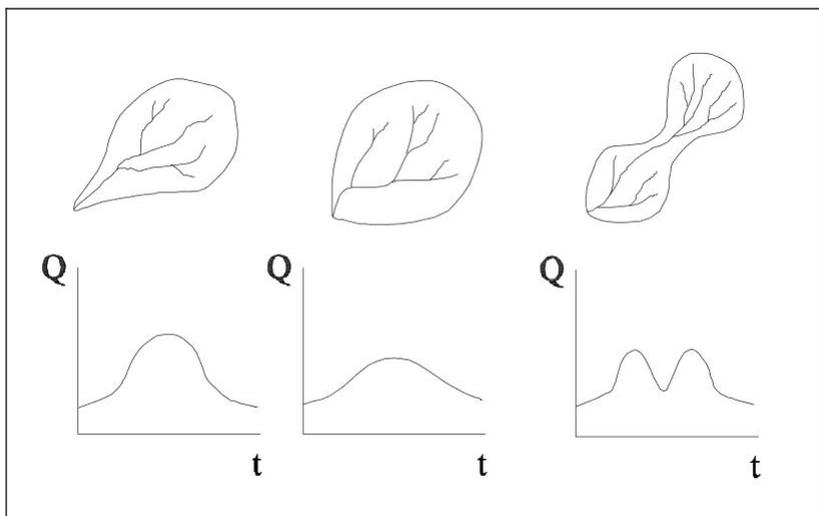


Figura 2.2: Forma de la cuenca e hidrograma probable

Fuente: (Ordoñez, 2011)

### 2.2.3. Longitud del cauce principal ( $L$ )

Este parámetro representa la longitud del río principal de la cuenca, donde van a drenar todos los afluentes y quebradas. En la elección de dicho cauce, se tiene en cuenta en cada caso, tanto la longitud del mismo como su desnivel, buscando siempre el curso que presenta en lo posible el máximo para ambos valores (Ordoñez, 2011; Fattorelli and Fernández, 2007). Lleva también unidades de longitud, tal como  $km$  o  $m$ .

### 2.2.4. Coeficiente o índice de compacidad ( $K_c$ )

También denominado índice de Graveliús, definido como la relación entre el perímetro  $P$  de la cuenca y el perímetro de un círculo de área  $A$  de la cuenca hidrográfica, es decir:

$$K_c = \frac{P}{2\sqrt{\pi A}} = \frac{0.28P}{\sqrt{A}} \quad (2.1)$$

donde  $K_c$  es coeficiente o índice de compacidad,  $A$  corresponde al área de la cuenca (en  $km^2$ ) y  $P$  es el perímetro de la cuenca (en  $km$ ) (Ordoñez, 2011).  $K_c$  es adimensional.

Se pueden establecer algunos criterios de clasificación según los valores del coeficiente de compacidad, tal como se ve en el cuadro 2.2.

Cuadro 2.2: Tipos de cuenca según el valor de  $K_c$ 

$K_c$	Tipo de cuenca
1.00 - 1.25	Cuenca redonda a oval redonda
1.25 - 1.50	Cuenca oval redonda a oval oblonga
1.50 - 1.75	Cuenca de oval oblonga a rectangular oblonga

Fuente: (Moreno Grande et al., 2015)

### 2.2.5. Factor de forma ( $R_f$ )

Horton sugiere un factor adimensional de forma, que puede calcularse de la siguiente expresión:

$$R_f = \frac{A}{L^2} \quad (2.2)$$

donde  $A$  corresponde al área de la cuenca (en  $km^2$ ) y  $L$  a la longitud del cauce principal (en  $km$ ).

También se puede caracterizar la cuenca basándonos en el valor de  $R_f$ , tal como se ve en el cuadro 2.3.

Cuadro 2.3: Tipos de cuenca según el valor de  $R_f$ 

Valores aproximados	Forma de la cuenca
< 0.22	Muy alargada
0.22 – 0.30	Alargada
0.30 – 0.37	Ligeramente alargada
0.37 – 0.45	Ni alargada, ni ensanchada
0.45 – 0.60	Ligeramente ensanchada
0.60 – 0.80	Ensanchada
0.80 – 1.20	Muy ensanchada
> 1.20	Rodeando el desagüe

Fuente: (Ordoñez, 2011)

### 2.2.6. Pendiente de la cuenca ( $S_c$ )

Este parámetro puede brindar un indicio de la velocidad media del flujo en la cuenca y su poder de arrastre y erosión. Puede ser calculado con diversos métodos, aquí se menciona el criterio de Alvord (ecuación 2.3), el cual consiste en calcular como media ponderada de las pendientes de todas las superficies elementales de la cuenca en las que la línea máxima de pendiente se mantiene constante (Ibáñez et al., 2011).

$$S_c = c \frac{\sum L_i}{A} \quad (2.3)$$

donde  $L_i$  representa la longitud de cada curva de nivel dentro de la cuenca,  $c$  es la equidistancia entre curvas y  $A$  es el área de la misma. Puede expresarse en %,  $m/m$  o  $m/km$ .

Teniendo en cuenta el área también podemos establecer algunos criterios de clasificación.

Cuadro 2.4: Tipos de cuenca según el valor de  $S_c$

Pendiente media en %	Tipo de relieve
0 – 3	Plano
3 – 7	Suave
7 – 12	Medianamente accidentado
12 – 20	Accidentado
20 – 35	Fuertemente accidentado
35 – 50	Muy fuertemente accidentado
50 – 75	Escarpado
> 75	Muy escarpado

Fuente: (Moreno Grande et al., 2015)

### 2.2.7. Altitud media ( $H$ )

Este parámetro representa la altitud promedio de la cuenca medido en  $m.s.n.m.$  y brinda indicios acerca de la climatología y los hábitats de la cuenca, tiene especial interés en zonas montañosas y es de vital importancia en los procesos de transposición de información de una cuenca a otra. Se puede calcular de la curva hipsométrica siendo el valor correspondiente a la ordenada media. En la curva hipsométrica el 50 % del área está situada por debajo de ella y el 50 % por encima. Puede calcularse también a partir de la expresión 2.4.

$$H = \frac{\sum H_i A_i}{A} \quad (2.4)$$

donde  $H_i$  es la  $i$ -ésima altura en  $m.s.n.m.$  correspondiente al área parcial  $A_i$  y  $A$  es el área total de la cuenca en  $km^2$ .

De igual manera, (Moreno Grande et al., 2015) consideran interpretar la altitud media en función del cuadro 2.5.

Cuadro 2.5: Interpretación de la Altitud media -  $H$ 

Altitud media - $H$	Interpretación
800 – 1000m.s.n.m.	Macro térmico, con las temperaturas siempre elevadas y constantes
1000 – 2750m.s.n.m.	Meso térmico o piso templado
2750 – 4700m.s.n.m.	Micro térmico o piso páramo
A partir de 4700m.s.n.m.	Gélido o de nieves perpetuas

Fuente: (Moreno Grande et al., 2015)

### 2.2.8. Pendiente del cauce principal ( $S$ )

Representa la relación existente entre el desnivel altitudinal del cauce y su longitud. A medida que la pendiente aumenta, mayor es la probabilidad de crecidas, por lógica, a mayor pendiente, mayores velocidades de escorrentía. Este parámetro, a diferencia de los demás, puede presentar mayores alteraciones en el transcurso del tiempo debido a factores como la erosión en cotas altas o sedimentación en cotas bajas. Su adecuada interpretación puede permitirnos realizar planes de manejo de cuencas y una adecuada colocación de infraestructura hidráulica.

En general, la pendiente del cauce principal varía a lo largo de su longitud, por ello también es recomendable acompañarlo de un gráfico de perfil longitudinal y estimar la pendiente de manera ponderada. Recomienda la siguiente metodología (Ortiz, 2008), indicando la ecuación 2.5.

$$S = \left[ \frac{\sum_{i=1}^n L_i}{\sum_{i=1}^n \left( \frac{L_i^2}{S_i} \right)^{1/2}} \right]^2 \quad (2.5)$$

donde  $S_i$  corresponde a la pendiente parcial de cada tramo,  $L_i$  es la longitud de cada tramo de pendiente  $S_i$  y  $n$  el número de tramos en que se ha dividido el perfil del cauce.

Para efectos de aplicación de la ecuación 2.5, se divide el relieve del cauce en tramos de similar pendiente. Si se divide el perfil del cauce principal en  $n$  tramos iguales, la ecuación 2.5 se transforma en:

$$S = \left[ \frac{n}{\sum_{i=1}^n \frac{1}{S_i^{1/2}}} \right]^2 \quad (2.6)$$

Asimismo, (Ortiz, 2008) recomienda emplear las ecuaciones con cautela en vista de que su deducción obedece a criterios de la hidráulica de canales. También puede expresarse en %,  $m/m$  o  $m/km$ .

### 2.2.9. Ancho promedio ( $B$ )

Es la razón entre el área de la cuenca y la longitud del cauce principal, indicando por lo tanto que valores mayores a 1 presentan formas no-alargadas y valores cercanos a 0 el caso opuesto. Tal como lo muestra la ecuación 2.7 (Jardí Porqueras, 1985).

$$B = \frac{A}{L} \quad (2.7)$$

donde  $A$  es el área de la cuenca en  $km^2$  y  $L$  es la longitud del cauce principal en  $km$ .

### 2.2.10. Altitud más frecuente

Esta dada por el máximo valor en porcentaje de la curva de frecuencia de altitudes, de igual manera se presenta en unidades de  $m.s.n.m.$  y expresa la altitud que más persiste en la cuenca, esto es, la altitud que representa una mayor área en porcentaje del total de la cuenca. Al igual que la altitud media puede darnos indicios acerca de la climatología de la cuenca (Ordoñez, 2011).

### 2.2.11. Coeficiente de masividad ( $C_m$ )

Es la relación entre la altitud media del relieve y la superficie proyectada (ecuación 2.8). Es adimensional.

$$C_m = H/A \quad (2.8)$$

### 2.2.12. Coeficiente orográfico ( $C_o$ )

Nos permite estudiar la similitud dinámica de dos cuencas, no depende del tamaño del sistema. También es adimensional.

$$C_o = H^2/A \quad (2.9)$$

Este parámetro es de gran importancia puesto que da indicios del relieve de la cuenca y puede ayudarnos a entender la degradación del suelo en la misma. Combina la altura y el coeficiente de masividad de la cuenca, por lo que influye en la energía potencial del agua, de gran utilidad para evaluar el potencial hidroenergético.

### 2.2.13. Rectángulo equivalente

Consiste en hacer un rectángulo equivalente en área y perímetro a los de la cuenca y por lo tanto con el mismo coeficiente de Graveliús ( $K_c$ ). Sirve para comparar la influencia de las características de la cuenca sobre la escorrentía, las curvas de nivel se transforman en rectas paralelas al lado menor del rectángulo, y el drenaje de la cuenca queda convertido en el lado menor del rectángulo. Para ubicar las curvas de nivel se utiliza una regla de tres, asignando a la mayor área acumulada el valor del lado mayor del rectángulo equivalente y en base a esto se calculan las siguientes longitudes de acuerdo a su correspondiente área, tal como lo muestra la figura 2.3. Las ecuaciones 2.10 y 2.11 se utilizan para determinar los lados mayor y menor del rectángulo equivalente (Moreno Grande et al., 2015).

$$L = \frac{K_c \sqrt{A}}{1.12} \left[ 1 + \sqrt{1 - \frac{1.12^2}{K_c^2}} \right] \quad (2.10)$$

$$l = \frac{K_c \sqrt{A}}{1.12} \left[ 1 - \sqrt{1 - \frac{1.12^2}{K_c^2}} \right] \quad (2.11)$$

donde  $L$  y  $l$  son los lados mayor y menor del rectángulo equivalente, respectivamente, y  $A$  es el área de la cuenca

### 2.2.14. Orden de la cuenca

Este parámetro refleja el grado de ramificación o bifurcación dentro de la cuenca. Para determinarlo existen varias metodología, tales como el criterio de Horton y el criterio de Strahler (figura 2.4).

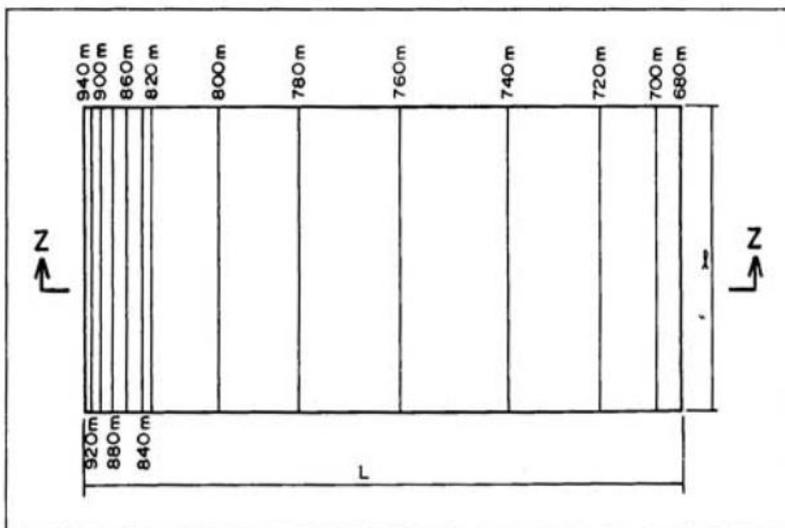


Figura 2.3: Rectángulo equivalente  
 Fuente: (Moreno Grande et al., 2015)

- **Horton.** Se realiza de las siguientes premisas: (i) los cauces de primer orden son los que no tienen tributarios; (ii) los cauces de segundo orden se forman en la unión de dos cauces de primer orden y, en general, los cauces de orden  $n$  se forman cuando dos cauces de orden  $n - 1$  se unen; (iii) cuando un cauce se une con un cauce de orden mayor, el cauce resultante hacia aguas abajo retiene el mayor orden; (iv) el orden de la cuenca es el mismo que el de su cauce de salida (Ordoñez, 2011).
- **Strahler.** Es muy parecido al de Horton, con la diferencia de que en el esquema de Strahler, un mismo río puede tener segmentos de distinto orden a lo largo de su curso, en función de los afluentes que llegan en cada tramo. El orden no se incrementa cuando a un segmento de un determinado orden confluye uno de orden menor. Esta falta no acomoda la distribución de tributarios de más baja orden que desembocan en un río de orden más alta, sin embargo, es el criterio más utilizado (Ordoñez, 2011).

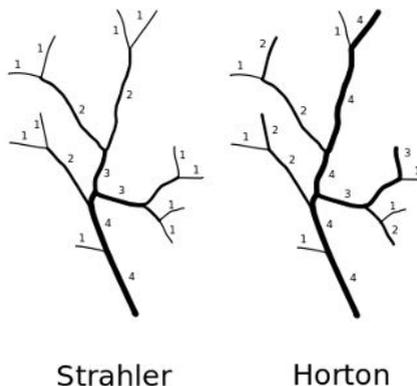


Figura 2.4: Criterios de Strahler y Horton para el orden de la cuenca  
Fuente: (Jasiewicz, 2018)

### 2.2.15. Relación de confluencias ( $r_c$ )

Este parámetro corresponde a la *primera ley de Horton*. En cada cuenca, los ríos de cada orden forman una serie geométrica inversa cuyo primer término es la unidad y la razón es la relación de confluencias. (Ortiz, 2008) señala que la relación de confluencias es un valor representativo de la cuenca y es un indicador de la potencialidad erosiva y de la rapidez de escurrimiento superficial. A mayor valor, mayor capacidad de erosión y de escurrimiento superficial. Cabe resaltar que el mismo autor indica que dos cuencas con la misma relación de confluencias promedio, son cuencas similares en su patrón de flujo y, por lo tanto, similares cinemáticamente.

La relación de confluencias se obtiene dividiendo el número total de ríos de un orden  $n$  por el número total de ríos de orden  $n + 1$ .

$$r_c = \frac{n_i}{n_{i+1}} \quad (2.12)$$

en la ecuación 2.12,  $n_i$  es el número total de ríos de orden “ $n$ ” y  $n_{i+1}$  el número total de ríos de orden “ $n+1$ ”. La relación de confluencias promedio es el valor representativo de la cuenca, se puede obtener de la ecuación 2.13.

$$N_r = \frac{r_c \times N - 1}{r_c - 1} \quad (2.13)$$

donde  $r_c$  es la relación de confluencias (parámetro adimensional),  $N_r$  el número total de ríos y  $N$  es el orden de la cuenca.

### 2.2.16. Relación de longitudes ( $r_L$ )

El parámetro descrito en este ítem corresponde a la *segunda ley de Horton* y se enuncia: “En una cuenca determinada, las longitudes medias de los ríos de la cuenca forman una serie geométrica directa cuyo primer término es la longitud promedio de los cursos elementales de la cuenca y la razón es la relación de longitudes”. La ecuación 2.14 indica cómo determinar la relación de longitudes.

$$r_L = \frac{L_i}{L_{i-1}} \quad (2.14)$$

El promedio de la relación de longitudes es el valor representativo de la cuenca, y está relacionado con la potencialidad erosiva y la magnitud de la escorrentía superficial, sin embargo también tiene influencia directa en la capacidad de almacenamiento momentáneo en la red de drenaje natural, capacidad que se ve reflejada en el pico del hidrograma de escorrentía directa del cauce principal. A mayor capacidad de almacenamiento en la red hidrográfica, menor caudal pico del hidrograma y viceversa (Ortiz, 2008).

### 2.2.17. Densidad de drenaje ( $D_d$ )

La *tercera ley de Horton* corresponde a este parámetro, y se enuncia como: “La red de drenaje de una cuenca puede ser caracterizada frecuentemente por una red suelta o densa, según que estén sueltos o concentrados los cursos de agua”. Con este parámetro se puede tener indicios acerca de las características sobre las cuales se ha desarrollado el sistema de cursos naturales. La densidad de drenaje y la causa de escurrimiento de la cuenca (Ortiz, 2008).

La ecuación 2.15 permite calcular dicho parámetro.

$$D_d = \frac{\sum_{i=1}^N L_i}{A} \quad (2.15)$$

donde  $L_i$  es la longitud total de los cursos de cada orden en  $km$ ,  $A$  es el área de la cuenca en  $km^2$ , y  $N$  es el orden de la cuenca. Las unidades son  $km/km^2$ .

### 2.2.18. Frecuencia de ríos ( $F_r$ )

Está dada por la ecuación 2.16, donde  $n_i$  es el número de ríos de orden  $i$  y  $A$  es el área de la cuenca en  $km^2$ .

$$F_r = \frac{n_i}{A} \quad (2.16)$$

La frecuencia de ríos representativa de la cuenca está dada por el promedio de los valores parciales obtenidos para cada orden. Cuando una red presenta una densidad de drenaje  $D_d$  superior a  $2.5km/km^2$  y una frecuencia de ríos de varias unidades de  $ríos/km^2$  podemos afirmar que es una red densa, es decir, que tiene un buen drenaje superficial. En cambio cuando la  $D_d$  es de apenas unas décimas de  $km/km^2$  y la frecuencia de ríos apenas algunas centésimas de  $ríos/km^2$ , podemos afirmar que la red tiene un mal drenaje superficial, esto ocurre generalmente en formaciones permeables y masivas (Ortiz, 2008).

### 2.2.19. Tiempo de concentración ( $T_c$ )

También se denomina tiempo de equilibrio o tiempo de viaje. Se define como el tiempo que toma la partícula hidráulicamente más lejana en viajar hasta el punto emisor de la cuenca. Suele asumirse que ocurre una lluvia uniforme sobre toda la cuenca durante un tiempo de por lo menos igual el tiempo de concentración (Ortiz, 2008). Existen muchas fórmulas empíricas las cuales se basan en otros parámetros, sin embargo por la naturaleza del tiempo de concentración es obvio que se verá influenciado por la cobertura vegetal e incluso por la intensidad de lluvia, razón por la cual algunos autores no lo consideran un parámetro sino una variable (Moreno Grande et al., 2015).

A continuación, se describen algunas fórmulas empíricas para su cálculo, extraídas de (Upegui, 2011) y adaptadas a las unidades que se utilizan en HidroCalc.

#### ■ Kirpich (1942)

$$T_c = 0.066 \left( \frac{L_f}{\sqrt{S_0}} \right)^{0.77} \quad (2.17)$$

donde  $T_c$  es el tiempo de concentración en *horas*,  $L_f$  es la longitud del cauce principal hasta la divisoria de aguas en *km* y  $S_0$  es la pendiente promedio del cauce principal en *m/m*.

- **Témez (1978)**

$$T_c = 0.3 \left( \frac{L_c}{S_0^{0.25}} \right)^{0.76} \quad (2.18)$$

donde  $T_c$  es el tiempo de concentración en *horas*,  $L_c$  es la longitud del cauce principal en *km* y  $S_0$  es la pendiente promedio del cauce principal en *m/m*.

- **Giandotti**

$$T_c = \frac{4\sqrt{A} + 1.5L_c}{25.3\sqrt{L_c S_0}} \quad (2.19)$$

donde  $T_c$  es el tiempo de concentración en *horas*,  $L_c$  es la longitud del cauce principal en *km*,  $A$  es el área de la cuenca en *km<sup>2</sup>* y  $S_0$  es la pendiente promedio del cauce principal en *m/m*.

- **SCS**

$$T_c = \frac{0.0136L^{0.8} \left( \frac{1000}{CN} - 9 \right)^{0.7}}{S^{0.5}} \quad (2.20)$$

donde  $T_c$  es el tiempo de concentración en *horas*,  $L_c$  es la longitud del cauce principal hasta la divisoria de aguas en *m*,  $CN$  es el número de curva SCS y  $S$  es la pendiente promedio de la cuenca en *m/m*.

## 2.2.20. Curva hipsométrica

Es una gráfica que representa algo similar a un perfil longitudinal promedio de la cuenca. Es el resultado de plotear el área acumulada que queda por encima de cada curva de nivel en el eje de las abscisas versus las alturas correspondientes en el eje de las ordenadas (Ortiz, 2008).

## 2.2.21. Frecuencia de altitudes

Es un gráfico que representa el grado de incidencia de las áreas comprendidas entre curvas de nivel con respecto al total del área de la cuenca (Ordoñez, 2011).

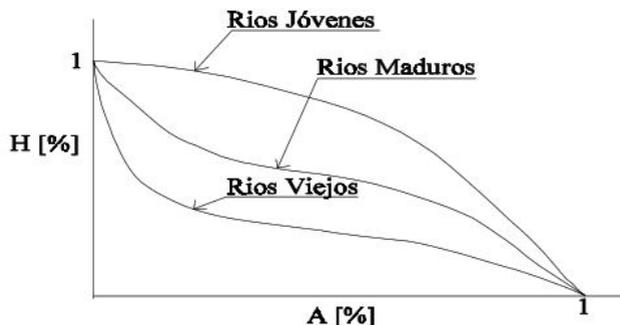


Figura 2.5: Cambio de forma de la curva hipsométrica con la edad del río  
Fuente: (Ibáñez et al., 2011)

### 2.2.22. Orientación de la cuenca

La orientación de la cuenca se determina de la orientación del cauce principal en relación al Norte. Es un parámetro poco estudiado y sin embargo es de gran importancia pues da indicios de la cantidad de radiación solar que recibe la cuenca, así como las corrientes de aire, factores de gran influencia en el aprovechamiento de energía solar y eólica, además que influyen considerablemente en la evapotranspiración (Ortiz, 2008; Moreno Grande et al., 2015).

### 2.2.23. Centro de gravedad ( $C.G$ )

Está dado por las coordenadas  $x$  e  $y$  de la cuenca, basándonos en la forma de su perímetro. Si la posición del centro de gravedad es equidistante a la línea del perímetro, este señala cierta simetría y por lo tanto un grado de redondez.

Como comentario adicional, basados en (Jardí Porqueras, 1985) podemos clasificar algunos de los parámetros descritos anteriormente de la siguiente manera:

- **Parámetros que definen el tamaño de la cuenca.** Área  $A$  y Perímetro  $P$ .
- **Parámetros que definen la simetría de la cuenca.** Centro de gravedad  $C.G$ .
- **Parámetros que definen la elongación de la cuenca.** Longitud del cauce principal  $L$ , Factor de forma  $R_f$  y el Ancho promedio  $B$ .

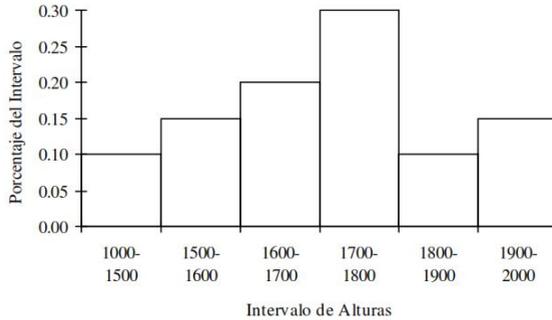


Figura 2.6: Representación de un Histograma Frecuencias de altitudes

*Fuente: (Ibáñez et al., 2011)*

- **Parámetros que definen la forma del contorno o perímetro de la cuenca.** Coeficiente o índice de compacidad  $K_c$ .

Finalmente, los parámetros Relación de confluencias  $r_c$ , Relación de longitudes  $r_L$  y Densidad de drenaje  $D_d$  corresponden a las *Leyes de Horton*.

## Capítulo 3

# Hidrología estadística

En este capítulo, se señalan los conceptos de hidrología estadística utilizados en la programación de HidroCalc 2. Se abarcan temas como el análisis de frecuencias para su aplicación en eventos máximos, análisis de consistencia y tendencias para series de tiempo, y regresión para la completación de datos faltantes o regionalización de variables hidrológicas.

HidroCalc 2 posee herramientas de estadística paramétrica y no paramétrica, asimismo, los algoritmos de resolución matricial permiten el manejo de grandes cantidades de datos.

Los parámetros de las distribuciones teóricas de probabilidad se estiman mediante los métodos de momentos ordinarios y momentos lineales. Se han considerado las distribuciones teóricas más utilizadas en hidrología para el análisis de eventos.

### 3.1. Conceptos básicos de estadística

El objetivo de la estadística en la hidrología, es el análisis de la información hidrológica en forma de muestras, a fin de inferir las características con que se debe esperar en el futuro el fenómeno que se estudia. Los estudios hidrológicos requieren del análisis de cuantiosa información hidrometeorológica, que puede consistir en datos de precipitación, caudales, temperatura y evaporación, entre otros. Los datos recopilados solamente representan una información en bruto, pero si se organizan y analizan en forma adecuada proporcionan al hidrólogo una herramienta de gran utilidad que le permite tomar decisiones en el diseño de estructuras hidráulicas.

cas. Para el análisis de la información, la hidrología utiliza los conceptos de probabilidades y estadística, siendo este campo una de las primeras áreas de la ciencia e ingeniería en usar los conceptos estadísticos en un esfuerzo para analizar los fenómenos naturales. (Adaptado de (Cherque Morán, 1989))

### 3.1.1. Términos básicos

- **Experimento aleatorio:** es todo aquel experimento que proporciona diferentes resultados aun cuando se repite de la misma manera.
- **Espacio muestral:** es el conjunto de los posibles resultados que se pueden obtener al realizar un experimento aleatorio.
- **Eventos:** son los resultados posibles que se pueden presentar de una experimento aleatorio, es un subconjunto del espacio muestral.

### 3.1.2. Variable aleatoria

Función  $X$ , definida sobre el espacio muestral. La función asigna un valor a cada punto o resultado del espacio muestral y es definida sobre un espacio de probabilidad. Para trabajar de manera sólida con variables aleatorias en general es necesario considerar un gran número de experimentos aleatorios, para su tratamiento estadístico. (Adaptado de (Villón Béjar, 2002))

### 3.1.3. Probabilidad

Es una medida del grado de incertidumbre de que ocurra un evento. Su definición clásica se muestra en la ecuación 3.1.

$$P(A) = N_a/N_s \quad (3.1)$$

Donde  $N_a$  son los resultados favorables o en los que ocurre el evento y  $N_s$  son los resultados igualmente posibles o el espacio muestral.

### 3.1.4. Periodo de retorno

El periodo de retorno se utiliza en muchas disciplinas para describir y cuantificar riesgos. En procesos estacionarios, el periodo de retorno de una evento particular, puede definirse como el inverso de la probabilidad

de que dicho evento puede ser excedido en cualquier año dado ((Cooley, 2013)).

$$Tr = \frac{1}{P(X \geq x_T)} \quad (3.2)$$

Donde  $Tr$  es el periodo de retorno en años y  $P(X \geq x_T)$  la probabilidad de excedencia del evento.

### 3.1.5. Probabilidad empírica

Teniendo un conjunto de datos ordenados de mayor a menor, la probabilidad de ocurrencia de cada uno se puede determinar mediante las ecuaciones del cuadro 3.1.

Cuadro 3.1: Ecuaciones de probabilidad empírica

Fórmula empírica	Probabilidad experimental acumulada
California	$\frac{m}{n}$
Hazen	$\frac{m - \frac{1}{2}}{n}$
Weibull	$\frac{m}{n + 1}$
Tukey	$\frac{3m - 1}{3n + 1}$

*Fuente: (Villón Béjar, 2002)*

Donde  $P$  es la probabilidad empírica acumulada,  $m$  es el número de orden y  $n$  es el número de datos.

La probabilidad de excedencia  $P(X \geq x)$  se calcula con los datos ordenados de mayor a menor, la probabilidad de no-excedencia  $P(X < x)$  se calcula con los datos ordenados de menor a mayor.

## 3.2. Análisis de frecuencias

En el diseño hidrológico, es necesario el conocimiento de la ley de probabilidades de la variable estudiada, con la finalidad de minimizar los riesgos o alcanzar niveles aceptables de seguridad en las estructuras hidráulicas. Para este fin, el hidrólogo estima estas probabilidades mediante el análisis de frecuencias. La esencia de este análisis es la búsqueda de la distribución teórica que se adapte o ajuste mejor a los datos con la finalidad de estimar valores específicos de la variable para probabilidades o periodos de retorno determinados.

### 3.2.1. Distribuciones teóricas de probabilidad

*“Una distribución de probabilidad es un reparto teórico de la población y por tanto es una función matemática, los valores que son más abundantes tendrán mayor probabilidad de aparecer al realizar la experiencia aleatoria que los valores más escasos” - (Campos Bonilla, 2017).*

La distribución de probabilidad es una función que asigna a cada suceso o evento de la variable la probabilidad de que dicho suceso o evento ocurra. Pueden representarse a través de sus formas acumulada y de densidad.

Por sus siglas en inglés, se abrevia las acumuladas como CDF y las de densidad como PDF. En forma práctica, se utilizan más las acumuladas dado que la probabilidad empírica también se determina de manera acumulativa.

- **Función de densidad - PDF:** Describe la probabilidad de que una variable aleatoria  $X$ , tome cualquier valor menor o igual a  $x$ .
- **Función acumulada - CDF:** Describe la probabilidad relativa, según la cual una variable aleatoria  $X$  tomará un valor cercano a  $x$ .

### Distribución Normal

Es considerada la distribución estadística más importante por la cantidad de fenómenos naturales, sociales y psicológicos que explica. Se le conoce también como el nombre de distribución de Gauss, Gaussiana, o de Laplace-Gauss (Juan et al., 2006).

Su PDF tiene forma acampanada y es simétrica respecto de un determinado parámetro estadístico. Otras de sus principales características es que desciende suavemente en ambas direcciones a partir del calor cen-

tral (único pico) y es asintótica, es decir, se acerca cada vez más al eje X pero sin llegar a tocarlo.

- Función de densidad

$$f(x) = \frac{1}{\sqrt{2\pi}S} e^{-\frac{1}{2}\left(\frac{x - \bar{X}}{S}\right)^2} \quad (3.3)$$

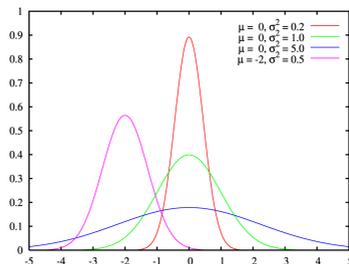


Figura 3.1: PDF - Normal 2P

Fuente: Wikipedia

- Función acumulada

$$F(x) = \frac{1}{\sqrt{2\pi}S} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{x - \bar{X}}{S}\right)^2} dx \quad (3.4)$$

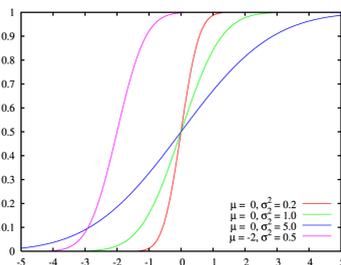


Figura 3.2: CDF - Normal 2P

Fuente: Wikipedia

Donde  $x$  es la variable independiente,  $\bar{X}$  es el parámetro de localización (media aritmética en este caso) y  $S$  el parámetro de escala (desviación

estándar).

En el caso de la distribución Normal, también se puede expresar la función acumulada en base a la variable estandarizada de los datos, tal como lo muestra la ecuación 3.5.

$$F(Z) = \frac{1}{\sqrt{2\Pi}} \int_{-\infty}^Z e^{-\frac{Z^2}{2}} dZ \quad (3.5)$$

### Distribución Gamma 2P

La función Gamma es de gran aplicación en meteorología, principalmente en análisis de precipitaciones y en otras variables con límites físicos bajos, ceros o incluso inferiores, y que sin embargo, no poseen límites superiores (Thom, 1958). Es un caso especial de la función Gamma-3P donde el parámetro de localización tiene valor 0.

#### ■ Función de densidad

$$f(x) = \frac{x^{\gamma-1} e^{-\frac{x}{\beta}}}{\beta^{\gamma} \Gamma(\gamma)} \quad (3.6)$$

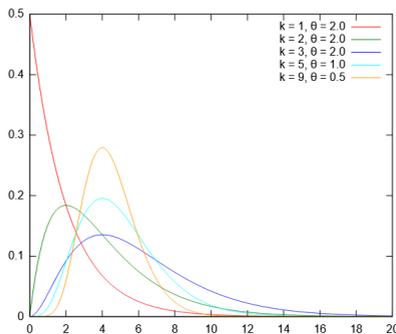


Figura 3.3: PDF - Gamma 2P

Fuente: Wikipedia

■ Función acumulada

$$F(x) = \int_0^x \frac{x^{\gamma-1} e^{-\frac{x}{\beta}}}{\beta \gamma \Gamma(\gamma)} dx \quad (3.7)$$

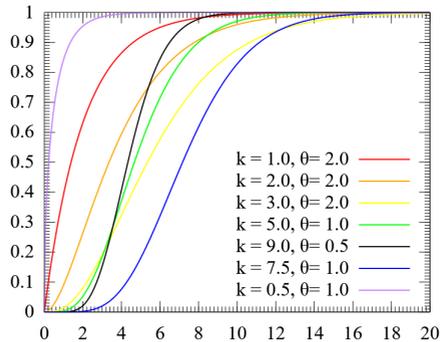


Figura 3.4: CDF - Gamma 2P

Fuente: Wikipedia

Donde  $\gamma$  es el parámetro de forma,  $\beta$  es el parámetro de escala y  $\Gamma(\gamma)$  es la función gamma completa, definida mediante la ecuación 3.8:

$$\Gamma(\gamma) = \int_0^{\infty} x^{\gamma-1} e^{-x} dx \quad (3.8)$$

Es importante mencionar que tanto la distribución Gamma de 2 parámetros y la de 3 parámetros pueden presentar dos tipos de parametrización, la primera es la ya descrita (forma y escala) y la segunda consta de parámetros de forma y tasa, inverso al de escala.

### Distribución Gamma 3P/Pearson III

La distribución Gamma 3P o Pearson tipo-III, es uno de los modelos más básicos en hidrología estadística a pesar de que su uso para modelar datos no transformados se ha restringido debido a las dificultades encontradas para ajustar esta distribución por métodos tradicionales, (Koutrouvelis and Canavos, 1999).

■ Función de densidad

$$f(x) = \frac{(x - x_0)^{\gamma-1} e^{-\frac{x - x_0}{\beta}}}{\beta^{\gamma} \Gamma(\gamma)} \tag{3.9}$$

■ Función acumulada

$$F(x) = \int_{x_0}^x \frac{(x - x_0)^{\gamma-1} e^{-\frac{x - x_0}{\beta}}}{\beta^{\gamma} \Gamma(\gamma)} dx \tag{3.10}$$

Donde  $x$  es la variable aleatoria gamma de 3 parámetros o Pearson tipo III,  $x_0$  es el parámetro de posición u origen,  $\beta$  es el parámetro de escala,  $\gamma$  es el parámetro de forma y  $\Gamma(\gamma)$  es la función gamma completa.

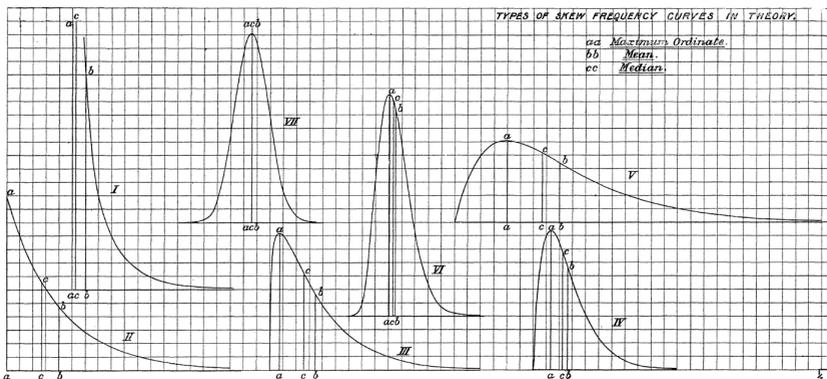


Figura 3.5: PDF - Gamma 3P

Fuente: (Pearson, 1895)

Como puede verse en la figura 3.5, las distribuciones gamma 2P y 3P, corresponden a distribuciones sesgadas, es decir, con mayores probabilidad de ocurrencia en valores extremos.

**Distribución Log-Normal 2P**

La distribución Log-normal 2P, es una de las distribuciones logarítmicas más utilizadas en hidrología. Se dice que una variable aleatoria  $X$  posee una distribución Log-normal 2P, si los valores de sus logaritmos poseen una distribución Normal ( $y = \ln x$ ).

- Función de densidad

$$f(x) = \frac{1}{x\sqrt{2\Pi}\sigma_y} e^{-\frac{1}{2}\left[\frac{\ln x - \mu_y}{\sigma_y}\right]^2} \quad (3.11)$$

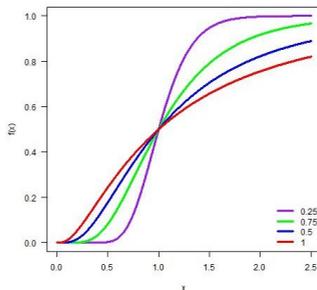


Figura 3.6: PDF - Log-normal 2P

Fuente: Wikipedia

- Función acumulada

$$F(x) = \frac{1}{x\sqrt{2\Pi}\sigma_y} \int_0^x e^{-\frac{1}{2}\left[\frac{\ln x - \mu_y}{\sigma_y}\right]^2} dx \quad (3.12)$$

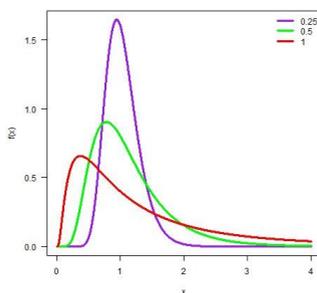


Figura 3.7: CDF - Log-normal 2P

Fuente: Wikipedia

Donde  $\mu_y$  es la media (parámetro de escala) y  $\sigma_y$  es la desviación estándar (parámetro de forma) de los logaritmos naturales de  $x$ .

### Distribución Log-Pearson tipo III

La distribución Log-Pearson tipo III, fue recomendada por el Water Resources Council (1967) - U.S.A, como un método base para el análisis de frecuencia de inundaciones, (Phine and Hira, 1983). Se dice que una variable aleatoria  $X$  posee una distribución Log-Pearson tipo III, si los valores de sus logaritmos poseen una distribución Pearson tipo III ( $y = \ln x$ ).

- Función de densidad

$$f(x) = \frac{(\ln x - x_0)^{\gamma-1} e^{-\frac{\ln x - x_0}{\beta}}}{x\beta\gamma\Gamma(\gamma)} \quad (3.13)$$

- Función acumulada

$$F(x) = \int_{x_0}^x \frac{(\ln x - x_0)^{\gamma-1} e^{-\frac{\ln x - x_0}{\beta}}}{x\beta\gamma\Gamma(\gamma)} dx \quad (3.14)$$

Donde  $x_0$  es el parámetro de posición (entre  $-\infty$  e  $\infty$ ),  $\beta$  es el parámetro de escala (entre 0 e  $\infty$ ) y  $\gamma$  es el parámetro de forma (entre 0 e  $\infty$ ).

### Distribución Gumbel

La distribución Gumbel se utiliza en el análisis de frecuencia de valores máximos y mínimos, por ello, es muy útil para modelar desastres naturales.

- Función de densidad

$$f(x) = \frac{1}{\alpha} e^{-\frac{x-\mu}{\alpha}} e^{-e^{-\frac{x-\mu}{\alpha}}} \quad (3.15)$$

- Función acumulada

$$F(x) = e^{-e^{-\frac{x-\mu}{\alpha}}} \quad (3.16)$$

Donde  $\alpha$  es el parámetro de escala (entre 0 e  $\infty$ ) y  $\mu$  es el parámetro de posición (entre  $-\infty$  e  $\infty$ ), (Villón Béjar, 2002).

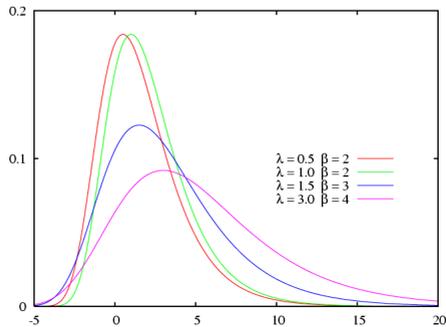


Figura 3.8: PDF - Gumbel

Fuente: Wikipedia

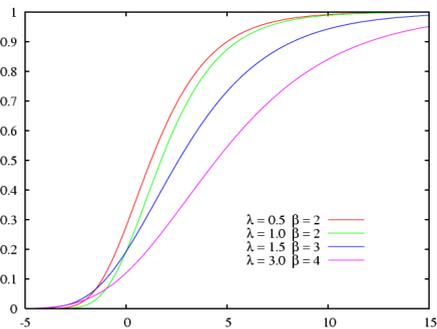


Figura 3.9: CDF - Gumbel

Fuente: Wikipedia

### 3.2.2. Estimación de parámetros de distribuciones

Los parámetros son **medidas descriptivas** de una población completa, que se pueden utilizar como las entradas para que una función de distribución de probabilidad (PDF/CDF, por sus siglas en inglés) genere curvas de distribución.

En ocasiones, el término “parámetros” se utiliza para referirse a los más comúnmente llamados estadísticos de una muestra, y aunque es una terminología más adecuada, puede resultar algo confusa ya que también debe diferenciarse entre muestrales y poblacionales. En este documento, utilizaremos el término “estadísticos” dado que son los que utilizaremos para encontrar lo que llamaremos *parámetros de las distribuciones*.

Los parámetros son los que le darán forma a la curva de densidad o acumulada de la distribución de probabilidad teórica que hayamos seleccionado para modelar nuestra variable aleatoria y que están sujetos a incertidumbre o errores de ajuste. En cuanto menos incertidumbre tengan estos parámetros, más seguros vamos a estar de que nuestra variable aleatoria tiene el comportamiento que indica la distribución que hemos elegido. Aquí es donde radica la importancia de su estimación.

En esencia, utilizaremos estos parámetros para **adaptar** la distribución de probabilidad teórica al **comportamiento real** de nuestra variable aleatoria. Para la estimación de los parámetros, existen muchos métodos, los más comunes son:

- Método gráfico
- Momentos ordinarios
- Momentos lineales
- Mínimos cuadrados
- Máxima verosimilitud
- Máxima entropía

HidroCalc 2, posee los métodos que se detallan a continuación.

### Método de momentos ordinarios

Este método de estimación de parámetros fue desarrollado por Karl Pearson en 1902. El cálculo de los momentos muestrales, se formuló a través de una analogía física, en la cual se le asigna una masa hipotética a cada uno de los valores de los datos o muestras obtenidas de una población (variable aleatoria). Esta masa es equivalente a la probabilidad de ocurrencia ( $1/n$ ) y que este sistema de masas hipotéticas rota alrededor del origen ( $x = 0$ ). De esta manera, el primer momento de cada observación  $x_i$  alrededor del origen  $x = 0$ , es el producto de su brazo de momento  $x_i$  y de su masa  $1/n$  y la suma de estos momentos para toda la información es la media de la muestra, lo cual es equivalente al centroide del cuerpo, (Chow et al., 1994), tal como lo muestra la ecuación 3.17.

$$\mu = \int_{-\infty}^{\infty} x f(x) dx \quad (3.17)$$

Luego es necesario expresar los **momentos poblacionales** (es decir, los

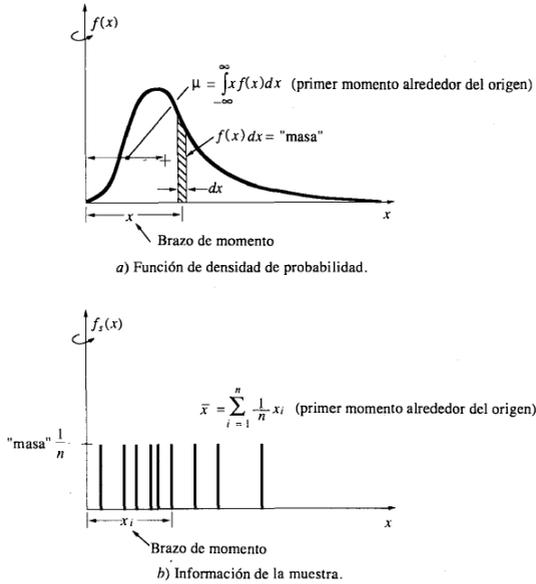


Figura 3.10: Momento de una distribución  
Fuente: (Chow et al., 1994)

valores esperados de las potencias de la variable aleatoria) como funciones de los parámetros de interés. Estas últimas expresiones se igualan a los **momentos muestrales** y se obtienen sistemas de ecuaciones cuyo número de ecuaciones es el mismo que el número de parámetros a estimar y cuya solución devuelve el valor de los parámetros.

De esta manera, siendo el problema la estimación de los  $k$  parámetros  $\theta_1, \theta_2, \dots, \theta_k$  que dan forma a la distribución  $f_W(w; \theta)$  de la variable aleatoria  $W$ , supongamos que los  $k$  momentos poblacionales pueden ser expresados como funciones de los  $\theta$ 's:

$$\begin{aligned}
 \mu_1 &\equiv E[W] = g_1(\theta_1, \theta_2, \dots, \theta_k), \\
 \mu_2 &\equiv E[W^2] = g_2(\theta_1, \theta_2, \dots, \theta_k), \\
 &\vdots \\
 \mu_k &\equiv E[W^k] = g_k(\theta_1, \theta_2, \dots, \theta_k).
 \end{aligned}
 \tag{3.18}$$

Asumiendo que extraemos una muestra de tamaño  $n$ , cuyos valores son  $w_1, \dots, w_n$  para  $j = 1, \dots, k$  los momentos muestrales se determinan mediante:

$$\hat{\mu}_j = \frac{1}{n} \sum_{i=1}^n w_i^j \quad (3.19)$$

Siendo  $\hat{\mu}_j$  una estimación de  $\mu_j$ . Finalmente, el estimador del método de los momentos para los parámetros  $\theta_1, \theta_2, \dots, \theta_k$  denotados por  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$  se define como la solución de las ecuaciones:

$$\begin{aligned} \hat{\mu}_1 &= g_1(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k), \\ \hat{\mu}_2 &= g_2(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k), \\ &\vdots \\ \hat{\mu}_k &= g_k(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k). \end{aligned} \quad (3.20)$$

### Método de momentos lineales

Los momentos lineales son una metodología alternativa para la estimación de parámetros de una población o muestra. Fue propuesta por Hosking en el año 1990. En la actualidad, también se utilizan para caracterizar variables hidrológicas como la precipitación máxima y realizar análisis regional de frecuencias. La metodología posee robustez aun en presencia de valores extremos.

Los momentos lineales o momentos-L, se derivaron de los momentos de probabilidad pesada o momentos ponderados propuestos anteriormente por Greenwood en 1979 y poseen ecuaciones que los relacionan directamente.

Para una variable aleatoria  $X$ , el  $r$ -ésimo momento lineal poblacional se define como:

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E X_{r-k} \quad (3.21)$$

Donde  $X_{k:n}$  denota el  $k$ -ésimo orden estadístico ( $k$ -ésimo valor más pequeño) de una muestra independiente de tamaño  $n$  de la distribución de  $X$  y  $E$  es el valor esperado.

Las expresiones desarrolladas para los cuatro primeros momentos lineales son:

$$\begin{aligned}
 \lambda_1 &= E[X_{1:1}] \\
 \lambda_2 &= \frac{1}{2}E[X_{2:2} - X_{1:2}] \\
 \lambda_3 &= \frac{1}{3}E[X_{3:3} - 2X_{2:3} + X_{1:3}] \\
 \lambda_4 &= \frac{1}{4}E[X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}]
 \end{aligned}
 \tag{3.22}$$

Para estimar estos momentos se puede proceder a través de los momentos pesados de probabilidad, sin embargo, también se pueden calcular a partir de las combinaciones lineales, (Hosking and Wallis, 2005).

$$\begin{aligned}
 \lambda_1 &= \binom{n}{1}^{-1} \sum_{i=1}^n X_i \\
 \lambda_2 &= \frac{1}{2} \binom{n}{2}^{-1} \sum_{i=1}^n ((\binom{i-1}{1}) - \binom{n-i}{1}) X_i \\
 \lambda_3 &= \frac{1}{3} \binom{n}{3}^{-1} \sum_{i=1}^n ((\binom{i-1}{2}) - 2\binom{i-1}{1}\binom{n-i}{1} + \binom{n-i}{2}) X_i \\
 \lambda_4 &= \frac{1}{4} \binom{n}{4}^{-1} \sum_{i=1}^n ((\binom{i-1}{3}) - 3\binom{i-1}{2}\binom{n-i}{1} + 3\binom{i-1}{1}\binom{n-i}{2} - \binom{n-i}{3}) X_i
 \end{aligned}
 \tag{3.23}$$

De manera análoga al método de momentos, el método de los momentos-L obtiene parámetros igualando los momentos-L muestrales a las cantidades correspondientes a la población o momentos poblacionales.

### 3.2.3. Bondad de ajuste: Smirnov-Kolmogorov

Las pruebas de bondad de ajuste se utilizan para evaluar las hipótesis sobre la distribución de una población, en otros términos, si la distribución empírica de probabilidades de una variable aleatoria se ajusta a una distribución teórica elegida *a priori*, (Massey Jr, 1951). Existen muchas pruebas y método para evaluar de bondad de ajuste, entre estos tenemos:

- Método gráfico
- Prueba de Smirnov-Kolmogorov
- Prueba de Chi-cuadrado
- Criterio de Cramer-Von Mises
- Prueba de Anderson-Darling

- Prueba de Shapiro-Wilk
- Criterio de información de Akaike

En análisis de regresión, algunos tópicos también se pueden relacionar a la bondad de ajuste, sin embargo, en este documento se detalla únicamente la prueba de Smirnov-Kolmogorov dado que es la prueba utilizada por HidroCalc 2.

### Smirnov-Kolmogorov

La prueba consiste en comparar la diferencia existente entre la probabilidad empírica de los datos de la muestra y la probabilidad teórica de la distribución elegida y se toma el valor absoluto máximo para compararlo con el valor crítico. Esto es:

$$\Delta = \max |F(x) - P(x)| \quad (3.24)$$

Donde  $\Delta$  es el estadístico de Smirnov-Kolmogorov,  $F(x)$  es la probabilidad de la distribución teórica y  $P(x)$  la probabilidad empírica.

Las hipótesis a contrastar son las siguientes:

- $H_0$ : Los datos analizados siguen una distribución  $F(x)$ .
- $H_1$ : Los datos analizados no siguen una distribución  $F(x)$ .

Los criterios a analizar serán:

- Si  $\Delta \leq \Delta_\alpha \rightarrow$  no rechazar  $H_0$ .
- Si  $\Delta > \Delta_\alpha \rightarrow$  rechazar  $H_0$ .

El valor crítico  $\Delta_\alpha$  se determina para el nivel de significancia  $\alpha$ , de tal manera que  $P(\Delta > \Delta_\alpha) = \alpha$ .

El valor  $\Delta_\alpha$  puede determinarse a través de tablas elaboradas para diferentes niveles de significancia ( $\alpha$ ), sin embargo, HidroCalc 2 resuelve el valor de  $x$  en la siguiente ecuación para determinar el valor crítico.

$$P(K \leq x) = \frac{\sqrt{2\pi}}{x} \sum_{k=1}^{\infty} e^{-(2k-1)^2 \pi^2 / (8x^2)} \quad (3.25)$$

Para mejorar la estimación de  $x$  en muestras pequeñas, se puede reemplazar  $x$  por la ecuación 3.26.

$$x + \frac{1}{6\sqrt{n}} + \frac{x-1}{4n} \quad (3.26)$$

La hipótesis nula se rechaza cuando:

$$\sqrt{n}\Delta_\alpha > K_\alpha \quad (3.27)$$

Donde  $K_\alpha$  se encuentra de:

$$Pr(K \leq K_\alpha) = 1 - \alpha \quad (3.28)$$

### 3.3. Análisis de consistencia y tendencia

La información hidrometeorológica debe pasar por el análisis de confiabilidad. (Ortiz, 2008) recomienda que el análisis de información debe estar relacionado con la *confiabilidad* (calidad), *continuidad* (datos faltantes) y la *longitud* (años de registro). Por lo tanto, el proceso de análisis debe cumplir con analizar la consistencia y homogeneidad (que incluye la detección y corrección de posibles errores sistemáticos), tendencia y completación de datos faltantes y extensión de series cortas a periodos largos.

Se tomará como base las definiciones del libro “*Hydrologic Time Series Analysis*”, de (Machiwal and Jha, 2012). Estos autores indican que tanto la *homogeneidad* o la *consistencia* implican que todo el registro hidrológico recolectado debe pertenecer a la misma población estadística, con estadísticas invariantes en el tiempo.

#### 3.3.1. Valores atípicos

Un valor atípico es un valor distante del resto de observaciones, pudiendo provenir de extremos o de errores de medición. Puede ocasionar distorsión en las medidas estadísticas, por esta razón es recomendable identificarlos y en ocasiones corregirlos o eliminarlos cuidadosamente.

El método (Tukey, 1977) se utiliza par identificar dichos valores y consiste en elaborar un gráfico estadístico conocido como gráfico de caja o *boxplot*. En este gráfico se representan cinco medidas estadísticas: el valor mínimo, el primer cuartil, la mediana, el tercer cuartil y el valor máximo (Minnaard et al., 2005).

El cuerpo de la caja queda representado por el rango intercuartílico (*RIC*), es decir, entre el primer y tercer cuartil, con la mediana representada por una línea dentro de la misma. Los valores máximo y mínimo corresponden respectivamente a los estimados de las ecuaciones 3.29 y 3.30.

$$Val_{max} = Q_3 + 1.5RIC \quad (3.29)$$

$$Val_{min} = Q_1 - 1.5RIC \quad (3.30)$$

Donde  $Val_{max}$  y  $Val_{min}$  son los valores máximo y mínimo,  $Q_3$  y  $Q_1$  son los valores correspondientes al tercer y primer cuartil respectivamente, y *RIC* es el rango intercuartílico, que se calcula de la ecuación 3.31.

$$RIC = Q_3 - Q_1 \quad (3.31)$$

Todo valor por sobre o bajo el valor máximo y mínimo se considera un valor *atípico*. Además se consideran valores *extremadamente atípicos* aquellos que sobrepasen los valores de calculados de las ecuaciones 3.32 y 3.33.

$$Ext_{max} = Q_3 + 3RIC \quad (3.32)$$

$$Ext_{min} = Q_1 - 3RIC \quad (3.33)$$

Donde  $Ext_{max}$  y  $Ext_{min}$  son los valores extremos máximos y mínimos.

El método de Tukey o gráfico de cajas, se limita a valores con distribución normal, por ello, HidroCalc 2 permite normalizar los valores, haciendo uso de la ecuación 3.34, (Dodge and Commenges, 2006).

$$X_Z = \frac{X - \mu}{\sigma} \quad (3.34)$$

Donde  $X_Z$  es la variable normalizada,  $\mu$  es la media de la muestra y  $\sigma$  la desviación estándar.

### 3.3.2. Estadística paramétrica y no paramétrica

Jacob Wolfowitz en 1942 acuñó el término de estadística “no-paramétrica” diciendo: “Podemos referirnos a la situación donde una distribución es determinada a través del pleno conocimiento de su conjunto de parámetros finitos, como el caso paramétrico, y denotar el caso opuesto, donde las formas funcionales de las distribuciones son desconocidas, nombrándolo caso no-paramétrico”. Partiendo de este punto, la estadística no-paramétrica puede ser definida por lo que no es: estadística tradicional basada en el conocimiento de las distribuciones y sus parámetros estadísticos (Kvam and Vidakovic, 2007).

La ventaja de usar pruebas no-paramétricas es que son más robustas que las paramétricas, ya que exigen menos condiciones de validez o supuestos, sin embargo, la ventaja de usar pruebas paramétricas es que son más potentes que las no-paramétricas, es decir que tienen una mayor capacidad para conducir a un rechazo de hipótesis nulas ( $H_0$ ).

Lo que interesa comprender bajo las definiciones dadas, es que pruebas como *T-Student* o *F-Fisher* corresponden a *pruebas paramétricas*, cuya asunción principal es la *normalidad* de las muestras y aun más, en el caso de *T-Student* se requiere la condición de homogeneidad de varianzas (*homocedasticidad*) (Rojas, 2003) y la ausencia de valores atípicos.

### 3.3.3. Homogeneidad en la media

La homogeneidad en la media puede ser evaluada mediante muchos métodos de estadística paramétrica y no paramétrica. En HidroCalc 2 puede hacerse a través de la prueba de Buishand (Test de desviaciones acumuladas) o la prueba de T-Student.

#### Test de desviaciones acumuladas - Buishand

Llamado también *Test de Buishand* o *Test de rangos de Buishand*, es una prueba no-paramétrica que por su origen bayesiano puede emplearse en variables con cualquier tipo de distribución. Sirve para identificar cambios en la media de la serie de datos. Fue propuesto por T.A. Buishand en 1982. Su autor menciona que esta prueba se basa en las desviaciones acumuladas de la media (ecuación 3.35).

$$S_0^* = 0; \quad S_k^* = \sum_{i=1}^k (Y_i - \bar{Y}), \quad k = 1, \dots, n \quad (3.35)$$

En caso de homogeneidad, los valores de  $S_k$  se encuentran alrededor de 0. Al existir una discontinuidad en la posición  $k$ , entonces  $S_k$  tendrá un valor máximo o un mínimo alrededor de esa posición. Se pueden escalar las sumas parciales mediante la ecuación 3.36.

$$S_k^{**} = S_k^* / D_Y \quad k = 0, \dots, n \quad (3.36)$$

Donde  $D_Y$  es la desviación estándar de la muestra  $(Y_1, \dots, Y_n)$ .

Los valores de  $S_k^{**}$  no están influenciados por la transformación lineal de los datos, es decir, no dependen de la unidad de la variable hidrológica. La sensibilidad a las desviaciones de la homogeneidad se define mediante la siguiente estadística:

$$Q = \max_{0 \leq k \leq n} |S_k^{**}| \quad (3.37)$$

Los valores altos de  $Q$  son un indicio de cambios en la media. Los valores críticos de  $Q$  se dan en el cuadro 3.2. También se puede usar el rango  $R$  como otro estadístico de prueba (Buishand, 1982).

$$R = \max_{0 \leq k \leq n} S_k^{**} - \min_{0 \leq k \leq n} S_k^{**} \quad (3.38)$$

Cuadro 3.2: Valores críticos del test de Desviaciones acumuladas

Tamaño muestra n	$Q/\sqrt{n}$			$R/\sqrt{n}$		
	90 %	95 %	99 %	90 %	95 %	99 %
10	1.05	1.14	1.29	1.21	1.28	1.38
20	1.10	1.22	1.42	1.34	1.43	1.60
30	1.12	1.24	1.46	1.40	1.5	1.70
40	1.13	1.26	1.5	1.42	1.53	1.74
50	1.14	1.27	1.52	1.44	1.55	1.78
100	1.17	1.29	1.55	1.50	1.62	1.86
$\infty$	1.22	1.36	1.63	1.62	1.75	2.0

Fuente: (Buishand, 1982)

### Prueba de T-Student

Es una prueba paramétrica cuyas premisas son: normalidad en los datos y muestras independientes. La prueba fue propuesta por William Sealy Gosset bajo el seudónimo de “estudiante” (Student), (Sánchez Turcios, 2015). Se puede ejecutar asumiendo, o no, la igualdad de varianzas (homocedasticidad), (Welch, 1938).

Es necesario identificar los periodos de análisis y luego determinar la media y varianza de cada uno. El estadístico  $t$  se determina para los casos mostrados a continuación, asimismo, el valor de los grados de libertad se determina diferente en cada caso.

- Para varianzas iguales, diferentes tamaños de muestra.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{X_1 X_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (3.39)$$

Donde  $\bar{X}_n$  es la media y  $S_{X_1 X_2}$  se determina de:

$$S_{X_1 X_2} = \sqrt{\frac{(n_1 - 1)S_{X_1}^2 + (n_2 - 1)S_{X_2}^2}{n_1 + n_2 - 2}} \quad (3.40)$$

Los grados de libertad se determinan a partir de:

$$g.l. = n_1 + n_2 - 2 \quad (3.41)$$

- Para varianzas desiguales, diferentes tamaños de muestra.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} \quad (3.42)$$

Donde  $\bar{X}_n$  es la media y  $S_{\bar{X}_1 - \bar{X}_2}$  se determina de:

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (3.43)$$

Los grados de libertad se determinan a partir de:

$$g.l. = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{(S_1^2/n_1)/(n_1 - 1) + (S_2^2/n_2)/(n_2 - 1)} - 2 \quad (3.44)$$

Con el nivel de confianza buscado y los grados de libertad ( $g.l.$ ) se determina el valor crítico en base a valores tabulados o a partir de la función acumulada de la distribución  $t$  - Student (ecuación 3.45), (Zogheib and Elsayehi, 2015).

$$F(t/n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \int_{-\infty}^t \left(1 + \frac{t^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} dt \quad (3.45)$$

Donde  $\Gamma$  es la función gamma y  $n$  son los grados de libertad. Se pueden emplear métodos numéricos para aproximar una solución a la ecuación 3.45, sin embargo existen métodos de aproximación, HydroCalc hace uso de la ecuación 3.46 obtenida por (Zogheib and Hlynka, 2009):

$$F(t/n) = (1 + e^{0.000345z^5 - 0.069547z^3 - 1.604326z})^{-1} \quad (3.46)$$

Donde  $z$  se obtiene de:

$$z = t \left(1 - \frac{1}{4n}\right) \left(1 + \frac{1}{2n}t^2\right)^{-\frac{1}{2}} \quad (3.47)$$

Finalmente, siendo  $t_c$  y  $t_t$  los valores calculado y tabulado (obtenido de la función de distribución acumulada), se tiene que:

- Si  $|t_c| \leq t_t(95\%)$ , significa que las medias son estadísticamente similares.
- Si  $|t_c| > t_t(95\%)$ , significa que las medias son estadísticamente diferentes.

### 3.3.4. Homogeneidad en la varianza

También es llamada “homocedasticidad”, significa que la varianza no cambie a lo largo de las observaciones. De igual manera que con la media, existen muchas pruebas para verificar la homocedasticidad, tanto de estadística paramétrica como no paramétrica. HidroCalc 2 permite evaluarlo mediante las pruebas de Fligner-Killeen y F-Fisher y Snedecor.

#### Fligner-Killeen

Es una prueba no-paramétrica que sirve para verificar la homogeneidad de un grupo de varianzas, basándonos en los rangos. Se utiliza comúnmente en información que no se ajusta a una distribución normal o en la cual existen valores atípicos (Zaiontz, 2018).

Este test compara la varianzas basándose en la mediana. El procedimiento de Fligner-Killeen (1976), modificado por Conover et al. (1981), para probar la homogeneidad de varianzas consiste en los siguientes pasos:

1° Se ordenan las variables  $|X_{i,j} - \tilde{X}_i|$  de menor a mayor, donde  $\tilde{X}_i$  es la mediana de las  $n_i$  observaciones de la población  $i$ .

2° Se define:

$$a_{N,i} = \Phi^{-1} \left( \frac{1}{2} + \frac{i}{2(N+1)} \right) \quad \text{para } i = 1, \dots, N \quad (3.48)$$

Donde  $\Phi(z)$  es la distribución acumulada  $N(0, 1)$  de  $-\infty$  a  $z$  y así  $\Phi^{-1}(p)$  es el percentil 100p de la distribución  $N(0, 1)$ .

3° Sea

$$\bar{a}_i = \sum_{j \in G_i}^N \frac{a_{N,j}}{n_i} \quad (3.49)$$

Donde  $G_i$  denota la muestra de la población  $i, i = 1, \dots, k$ . Y

$$\bar{a} = \sum_{j=1}^N \frac{a_{N,j}}{N} \quad (3.50)$$

Luego, el estadístico de prueba es:

$$x = \frac{\sum_{i=1}^k n_i (\bar{a}_i - \bar{a})^2}{\sum_{j=1}^N (a_{N,j} - \bar{a})^2 / (n - 1)} \quad (3.51)$$

Este estadístico bajo  $H_0$  se distribuye aproximadamente  $\chi_{k-1}^2$ . La prueba de Fligner es menos sensible a desviaciones del supuesto de normalidad (Correa et al., 2006).

### F-Fisher y Snedecor

El nombre de esta prueba paramétrica fue acuñado por George W. Snedecor en honor a Sir Ronald A. Fisher, quién desarrolló el estadístico como la razón de varianzas en la década de 1920.

El cálculo de este estadístico se realiza a partir de las varianzas, tal como se muestra en la ecuación 3.52.

$$F_c = \frac{S_1^2(x)}{S_2^2(x)} \quad (3.52)$$

Donde  $F_c$  es el F calculado y  $S_1^2(x)$  es la varianza mayor y  $S_2^2(x)$  la menor.

El valor crítico se estima a partir de tablas o de la función acumulada de Fisher, la cual tiene la forma de la ecuación 3.53.

$$F(x; d_1, d_2) = I \frac{d_1 x}{d_1 x + d_2} \left( \frac{d_1}{2}, \frac{d_2}{2} \right) \quad (3.53)$$

Donde  $I$  es la función beta incompleta. Para fines prácticos,  $d_1$  y  $d_2$  son los grados de libertad del numerador y el denominador, que se estiman de la siguiente manera:

$$d_1 = n_1 - 1 \quad (3.54)$$

$$d_2 = n_2 - 1 \quad (3.55)$$

Finalmente, se rechaza o no la hipótesis nula a través de los siguientes criterios, (Villón Béjar, 2004):

- Si  $F_c \leq F_t(95\%)$ , las varianzas son estadísticamente similares.
- Si  $F_c > F_t(95\%)$ , las varianzas son no estadísticamente similares.

### 3.3.5. Tendencias

Para el caso de las tendencias, HidroCalc 2 solamente permite el uso de la prueba paramétrica de Mann-Kendall, la cual se detalla a continuación.

#### Prueba de Mann-Kendall

Es una prueba no-paramétrica para explorar la tendencia de una serie temporal, sin necesidad de especificar el tipo de tendencia, es decir, lineal o no-lineal. Fue propuesta por Mann en 1945 y Kendall en 1975 obtuvo su distribución estadística. Se ha encontrado que es una herramienta excelente para la detección de tendencias.

Considerando la serie temporal  $x(t = 1, 2, \dots, n)$ , cada valor de la serie ( $x_t$ ) es comparado con los valores consecutivos ( $x_{t+1}$ ) y una nueva serie  $z_k$  es generada como sigue:

$$\begin{aligned} z_k &= 1 \quad \text{para } x_t > x_{t'} \\ z_k &= 0 \quad \text{para } x_t = x_{t'} \\ z_k &= -1 \quad \text{para } x_t < x_{t'} \end{aligned} \quad (3.56)$$

donde  $k$  es dado por:

$$k = (t' - 1)(2n - t')/2 + (t - t') \quad (3.57)$$

El estadístico Mann-Kendall ( $S$ ) es definido como sigue:

$$S = \sum_{t'=1}^{n-1} \sum_{t=t'+1}^n Z_k \quad (3.58)$$

Así, este estadístico representa el número de diferencias positivas menos el número de diferencias negativas para todas las diferencias consideradas. Además, la estadística de prueba anterior para  $n > 40$  se puede escribir como:

$$u_c = \frac{S + m}{\sqrt{V(S)}} \quad (3.59)$$

Donde:

$$V(S) = \frac{1}{18} \left( n(n-1)(2n+5) - \sum_{i=1}^g e_i(e_i-1)(2e_i+5) \right) \quad (3.60)$$

En la ecuación 3.59 y 3.60,  $m = 1$  para  $S < 0$  y  $m = -1$  para  $S > 0$ ,  $g$  es el número de grupos vinculados, y  $e_i$  es el número de datos en el  $i^{\text{ésimo}}$

grupo vinculado. El valor de la prueba estadística  $u_c$  es tomado como cero para  $S = 0$ . Ahora, si el valor absoluto calculado de  $u_c$  es mayor que el valor crítico de la distribución normal estándar, no se puede rechazar la hipótesis de una tendencia ascendente o descendente en el nivel de significancia  $\alpha$ . Cabe señalar que Kendall (1975) sugirió usar la prueba de Mann-Kendall incluso para  $n$  valores tan bajos como 10, siempre que no haya demasiados valores vinculados. Hirsch et al. (1982) informaron la aplicación de esta prueba a series temporales estacionales.

### 3.4. Análisis de regresión

El análisis de regresión se trata de encontrar una relación entre una variable dependiente y una o varias variables independientes. En términos simples, se busca explicar una variable a través de otra u otras. En el proceso, se busca la variación de la variable dependiente en función de la variación una de las independientes, manteniendo fijas a las demás.

Este análisis se emplea en muchas áreas de investigación, para el caso de HidroCalc 2, se han implementado la regresión lineal simple y múltiple lineal y exponencial, las aplicaciones recomendadas incluyen completación de series de tiempo hidrológicas y análisis regional de precipitaciones máximas.

#### 3.4.1. Regresión lineal simple

Es un caso especial de la regresión lineal, en donde sólo se maneja una variable independiente, por esta razón, solo cuenta con dos parámetros. Su ecuación general es la siguiente:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad (3.61)$$

donde  $\epsilon_i$  es el error asociado a la medición del valor  $X_i$ .

Además,

$$\hat{\beta}_1 = \frac{\Sigma x \Sigma y - n \Sigma xy}{(\Sigma x)^2 - n \Sigma x^2} \quad (3.62)$$

$$\hat{\beta}_0 = \frac{\Sigma y - \hat{\beta}_1 \Sigma x}{n} \quad (3.63)$$

Extraído de: (Villón Béjar, 2002).

### 3.4.2. Regresión múltiple

#### Lineal

En el modelo de regresión lineal múltiple, se supone que la función de regresión que relaciona la variable dependiente con las variables independientes es lineal, es decir:

$$Y_t = \beta_0 + \sum \beta_i X_i + \varepsilon \quad (3.64)$$

donde:

- $\beta_0$  es el término independiente. Es el valor esperado de Y cuando  $X_1, \dots, X_p$  son cero.
- $\beta_i$  representa los coeficientes parciales de la regresión, los cuales miden el cambio en  $Y_t$  por cada cambio unitario en  $X_i$ , manteniendo a los demás términos, constantes.
- $\varepsilon$  es el error de observación debido a variables no controladas

#### Exponencial

El modelo de regresión exponencial múltiple consiste en encontrar una función exponencial de varias variables independientes que se ajusten mejor a una variable dependiente, tiene la forma de la ecuación 3.65.

$$y = a_0 x_1^{a_1} x_2^{a_2} x_3^{a_3} \dots \quad (3.65)$$

En HidroCalc 2, el análisis de regresión es validado mediante el coeficiente de correlación de Pearson, cuya fórmula es la siguiente:

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (3.66)$$

Donde  $x$  e  $y$  son las variables analizadas. La interpretación de  $r_{xy}$  implica que valores cercanos a 1 tienen una buena correlación positiva, siendo lo opuesto con valores cercanos a 0. Valores cercanos a -1 indican una correlación negativa. Asimismo, se puede hallar el coeficiente de determinación elevando  $r_{xy}^2$ .

## Capítulo 4

# Precipitación

La precipitación es la fuente primaria de agua en la superficie terrestre, su estudio es el punto de partida desde el punto de vista hidrológico. El proceso de precipitación incluye a toda forma de humedad como granizo, nieve y lluvia.

La precipitación posee variación espacial y temporal, acorde con los patrones generales de circulación atmosférica y con factores locales. El promedio de varios años de observación se conoce como su valor *normal*.

### Precipitación diaria

Es el resultado de la medición de un día pluviométrico, entendiendo a tal como la suma de las medidas de los pluviómetros a las 19:00 y a las 07:00 del día siguiente, (Basualdo, 2015).

### Corrección de medición por factor

Quando la precipitación diaria se usa en el análisis de eventos máximos (precipitación máxima en 24 horas), es necesario aplicar un factor de corrección recomendado por la Organización Meteorológica Mundial (WMO, por sus siglas en inglés) de 1.13. "Los estudios de miles de estaciones-años de datos de lluvia indican que multiplicar los resultados de un análisis de frecuencia de las cantidades máximas de lluvia anual para un único intervalo de tiempo fijo de cualquier duración de 1 a 24 horas por 1,13 producirá valores que se aproximan mucho a los que se obtendrá a partir de un análisis basado en verdaderos máximos", (Organization, 1986).

## Precipitación mensual

Es el resultado de la suma de las mediciones diarias de precipitación en un mes determinado.

## Precipitación anual

Es el resultado de la suma de las mediciones diarias de precipitación en un año determinado.

### 4.1. Curvas IDF

Las curvas IDF representan una relación matemática entre la Intensidad de una precipitación, su Duración y la Frecuencia con la que se observa.

- **Intensidad** es la precipitación por unidad de tiempo, HydroCalc utiliza mm/hora.
- **Duración** es el tiempo que transcurre entre el inicio y el fin de la tormenta.
- **Frecuencia** es el inverso del periodo de retorno.

HydroCalc 2 no presenta un módulo exclusivamente para el estudio de estas curvas elaboradas a partir de mediciones pluviográficas. Sin embargo, el programa posee un módulo a través del cual se pueden realizar estas curvas a partir de las precipitaciones máximas en 24 horas. Este proceso es útil en lugares donde no se dispone de información directa para la elaboración de las curvas IDF.

El proceso consiste en estimar las láminas de precipitación para diferentes duraciones de tormenta. Se puede hacer a partir de las metodologías de Dick y Peschke (ecuación 4.1) o las tormentas tipo del SCS (cuadro 4.1). Luego se calcula la intensidad dividiendo la lámina precipitada (mm) entre la duración convertida en horas. Finalmente, a través de un análisis de regresión exponencial múltiple, se determina la ecuación de intensidad en función de la duración y el periodo de retorno.

$$P_d = P_{24h} \left( \frac{d}{1440} \right)^{0.25} \quad (4.1)$$

Donde  $P_d$  es la precipitación total (mm) para la duración  $d$  obtenida a partir de la precipitación máxima en 24 horas ( $P_{24h}$  en mm).

Las distribuciones de lluvia según el SCS para diferentes tipos de tormenta estudiados en el territorio de los Estados Unidos se muestran en el cuadro ??.

Cuadro 4.1: Distribuciones de lluvia SCS

Hora $t$	$t/24$	Tipo I	Tipo IA	Tipo II	Tipo III
0	0.000	0.000	0.000	0.000	0.000
2.0	0.083	0.035	0.050	0.022	0.020
4.0	0.167	0.076	0.116	0.048	0.043
6.0	0.250	0.125	0.206	0.080	0.072
7.0	0.292	0.156	0.268	0.098	0.089
8.0	0.333	0.194	0.425	0.120	0.115
8.5	0.354	0.219	0.480	0.133	0.130
9.0	0.375	0.254	0.520	0.147	0.148
9.5	0.396	0.303	0.550	0.163	0.167
9.75	0.406	0.362	0.564	0.172	0.178
10.0	0.417	0.515	0.577	0.181	0.189
10.5	0.438	0.583	0.601	0.204	0.216
11.0	0.459	0.624	0.624	0.235	0.250
11.5	0.479	0.654	0.645	0.283	0.298
11.75	0.489	0.669	0.655	0.357	0.339
12.0	0.500	0.682	0.664	0.663	0.500
12.5	0.521	0.706	0.683	0.735	0.702
13.0	0.542	0.727	0.701	0.772	0.751
13.5	0.563	0.748	0.719	0.799	0.785
14.0	0.583	0.767	0.736	0.820	0.811
16.0	0.667	0.830	0.800	0.880	0.886
20.0	0.833	0.926	0.906	0.952	0.957
24.0	1.000	1.000	1.000	1.000	1.000

Fuente: (Chow et al., 1994)

## 4.2. Hietogramas

El hietograma, es un gráfico que representa la precipitación en intervalos de tiempo regulares. Para su elaboración, HidroCalc utiliza el método del bloque alterno y también lo hace a partir de las tormentas tipo del SCS (cuadro 4.1).

### Método del bloque alterno

Según (Chow et al., 1994), el método del bloque alterno es una manera de desarrollar un hietograma de diseño a partir de una curva IDF.

El método requiere de establecer una duración total  $T_d$  distribuida en  $n$  intervalos de tiempo  $\Delta t$ , de esta forma  $T_d = n\Delta t$ .

Posteriormente, se selecciona el periodo de retorno y se calcula la intensidad a partir de la curva IDF correspondiente y para cada una de las duraciones  $\Delta t$ ,  $2\Delta t$ ,  $3\Delta t$ , ..., la precipitación se calcula multiplicando la intensidad por la duración.

Se determinan las diferencias entre los valores consecutivos de precipitación y se encuentra la cantidad de precipitación que debe añadirse por cada unidad adicional de tiempo  $\Delta t$ . 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, (Chow et al., 1994).

## Capítulo 5

# Procesos precipitación - escorrentía

### 5.1. Caudales máximos

#### 5.1.1. Fórmulas empíricas

##### Método racional

Según (Aparicio Mijares, 1999), Este es posiblemente el método más antiguo de relación lluvia-escorrentía, probablemente elaborado entre 1851 y 1889. Toma en cuenta el área, la intensidad de precipitación y el coeficiente de escorrentía. Se utiliza principalmente en drenaje urbano.

La fórmula se describe de la siguiente manera:

$$Q = 0.278CiA_c \quad (5.1)$$

Donde  $C$  es un coeficiente de escurrimiento adimensional, que representa la fracción de la lluvia que escurre en forma directa y  $Q$  es el gasto máximo posible en  $m^3/s$  que puede producirse con una lluvia de intensidad  $i$  en  $mm/h$  en una cuenca de área  $A_c$  en  $Km^2$ . Dicho coeficiente varía entre 0 y 1.

No se recomienda aplicar el método racional en cuencas mayores a  $50Km^2$ .

### Método racional modificado

Este método está basado en el método racional, con la diferencia que incluye un coeficiente que permite utilizarlo para cuencas de hasta  $770 \text{ Km}^2$  y con tiempos de concentración de entre 0.25 y 24 horas. La fórmula es la siguiente:

$$Q = 0.278 C_i A_c K \quad (5.2)$$

Donde  $C$  es el coeficiente de escorrentía,  $Q$  el caudal unitario en  $\text{m}^3/\text{s}$ ,  $i$  representa la intensidad de lluvia en  $\text{mm}/\text{h}$ ,  $A_c$  el área de la cuenca en  $\text{Km}^2$  y  $K$  el coeficiente de uniformidad, adimensional.

El coeficiente de uniformidad puede ser calculado mediante la siguiente ecuación:

$$K = 1 + \frac{T_c^{1.25}}{T_c^{1.25} + 14} \quad (5.3)$$

Donde  $T_c$  es el tiempo de concentración en horas. Extraído de (Comunicaciones, 2015).

### Métodos de envolventes

Los métodos de envolventes toman en cuenta únicamente el área de la cuenca. No analizan la relación entre la lluvia y la escorrentía y deben emplearse sólo como aproximaciones gruesas de los caudales máximos probables, o bien cuando se carezca completamente de información, algo que hoy en día es poco frecuente. La idea fundamental de estos métodos es relacionar el caudal máximo  $Q$  con el área de la cuenca  $A_c$  en forma de la ecuación 5.4:

$$Q = \alpha A_c^\beta \quad (5.4)$$

En donde  $Q$  es el caudal máximo y  $\alpha$  y  $\beta$  son parámetros empíricos, que también pueden ser función de  $A_c$ .

La ecuación 5.5 es la planteada por **Creager**:

$$q = 1.303 C_c (0.386 A_c)^\alpha A_c^{-1} \quad (5.5)$$

Donde  $q$  es el caudal máximo por unidad de área,  $q = Q/A_c$ .

$$\alpha = \frac{0.936}{A_c^{0.048}} \quad (5.6)$$

Y  $C_c$  es un coeficiente empírico y  $A - c$  está en  $Km^2$ .

**Lowry** plantea su método mediante la ecuación 5.7:

$$q = \frac{C_L}{(A_c + 259)^{0.85}} \quad (5.7)$$

Donde  $C_L$  es otro coeficiente empírico. Los valores de  $C_c$  y  $C_L$  se determinan por regiones, llevando a una gráfica logarítmica los gastos unitarios máximos  $Q$  registrados contra sus respectivas áreas de cuenca y seleccionando el valor de  $C_c$  y  $C_L$  que envuelva a todos los puntos medidos.

El valor de  $C_c = 200$  es la envolvente para todos los puntos que analizó *Creager*, aunque  $C_c = 100$  da valores mucho más razonable y se usa como envolvente mundial. Para la fórmula de Lowry se puede tomar un valor  $C_L = 3500$  como mundial. Extraído de : (Aparicio Mijares, 1999).

### Método del número de curva

El método ha sido desarrollado por el Servicio de Conservación de Suelos (SCS) de los Estados Unidos. Es aplicable tanto a cuencas medianas como pequeñas. Se considera una serie de curvas que llevan el número  $N$ , que varía de 0 a 100. Un  $N = 1$  indica que toda la lluvia se infiltra, y un  $N = 100$ , que toda la lluvia escurre, estos números representan una analogía con el coeficiente de escorrentía o escurrimiento. Para la aplicación del método se necesitan datos de precipitación y parámetros morfológicos de la cuenca. El método está basado en la siguiente relación:

$$\frac{F}{S} = \frac{Q}{Pe} \quad (5.8)$$

Donde  $F$  es la infiltración real acumulada,  $S$  la infiltración potencial máxima,  $Q$  la escorrentía total acumulada y  $Pe$  la escorrentía potencial o exceso de precipitación. Todos los parámetros descritos se expresan en unidades de longitud.

La ecuación 5.8 es válida a partir del inicio de la escorrentía, donde  $Pe$  se determina de la ecuación 5.9.

$$Pe = P - Ia \quad (5.9)$$

La infiltración real acumulada  $F$ , se define a través de la ecuación 5.10.

$$F = Pe - Q \quad (5.10)$$

Donde la son las abstracciones iniciales y se definen como la precipitación acumulada hasta el inicio de la escorrentía, además es una función de la intercepción, el almacenamiento en depresiones e infiltración antes del inicio de la escorrentía.

Sustituyendo 5.10 en 5.8 resulta:

$$\frac{Pe - Q}{S} = \frac{Q}{Pe} \quad (5.11)$$

De donde se obtiene:

$$Q = \frac{Pe^2}{PE + S} \quad (5.12)$$

Reemplazando 5.12 en 5.9 se obtiene:

$$Q = \frac{(P - Ia)^2}{P - Ia + S} \quad (5.13)$$

Los autores del método obtuvieron la relación de la ecuación 5.14 a través de datos experimentales:

$$Ia = 2S \quad (5.14)$$

Dicha relación es bastante acertada para situaciones promedio, si se reemplaza la ecuación 5.14 en 5.13 se obtiene:

$$Q = \frac{(P - 0.2S)^2}{P + 0.8S} \quad (5.15)$$

En donde Q es la escorrentía total acumulada, P la precipitación y S la infiltración potencial máxima.

La ecuación 5.15 es la ecuación fundamental del método, todos los términos deben ser expresados en las mismas unidades. Adicionalmente, el SCS estudió un gran número de pequeñas cuencas para establecer una relación para estimar S a partir del número de curva N, la relación se muestra en la ecuación 5.16.

$$N = \frac{1000}{10 + S} \quad (5.16)$$

Dicha relación también puede expresarse mediante la ecuación 5.17.

$$S = \frac{1000}{N} - 10 \quad (5.17)$$

Donde  $S$  está expresado en pulgadas. La misma relación, con  $S$  en centímetros es descrita como sigue:

$$S = \frac{2540}{N} - 25.4 \quad (5.18)$$

Sustituyendo esta última expresión en la ecuación 5.15 resulta:

$$Q = \frac{[N(P + 5.08) - 508]^2}{N[N(P - 20.32) + 2032]} \quad (5.19)$$

Donde  $Q$  es la escorrentía total acumulada en cm,  $P$  la precipitación en cm y  $N$  el número de curva. Extraído de: (Villón Béjar, 2004).

### 5.1.2. Hidrogramas sintéticos

El hidrograma unitario desarrollado a partir de la información de lluvia y caudal en una cuenca, se aplica solamente para la cuenca y para el punto del cauce donde se midió la información de caudales. El procedimiento del hidrograma sintético se utiliza para desarrollar hidrogramas unitarios para otros puntos en la corriente dentro de la misma cuenca o para cuencas adyacentes de comportamiento similar.

#### Hidrograma SCS triangular

El hidrograma del Servicio de Conservación de Suelos (SCS) de EEUU o hidrograma Triangular, consta de dos partes. En la primera se hace una estimación del volumen de escorrentía resultante de una precipitación - escurrimiento directo, en la segunda se determina el tiempo de distribución del escurrimiento, incluyendo el caudal de punta.

Para precipitaciones ( $P$ ) menores que las abstracciones iniciales ( $I_a$ ), no tiene lugar el escurrimiento superficial ( $Q$ ).  $I_a$  consiste principalmente en pérdidas por intercepción, almacenamiento en depresiones e infiltración, antes de que se produzca el escurrimiento.

Para cantidad de lluvia en aumento, la curva  $Q$  en relación con  $P$  se aproxima asintóticamente a una línea recta paralela  $S$  denominada retención potencial máxima, que es la cantidad de lluvia que la cuenca puede absorber. Como hay muchas curvas que cumplen la condición de pasar por el punto  $I_a, Q$  y aproximarse asintóticamente a  $Q = P - S$ , se necesita definir otra condición para establecer la forma de la curva.

Esto se hace estableciendo que la relación entre la retención real  $P -$

$I_a - Q$ , y la retención potencial máxima  $S$  es igual a la relación entre la escorrentía real  $Q$  y la escorrentía potencial máxima  $P - I_a$ , tal como se muestra en la ecuación 5.20.

$$\frac{P - I_a - Q}{S} = \frac{Q}{P - I_a} \quad (5.20)$$

La experiencia práctica ha demostrado que  $I_a$  es aproximadamente el 20 % de la retención potencial máxima, así  $I_a = 0.2S$ , por lo que la ecuación de escurrimiento puede describirse como:

$$Q = \frac{(P - 0.2S)^2}{P + 0.8S} \quad (5.21)$$

El valor de  $S$  (en pulgadas) se relaciona con el número de curva de escorrentía (CN) por la definición:

$$CN = \frac{1000}{10 + S} \quad (5.22)$$

Donde CN representa el número de curva.

Para la distribución en el tiempo de la escorrentía, se procede de la siguiente manera: se hace uso del hidrograma unitario adimensional desarrollado por el SCS, cuya forma está predeterminada. La escala de tiempo del hidrograma se expresa en unidades del periodo de ascenso ( $T_p$ ) del mismo y los caudales de escurrimiento se expresan en unidades del caudal de escorrentía de punta ( $qp$ ).

Para convertir las ordenadas adimensionales del hidrograma en valores reales, debe conocerse el periodo de elevación ( $T_p$ ) y el caudal pico o de punta ( $qp$ ). Este último se obtiene utilizando la relación obtenida del hidrograma triangular sintético, como se muestra en la ecuación 5.23.

$$q_p = \frac{2Q}{T_p + T_R} \quad (5.23)$$

Donde  $qp$  es el caudal pico o de punta,  $Q$  el volumen de escurrimiento directo (mm),  $T_p$  es el periodo de elevación o tiempo hasta el caudal pico,  $T_R = 1.67T_p$  representa el tiempo desde el caudal pico hasta el final del escurrimiento directo - recesión. Si esta última relación es introducida en la ecuación 5.23 se obtiene:

$$q_p = 0.75 \frac{Q}{T_p} \quad (5.24)$$

Considerando el área de aporte en  $Km^2$ , de manera de llevar el valor puntual a un valor representativo para toda la cuenca, es escurrimiento

directo en mm, el tiempo de elevación en horas, el caudal punto queda expresado en  $m^3/s$ .

$$q_p = 0.208 \frac{A \times Q}{T_p} (m^3/s) \quad (5.25)$$

Dentro de esta metodología, se toma normalmente como periodo de tiempo unitario  $t_u = \frac{T_p}{4}$ . El periodo de elevación puede estimarse a partir del tiempo de concentración, según la relación empírica  $T_p = 0.7 \times T_c$ .

### Hidrograma sintético de Snyder

El hidrograma sintético de Snyder relaciona las características del hidrograma (tasa de flujo pico, flujo base, etc.) con las características de la cuenca.

Se obtuvo a partir de un estudio en cuencas localizadas principalmente en los montes Apalaches de los Estados Unidos y con tamaño que variaban desde cerca de 10 hasta 10000  $mi^2$ , Snyder encontró relaciones sintéticas para algunas características de un hidrograma unitario estándar. Algunas relaciones del mismo tipo fueron encontradas más tarde. A partir de las relaciones, pueden calcularse cinco características de un hidrograma unitario requerido para una duración en exceso de lluvia dada: el caudal pico por unidad de área de la cuenca,  $q_{pR}$ , el retardo de cuenca,  $t_{pR}$  (diferencia de tiempo entre el centroide del hidrograma de exceso de lluvia y el pico del hidrograma unitario), el tiempo base  $t_b$ , y los anchos  $W$  (en unidades de tiempo) del hidrograma unitario al 50 % y 75 % del caudal pico. Utilizando estas características puede dibujarse el hidrograma unitario requerido.

Snyder definió el hidrograma unitario estándar como aquel cuya duración de lluvia  $t_r$  está relacionada con el retardo de cuenca  $t_p$  por la ecuación 5.26.

$$t_p = 5.5t_r \quad (5.26)$$

Para un hidrograma unitario estándar encontró que:

- El retardo de la cuenca es:

$$t_p = C_1 C_t (LL_c)^{0.3} \quad (5.27)$$

Donde  $t_p$  está en horas,  $L$  es la longitud de la corriente principal en  $Km$  desde la salida de la cuenca hasta la divisoria de aguas arriba,

$L_c$  es la distancia en  $Km$  desde la salida de la cuenca hasta el punto de la corriente más cercano al centroide del área de la cuenca,  $C_1 = 0.75$  y  $C_t$  es un coeficiente basado en cuencas instrumentadas en la misma región.

- El caudal pico por unidad de área de drenajes en  $m^3/s \cdot Km^2$  del hidrograma unitario estándar es:

$$q_p = \frac{C_2 C_p}{t_p} \quad (5.28)$$

Donde  $C_2 = 2.75$  y  $C_p$  es un coeficiente basado en cuencas instrumentadas en la misma región.

Para calcular  $C_t$  y  $C_p$  de una cuenca instrumentada, los valores de  $L$  y  $L_c$  se miden utilizando un mapa de la cuenca. A partir de un hidrograma unitario deducido en la cuenca se obtienen los valores de su duración efectiva  $t_R$  en horas, su tiempo de retardo en la cuenca  $t_{pR}$  en horas y su caudal pico por unidad de área de drenaje,  $q_{pR}$ , en  $m^3/s \cdot Km^2 \cdot cm$ . Si  $t_{pR} = 5.5t_R$ , entonces  $t_r = t_R$ ,  $t_R = t_p$  y  $q_{pR} = q_p$ , y  $C_t$  y  $C_p$  se calculan utilizando las ecuaciones 5.27 y 5.28. Si  $t_{pR}$  es muy diferente de  $5.5t_R$ , el retardo de cuenca estándar es:

$$t_p = t_{pR} + \frac{t_r - t_R}{4} \quad (5.29)$$

Y las ecuaciones 5.26 y 5.29 se resuelven simultáneamente para encontrar  $t_r$  y  $t_p$ . Luego se calculan los valores de  $C_t$  y  $C_p$  de las ecuaciones 5.27 y 5.28 con  $q_{pR} = q_p$  y  $t_{pR} = t_p$ .

Cuando una cuenca no instrumentada parece ser similar a una cuenca instrumentada, los coeficientes  $C_t$  y  $C_p$  para la cuenca instrumentada pueden utilizarse en las ecuaciones anteriores para deducir el hidrograma unitario sintético requerido para la cuenca no instrumentada.

- La relación entre  $q_p$  y el caudal pico por unidad de área de drenaje  $q_{pR}$  del hidrograma unitario requerido es:

$$q_{pR} = \frac{q_p t_p}{t_{pR}} \quad (5.30)$$

- El tiempo base  $t_b$  en horas del hidrograma unitario puede determinarse utilizando el hecho de que el área bajo el hidrograma unitario

es equivalente a una escorrentía directa de 1 cm. Suponiendo una forma triangular para el hidrograma unitario, el tiempo base puede estimarse por:

$$t_b = \frac{C_3}{q_{pR}} \quad (5.31)$$

Donde  $C_3 = 5.56$ .

- El ancho en horas de un hidrograma unitario a un caudal igual a cierto porcentaje del caudal pico  $q_{pR}$  está dado por:

$$W = C_w q_{pR}^{-1.08} \quad (5.32)$$

Donde  $C_w = 1.22$  para un ancho del 75 % y 2.14 para un ancho de 50 %. Usualmente un tercio de este ancho se distribuye antes del momento en que ocurre el pico del hidrograma unitario y dos tercios después de dicho pico.

Extraído de (Chow et al., 1994).

## 5.2. Caudales mensuales

### 5.2.1. Modelo de Témez

El modelo de Témez es un método matemático simplificado de lluvia-escorrentía de paso mensual, cuyas entradas principales son las precipitaciones mensuales, evapotransporaciones mensuales y las salidas corresponden a las aportaciones del río para esos periodos.

La escorrentía total es la suma de la componente superficial y de la subterránea. La componente superficial es la fracción no infiltrada ni evaporada de la precipitación, mientras que la componente subterránea resulta de un modelo simple tipo celda. De manera resumida, el conjunto de ecuaciones empleadas en el modelo son las siguientes:

- **Evapotranspiración potencial**

$$EP_i = EO_i * K \quad (5.33)$$

- **Umbral de escorrentía,  $P_0$**

$$P_0 = C(H_{max} - H_{i-1}) \quad (5.34)$$

- **Excedente,  $T_i$**

$$T_i = 0 \Rightarrow P_i \leq P_0 \quad (5.35)$$

$$T_i = \frac{(P_i - P_0)^2}{P_i + \delta - 2P_0} \Rightarrow P_i > P_0 \quad (5.36)$$

Siendo,

$$\delta = ETP + H_{max} - H_{i-1} \quad (5.37)$$

- **Infiltración,  $I_i$**

$$I = I_{max} \frac{T_i}{T_i + I_{max}} \quad (5.38)$$

- **Aportación mínima,  $A_{min}$**

$$A_{min} = \frac{-D * 86.4(1 - \beta^2)f_{i-1}}{2S \ln \beta} \quad (5.39)$$

- **Aportación de un intervalo,  $A_F$**

$$A_F = A_{min} - I * \beta + I \quad (5.40)$$

- **Aportación buscada,  $A_i$**

$$A_i = A_{min} + T_i + \beta I \quad (5.41)$$

- **Total de agua disponible,  $X$**

$$X = H_{i-1} + P_i - T \quad (5.42)$$

- **Humedad remanente al final del intervalo,  $H_i$**

$$H_i = 0 \Rightarrow X \leq ETP \quad (5.43)$$

$$H_i = X - ETP \Rightarrow X \geq ETP \quad (5.44)$$

- **Aportación subterránea al cauce,  $f_i$**

$$f_i = \frac{S\beta}{86.4} \left[ \frac{86.4\beta f_{i-1}}{S} - \frac{2I \ln \beta}{D} \right] \quad (5.45)$$

### 5.2.2. Modelo de Lutz Scholz

El modelo propuesto por Lutz Scholz es de naturaleza combinada, dado que cuenta con una estructura determinística en la parte del cálculo de los caudales mensuales del año promedio, y una estructura estocástica para la generación de series extendidas de caudal. Fue desarrollado entre los años 1979 - 1980, en el marco del Plan Meris II y a través de la Cooperación Técnica de la República de Alemania.

El modelo requiere de parámetros físicos y meteorológicos fácilmente adquiribles en las cuencas peruanas.

#### ■ Ecuación de balance hídrico

La ecuación que describe el balance hídrico a nivel mensual en mm/mes es la siguiente:

$$CM_i = P_i - D_i + G_i - A_i \quad (5.46)$$

En donde  $CM_i$  es caudal mensual (mm/mes),  $P_i$  = la precipitación mensual sobre la cuenca (mm/mes),  $D_i$  = el déficit de escurrimiento (mm/mes) y  $A_i$  es el abastecimiento de la retención (mm/mes).

Asumiendo que para periodos largos (en este caso 1 año), el gasto y abastecimiento de la retención tienen el mismo valor, es decir  $G_i = A_i$ ; y que para el año promedio, una parte de la precipitación retorna a la atmósfera por evaporación.

Reemplazando (P-D) por (C\*P), y tomando en cuenta la transformación de unidades (mm/mes a  $m^3/s$ ) la ecuación 5.46 se convierte en:

$$Q = c' * C * P * AR \quad (5.47)$$

Que es la expresión básica del método racional, en donde  $Q$  es el caudal en  $m^3/s$ ,  $c'$  el coeficiente de conversión del tiempo (mes/s),  $C$  es el coeficiente de escurrimiento,  $P$  es la precipitación total mensual en (mm/mes) y  $AR$  el área de la cuenca en  $m^2$ .

#### ■ Coeficiente de escurrimiento

Se considera la fórmula propuesta por L. turc:

$$C = \frac{P - D}{P} \quad (5.48)$$

en donde  $C$  es el coeficiente de escurrimiento (mm/año),  $P$  la precipitación total (mm/año) y  $D$  el déficit de escurrimiento (mm/año).

Para la determinación de D se puede emplear la ecuación 5.49.

$$D = P \frac{1}{\left(0.9 + \frac{P^2}{L^2}\right)^{1/2}} \quad (5.49)$$

Además, el valor de L puede determinarse de la ecuación 5.50.

$$L = 300 + 25T + 0.05T^3 \quad (5.50)$$

En donde L es el coeficiente de temperatura y T la temperatura media anual ( $^{\circ}C$ ).

#### ■ Precipitación efectiva

Para determinar la precipitación efectiva, se asume que los caudales promedio observados en la cuenca pertenecen a un estado de equilibrio entre gasto y abastecimiento de la retención. La precipitación efectiva se determina para el coeficiente de escurrimiento promedio, de esta forma, la relación entre precipitación efectiva y precipitación total, resulta igual al coeficiente de escorrentía.

Para fines hidrológicos, se toma como precipitación efectiva la parte de la precipitación total mensual que corresponde al déficit según el método del USBR (precipitación efectiva en antítesis de la precipitación efectiva para los cultivos).

Para facilitar el cálculo, se establece el polinomio de quinto grado:

$$PE = a_0 + a_1P + a_2P^2 + a_3P^3 + a_4P^4 + a_5P^5 \quad (5.51)$$

Donde PE representa la precipitación efectiva en (mm/mes), P es la precipitación total mensual (mm/mes) y  $a_i$  es el coeficiente del polinomio.

La tabla 5.1 muestra los valores límite de la precipitación efectiva y la tabla 5.2 muestra los tres juegos de coeficientes  $a_i$  que permiten alcanzar por interpolación valores de C, comprendidos entre 0.15 y 0.45.

De esta forma, es posible llegar a la relación entre la precipitación efectiva y precipitación total:

$$C = \frac{Q}{P} = \sum_{i=1}^{12} \frac{PE_i}{P} \quad (5.52)$$

Cuadro 5.1: Límite superior para la precipitación efectiva

Curva	Precipitación efectiva
Curva I	PE = P-120.6 para $P > 177.8\text{mm/mes}$
Curva II	PE = P-86.4 para $P > 152.4\text{mm/mes}$
Curva III	PE = P-59.7 para $P > 127.0\text{mm/mes}$

Fuente:

Cuadro 5.2: Coeficientes para el cálculo de la Precipitación Efectiva

	Curva I	Curva II	Curva III
$a_0$	-0.018	-0.021	-0.028
$a_1$	+0.01850	+0.1358	+0.2756
$a_2$	+0.001105	-0.002296	-0.004103
$a_3$	-1204 E-8	+4349 E-8	+5534 E-8
$a_4$	+144 E-9	-89.0 E-9	+124 E-9
$a_5$	-285 E-12	-879 E-13	-142 E-11

Fuente:

En donde C es el coeficiente de escurrimiento o escorrentía, Q es el caudal anual, P representa la precipitación total anual y  $\sum_{i=1}^{12} PE_i$  es la suma de la precipitación efectiva mensual.

#### ■ Retención de la cuenca

Suponiendo que existe equilibrio entre el gasto y el abastecimiento de la reserva de la cuenca y además que el caudal total es igual a la precipitación efectiva anual, la contribución de la reserva hídrica al caudal se puede calcular según las ecuaciones 5.53 y 5.54.

$$R_i = CM_i - P_i \quad (5.53)$$

$$CM_i = PE_i + G_i - A_i \quad (5.54)$$

En las que  $CM_i$  representa el caudal mensual (mm/mes),  $PE_i$  a la precipitación efectiva mensual (mm/mes),  $R_i$  es la retención de la cuenca (mm/mes),  $G_i$  el gasto de la retención (mm/mes),  $A_i$  el abastecimiento de la retención (mm/mes).

Sumando los valores de  $G$  o  $A$  respectivamente, se halla la retención total de la cuenca para el año promedio, que para el caso de las cuencas de la sierra varía de 43 a 188 (mm/año).

■ **Relación entre descargas y retención**

En la estación seca, el gasto de la retención alimenta los ríos, constituyendo el caudal o descarga básica. La reserva o retención de la cuenca se agota al final de la estación seca; durante esta estación la descarga se puede calcular en base a la ecuación 5.55.

$$Q_t = Q_0 \exp^{-\alpha(t)} \quad (5.55)$$

En la cual,  $Q_t$  es la descarga en el tiempo  $t$ ,  $Q_0$  la descarga inicial,  $a$  el coeficiente de agotamiento y  $t$  el tiempo.

■ **Coefficiente de agotamiento**

Mediante la ecuación 5.55, se puede calcular el coeficiente de agotamiento  $a$ , en base a datos hidrométricos. El coeficiente no es constante, sino que disminuye gradualmente a lo largo de la estación seca. Esta variación puede desprejarse empleando un valor promedio. El coeficiente de agotamiento de la cuenca tiene una dependencia logarítmica del área de la cuenca:

$$a = f(\ln AR) \quad (5.56)$$

En principio, es posible determinar el coeficiente de agotamiento real mediante aforos sucesivos en el río durante la estación seca, pero cuando no se puede, se puede recurrir a las ecuaciones desarrolladas para la determinación del coeficiente  $a$  para cuatro clases de cuencas:

- Cuencas con agotamiento muy rápido. Debido a temperaturas elevadas ( $> 10^\circ C$ ) y retención que va de reducida (50 mm/año) a mediana (80 mm/año):

$$a = -0.00252(\ln AR) + 0.034 \quad (5.57)$$

- Cuencas con agotamiento rápido. Retención entre 50 y 80 mm/año y vegetación poco desarrollada (puna):

$$a = -0.00252(\ln AR) + 0.030 \quad (5.58)$$

- Cuencas con agotamiento mediano. Retención mediana (80 mm/año) y vegetación mezclada (pastos, bosques y terrenos cultivados):

$$a = -0.00252(\ln AR) + 0.023 \quad (5.59)$$

En donde  $a$  es el coeficiente de agotamiento por día,  $AR$  el área de la cuenca ( $Km^2$ ),  $EP$  la evapotranspiración potencial anual (mm/año),  $T$  la duración de la temporada seca (días) y  $R$  es la retención total de la cuenca (mm/año).

■ **Almacenamiento hídrico**

Tres tipos de almacenes hídricos naturales que inciden en la retención de la cuenca son considerados:

- Acuíferos
- Lagunas y pantanos
- Nevados

■ **Abastecimiento de la retención**

El abastecimiento durante la estación lluviosa es uniforme para cuencas ubicadas en la misma región climática. Los coeficientes mensuales, expresados en porcentaje del almacenamiento total anual se muestran en la tabla 5.3.

Cuadro 5.3: Coeficientes para el cálculo de la Precipitación Efectiva

Región	Oct	Nov	Dic	Ene	Feb	Mar	Total
Cusco	0	5	35	40	20	0	100
Huancavelica	10	0	35	30	20	5	100
Junín	10	0	25	30	30	5	100
Cajamarca	25	5	0	20	25	35	100

*Fuente:*

La lámina de agua  $A_i$  que entra en la reserva de la cuenca se muestra en forma de déficit mensual de la precipitación efectiva  $PE_i$ . Se calcula mediante la ecuación 5.60.

$$A_i = a_i \left( \frac{R}{100} \right) \quad (5.60)$$

Siendo  $A_i$  abastecimiento mensual déficit de la precipitación efectiva (mm/mes),  $a_i$  es el coeficiente de abastecimiento (%) y  $R$  la retención de la cuenca (mm/año).

■ **Determinación del caudal mensual para el año promedio**

Está basado en la ecuación fundamental que describe el balance

hídrico mensual a partir de los componentes descritos anteriormente:

$$CM_i = PE_i + G_i - A_i \quad (5.61)$$

Donde  $CM_i$  es el caudal del mes  $i$  (mm/mes),  $PE_i$  la precipitación efectiva del mes  $i$  (mm/mes),  $G_i$  es el gasto de la retención del mes  $i$  (mm/mes) y  $A_i$  el abastecimiento del mes  $i$  (mm/mes).

■ **Generación de caudales mensuales para periodos extendidos**

Para generar una serie sintética de caudales para periodos extendidos, se ha implementado un modelo estocástico que consiste en una combinación de un proceso markoviano de primer orden, según la ecuación 5.62 con una variable de impulso, que en este caso es la precipitación efectiva de la ecuación 5.63.

$$Q_t = f(Q_{t-1}) \quad (5.62)$$

$$Q = g(PE_t) \quad (5.63)$$

Con la finalidad de aumentar el rango de valores generados y obtener una optima aproximación a la realidad, se utiliza además una variable aleatoria.

$$Z = z(S)\sqrt{1-r^2} \quad (5.64)$$

La ecuación integral para la generación de caudales mensuales es:

$$Q_t = B1 + B2(Q_{t-1}) + B3(PE_t) + z(S)\sqrt{1-r^2} \quad (5.65)$$

En donde  $Q_t$  es el caudal del mes  $t$ ,  $Q_{t-1}$  el caudal del mes anterior,  $PE_t$  la precipitación efectiva del mes  $t$  y  $B1$  el factor constante o caudal básico.

Los parámetros  $B1$ ,  $B2$ ,  $B3$ ,  $r$  y  $S$  sobre la base de los resultados del modelos para el año promedio por un cálculo de regresión con  $Q_t$  como valor dependiente y  $Q_{t-1}$  y  $PE_t$ , como valores independientes.

El proceso de generación requiere de un valor inicial, el cual puede ser obtenido en una de las siguiente formas:

- Empezar el cálculo en el mes para el cual se dispone de un aforo.
- Tomar como valor inicial el caudal promedio de cualquier mes.

- Empezar con un caudal cero, calcular un año y tomar el último valor como  $Q_0$  sin considerar estos valores en el cálculo de los parámetros estadísticos del período generado.

#### ■ Restricciones del modelo

El modelo presenta ciertas restricciones de uso o aplicación tales como:

- El uso de los modelos parciales, únicamente dentro del rango de calibración establecido.
- Su uso es únicamente para el cálculo de caudales mensuales promedio.
- Los registros generados en el período de secas presentan una mayor confiabilidad que los valores generados para la época lluviosa.
- La aplicación del modelo se restringe a las cuencas en las que se ha calibrado sus parámetros (sierra peruana).

Es importante tener en cuenta las mencionadas restricciones a fin de garantizar el buen desempeño del modelo.

## Capítulo 6

# Redes Neuronales Artificiales

Las redes neuronales son una alternativa a la computación clásica, basadas en algoritmos inspirados en el funcionamiento del cerebro y sus neuronas, éstas tienen un buen desempeño en problemas que la segunda no puede resolver, principalmente es empleada en campos relacionados con el reconocimiento de patrones, planeamiento, predicción, control y optimización. A diferencia de la computación clásica que es secuencial, la red neuronal trabaja con entradas paralelas y el resultado no es almacenado, es un estado de la red cuando se logra el equilibrio. Su aprendizaje no se almacena en instrucciones, sino en topología y los valores de sus conexiones. Sin embargo no son un reemplazo a los sistemas tradicionales, sino más bien se consideran un complemento. Pueden resolver problemas físicos sin conocer, valga la redundancia, las relaciones físicas de las variables. Para entrenarlas, se presentan de manera repetitiva y sistemática una serie de entradas y sus respectivas salidas (Delgado, 1998).

### Analogía entre RNA y Redes neuronales biológicas

Cajal, 1991, demostró que el tejido del cerebro no es una masa continua sino una red de unidades discretas denominadas neuronas (figura 6.1). Una neurona biológica está compuesta por el *soma* o cuerpo que contiene el núcleo y recibe información de otras neuronas a través de la sinapsis en su superficie. Las *dendritas* son estructuras que parten del soma con ramificaciones, se especializan en la recepción de señales de otras células nerviosas por medio de conexiones sinápticas (Delgado, 1998). El *axón* permite enviar impulsos a otras células nerviosas. La *sinapsis* es una conexión entre dos células nerviosas. Las sinapsis pueden ser excitativas

o inhibitorias según el neurotransmisor que se libere, cada neurona recibe de 10000 a 100000 sinapsis y su axón realiza una cantidad similar de sinapsis.

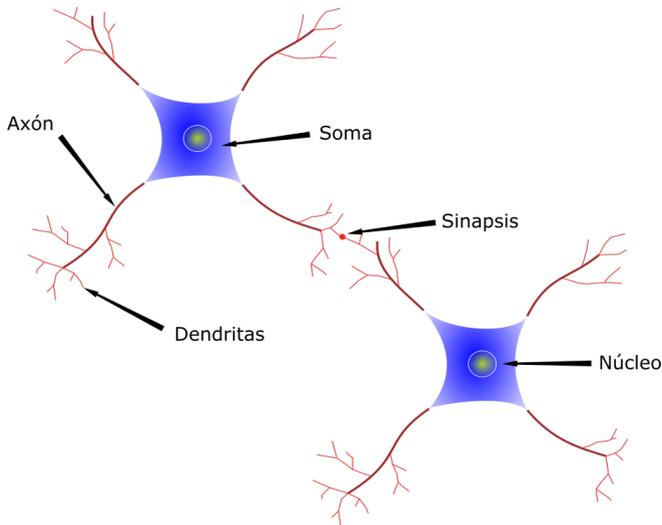


Figura 6.1: Neuronas biológicas.

*Fuente: Elaboración propia.*

En estado de reposo, la membrana de la neurona mantiene una diferencia de potencial de  $-70\text{mv}$ . En este estado es más permeable a los iones de potasio que a los iones de sodio. Cuando se estimula la célula, la permeabilidad al sodio se incrementa, y entonces se produce una entrada de cargas positivas que produce un impulso y cuando éste alcanza los terminales del axón de la neurona presináptica, este induce la liberación de moléculas neurotransmisoras. Los transmisores se difunden y alcanzan los receptores de la membrana postsináptica. La frecuencia de los impulsos y la intensidad de excitación obedecen a una función (figura 6.2) (Delgado, 1998).

El psicólogo Hebb en 1949, propone que el aprendizaje en las neuronas ocurre por la actividad neuronal coincidente, esto se conoce como la **ley de Hebb o aprendizaje Hebbiano**: “Cuando el axón de la célula A excita a la célula B y repetidamente toma lugar en su activación, ocurre algún proceso de crecimiento o cambio metabólico en una o ambas células tal que la eficacia de A, como una de las células que dispara a B, se incrementa”, esto se conoce como mecanismo asociativo pre-post, figura 6.3.

Tauc y Kandel en 1963 propusieron una segunda regla de aprendizaje,

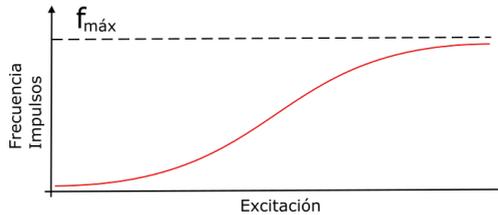


Figura 6.2: Variación de la frecuencia de descarga neuronal según la excitación.  
 Fuente: (Delgado, 1998).

encontraron que la conexión sináptica entre dos neuronas se puede fortalecer sin actividad de la célula postsináptica, una tercera neurona, denominada neurona moduladora, incrementa la liberación del neurotransmisor de la célula presináptica, figura 6.4.

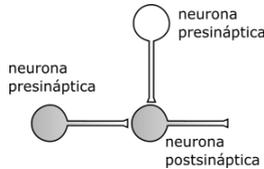


Figura 6.3: Coincidencia pre-post.  
 Fuente: (Delgado, 1998).

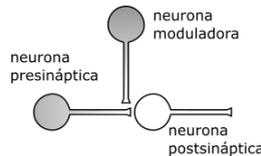


Figura 6.4: Coincidencia pre-moduladora.  
 Fuente: (Delgado, 1998).

W.S. McCulloch y W. Pitts en 1943 proponer un modelo simple de la neurona biológica.

Las entradas son  $x_i$ ,  $i = 1, \dots, n$ , las sinapsis se modelan por los pesos  $\omega_i$ . La neurona se dispara si la suma ponderada de las entradas excede el valor umbral  $\omega_0$ , en este caso la salida de la neurona toma dos valores  $(-1, +1)$ . En notación matemática:

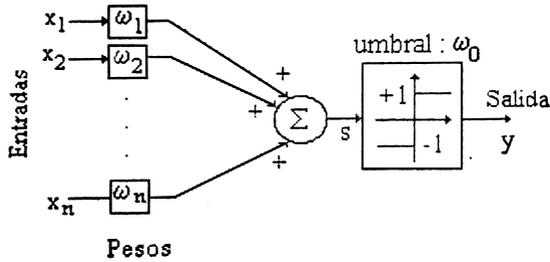


Figura 6.5: Modelo de neurona artificial.  
Fuente: (Delgado, 1998).

$$y = \sigma\left(\sum_{i=1}^n \omega_i \cdot x_i + \omega_0\right) \quad (6.1)$$

donde:

$$\sigma(\theta) = \begin{cases} -1, & \theta < 0 \\ +1, & \theta \geq 0 \end{cases} \quad (6.2)$$

la función  $\sigma(\theta)$  se denomina función de activación. Hay varios tipos de funciones de activación, se muestran en la figura 6.6.

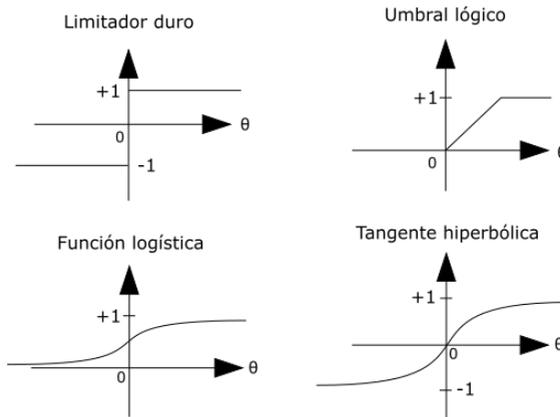


Figura 6.6: Funciones de activación comunes.  
Fuente: (Delgado, 1998).

## Red Neuronal de una Capa

En la figura 6.7 se muestra una red neuronal con nodos de entrada y una capa de neuronas a la salida. Cada círculo de la figura 6.7 corresponde al modelo de la figura 6.5. Sin embargo estas redes neuronales tienen una importante limitación: sólo pueden resolver problemas separables linealmente. Esta limitación de las redes neuronales de una capa causó que la investigación en redes neuronales entrara en un periodo de hibernación científica que duró un aproximado de 20 años.

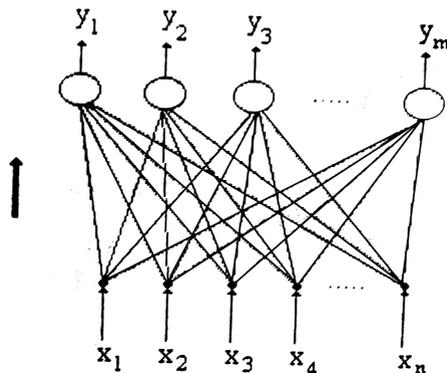


Figura 6.7: Red Neuronal de una Capa.

Fuente: (Delgado, 1998).

## Algoritmos para el entrenamiento de una neurona

Existen dos tipos de algoritmos: (i) Supervisados: algoritmos de corrección del error y el descenso de gradiente y (ii) No supervisados: no requieren patrones entrada-salida, la neurona establece sus propios pesos a partir de los datos de entrada.

A continuación pasaremos a describir los supervisados, que son lo más comunes.

### ■ Algoritmos de corrección del error

- **$\alpha$ -Least Mean Square ( $\alpha$ -LMS).** Este algoritmo no funciona correctamente para datos fuertemente correlacionados debido a que en estos casos su aprendizaje es muy lento, razón por la cual no se verá a profundidad.

- **Regla del perceptrón.** Se emplea en neuronas con función de activación, en estos casos “S” es la salida de las neuronas y “y” es la salida de la función de activación. El error estará dado por:

$$\hat{e}^k = d^k - y^k, \quad (6.3)$$

en donde:

$$y^k = \sigma(S^k) \quad (6.4)$$

y  $d^k$  es la salida deseada. Es importante resaltar que el superíndice  $k$  no representa exponente, es sólo un indicador del número de iteración. La actualización del vector de pesos se hace con la ecuación 6.5.

$$\omega^{k+1} = \omega^k + \alpha \cdot \hat{e}^k \cdot x^k \quad (6.5)$$

- **Algoritmos basados en el gradiente** De igual manera se emplean los dos tipos anteriores. Se enuncia (Delgado, 1998): “Un algoritmo basado en el gradiente toma un vector de pesos inicial, calcula la función de error y su gradiente, y luego obtiene un nuevo vector de pesos modificando el vector de pesos inicial en dirección opuesta al gradiente de la función de error, por lo general, la función de error cuadrático. El proceso se repite hasta que el error se encuentra en el límite establecido”. La regla está dada por:

$$\omega^{k+1} = \omega^k + \alpha(-\nabla^k) \quad (6.6)$$

donde  $\alpha$  es la tasa o tasa de aprendizaje y  $\nabla$  es el gradiente de la función de error respecto de  $\omega^k$

- **$\alpha$ -LMS.** No se emplea una función de propagación, emplea un combinador lineal. Para el patrón de entrada  $x^k$ , el error cuadrático es:

$$(e^k)^2 = \{d^k - (\omega^k)^T \cdot x^k\}^2 \quad (6.7)$$

de esta función debe obtenerse su vector gradiente:

$$\nabla^k = \frac{\partial (e^k)^2}{\partial \omega^k} = \left[ \frac{\partial (e^k)^2}{\partial \omega_0^k} \cdot \frac{\partial (e^k)^2}{\partial \omega^k} \cdots \frac{\partial (e^k)^2}{\partial \omega_n^k} \right]^T \quad (6.8)$$

por supuesto, al hablar de vector “T” significa transpuesta, y el nuevo vector de pesos será:

$$\omega^{k+1} = \omega^k - \alpha \cdot \frac{\partial (e^k)^2}{\partial \omega^k} \quad (6.9)$$

Derivando la ecuación 6.9 y reemplazando en 6.6, finalmente se obtiene:

$$\omega^{k+1} = \omega^k + 2 \cdot \alpha \cdot e^k \cdot x^k \quad (6.10)$$

Finalmente la ecuación 6.10 se utiliza para activar los pesos de una neurona sin función de activación.

- **Propagación inversa o “Back Propagation”**. En este tipo de algoritmos se necesita una función de activación. La figura 6.8 muestra una neurona con función de activación *tangente hiperbólica*, la cual es una función derivable por ser continua.

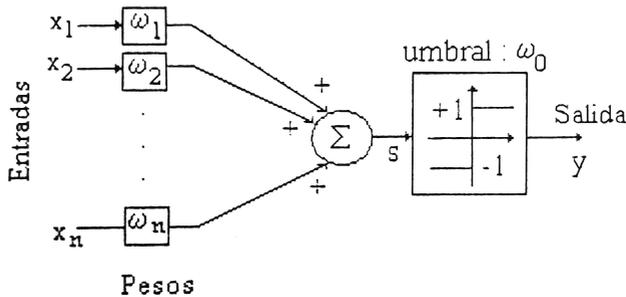


Figura 6.8: RNA con función de activación tangente hiperbólica.  
Fuente: (Delgado, 1998).

La salida de esta neurona estará dada por:

$$y^k = \tanh(S^k) = \sigma(S^k) \quad (6.11)$$

El error no-lineal está dado por:

$$\tilde{e}^k = d^k - \sigma(S^k) = d^k - \sigma((\omega^k)^T \cdot x^k) \quad (6.12)$$

El gradiente del error cuadrático es:

$$\nabla^k = \frac{\partial (\tilde{e}^k)^2}{\partial \omega^k} = 2 \cdot \tilde{e}^k \frac{\partial \tilde{e}^k}{\partial \omega^k} \quad (6.13)$$

Reemplazando el error (ecuación 6.12) en la derivada y empleando la regla de la cadena, finalmente tendremos:

$$\nabla^k = -2\alpha \tilde{e}^k \sigma'(S^k) \cdot x^k \quad (6.14)$$

Sustituyendo en la ecuación 6.10 (ecuación de actualización de pesos):

$$\omega^{k+1} = \omega^k + 2\alpha^k \sigma'(S^k) x^k \quad (6.15)$$

La derivada de la función de activación tangente hiperbólica es:

$$\sigma'(S^k) = \frac{d \tanh(S^k)}{dS^k} = 1 - (y^k)^2 \quad (6.16)$$

Se concluye que la expresión para actualizar el vector de pesos en una neurona con función de activación tangente hiperbólica es:

$$\omega^{k+1} = \omega^k + 2\alpha^k \{1 - (y^k)^2\} \cdot x^k \quad (6.17)$$

### Neuronas estáticas

Los problemas adecuados para solucionar mediante redes neuronales son aquellos que no tienen solución computacional precisa o requieren algoritmos muy extensos como en el caso del reconocimiento de imágenes (Delgado, 1998). Entre las ventajas de las redes neuronales podemos enunciar las siguientes: sintetizar algoritmos a través de un proceso de aprendizaje, no se requiere conocer los detalles matemáticos, sólo estar familiarizado con los datos de trabajo, es óptimo para desarrollar problemas no-lineales y es robusta, pueden fallar algunos elementos y seguir funcionando. Entre las desventajas están: deben entrenar para cada problema y consumen mucho CPU, necesitan muchos datos y finalmente, su aspecto es complejo para un observador externo.

- **Red tipo perceptrón.** Poseen una capa de conexiones modificables entre sus nodos de entrada y las neuronas de salida, el conocimiento se almacena en el patrón de conexiones.
- **Red multicapa.** La red neuronal multicapa (figura 6.9) puede resolver problemas de clasificación de patrones no separables linealmente como la función lógica XOR.

Para esta red neuronal, la salida de las neuronas está dada por:

$$h_i = \sigma(S_i) = \sigma\left(\sum_{j=1}^N \omega_{ij} \cdot x_{ij} + \omega_{i0}\right) \quad (6.18)$$

donde  $i = 1, \dots, N$ ,  $\omega_{ij}$  es el peso que une la neurona  $i$  con la entrada  $j$  y  $\omega_{i0}$  es el peso asociado a la entrada umbral  $x_0 = 1$ . Las salidas de la red son:

$$f_l = \sum_{i=1}^N c_{li} h_i \quad (6.19)$$

donde  $l = 1, \dots, m$ ,  $c_{li}$  es el peso que une la salida  $l$  con la neurona  $i$ . La red no tiene entradas umbral ni pesos umbral en las salidas

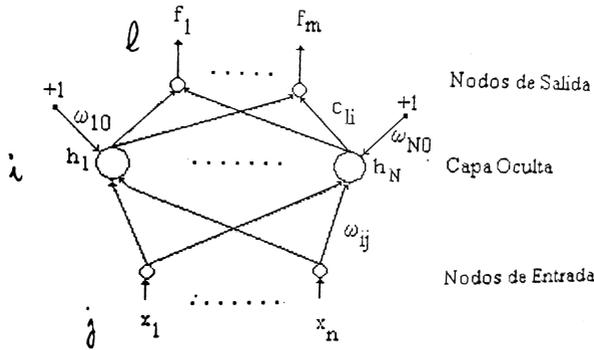


Figura 6.9: RNA Multicapa n-N-m  
Fuente: (Delgado, 1998).

cuando se trata de nodos. Es posible tener neuronas en la capa de salida, en este caso:

$$f_i = \sigma \left( \sum_{i=1}^N c_{ij} h_i + c_{i0} \right) \quad (6.20)$$

Esto es, la suma ponderada pasa por la función de activación y existe un peso umbral por neurona de salida.

- **Entrenamiento con algoritmo de propagación inversa.** Se necesita que las neuronas tengan una función de activación continua, se utiliza el gradiente de la función de error para actualizar los pesos.

Durante la presentación de un patrón, el error cuadrático de las salidas estará dado por:

$$e^2 = (d_1 - y_1)^2 + (d_2 - y_2)^2 \quad (6.21)$$

Donde  $d_1$  y  $d_2$  son las salidas deseadas:

$$y_1 = c_{11}h_1 + c_{12}h_2 + c_{13}h_3 \quad (6.22)$$

$$y_2 = c_{21}h_1 + c_{22}h_2 + c_{23}h_3 \quad (6.23)$$

Calculando el gradiente del error cuadrático:

$$\frac{\partial e^2}{\partial c_{li}} = -2e_i h_i \quad (6.24)$$

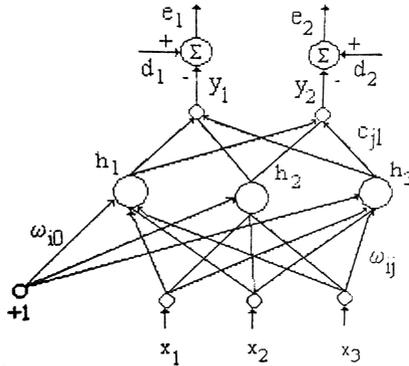


Figura 6.10: RNA Multicapa 3-3-2.  
Fuente: (Delgado, 1998).

Entonces la regla para actualizar los pesos de salida es:

$$c_{li}^{k+1} = c_{li}^k + 2\alpha e_i h_i \quad (6.25)$$

donde  $l$  va desde  $1, \dots, m$  e  $i$  desde  $1, \dots, N$ . La salida de las neuronas ocultas está dada por:

$$h_1 = \sigma(S_1), \quad h_2 = \sigma(S_2), \quad h_3 = \sigma(S_3) \quad (6.26)$$

El gradiente del error cuadrático respecto de los pesos de la capa oculta de la red neuronal es:

$$\frac{\partial e^2}{\partial \omega_{ij}} = -2\delta_l \sigma'(S_i) x_j \quad (6.27)$$

donde  $\delta$  es el error y está dado por:

$$\delta_i = \sum_{l=1}^m c_{li} e_i = \sum_{l=1}^m c_{li} (d_l - y_l) \quad (6.28)$$

Entonces, la regla para actualizar los pesos de entrada es:

$$\omega_{ij}^{k+1} = \omega_{ij}^k + 2\alpha \delta_i \sigma'(S_i) x_i \quad (6.29)$$

donde  $j$  va desde  $1, \dots, n$  y la derivada de la función de activación es:

$$\sigma'(\theta) = \frac{\partial \sigma(\theta)}{\partial \theta} \quad (6.30)$$

El algoritmo se denomina propagación inversa porque los errores se propagan en sentido inverso a las entradas para actualizar los pesos de la red

- **Generalización en redes neuronales.** Uno de los aspectos más sobresaliente de las RNA es su habilidad para reconocer o clasificar patrones nunca antes presentados a la red.

(Extraído de: (Delgado, 1998))



# Bibliografía

- Aparicio Mijares, F. J. (1999). *Fundamentos de hidrología de superficie*.
- Basualdo, A. (2015). Manual de buenas prácticas para la generación, el almacenamiento y la difusión de información climática en instituciones y organismos del mercosur. Technical report, IICA, San José (Costa Rica).
- Buishand, T. A. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of hydrology*, 58(1-2):11–27.
- Campos Bonilla, B. (2017). Beir1718-distribuciones de probabilidad.
- Chereque Morán, W. (1989). Hidrología: para estudiantes de ingeniería civil.
- Chow, V. T., Maidment, D. R., Mays, L. W., Saldarriaga, J. G., et al. (1994). *Hidrología aplicada*.
- Comunicaciones, M. (2015). Manual de hidrología, hidráulica y drenaje. *Proyecto. Lima: Ministerio de Transportes y Comunicaciones, Lima*.
- Cooley, D. (2013). Return periods and return levels under climate change. In *Extremes in a changing climate*, pages 97–114. Springer.
- Correa, J. C., Iral, R., and Rojas, L. (2006). Estudio de potencia de pruebas de homogeneidad de varianza. *Revista colombiana de estadística*, 29(1):57–76.
- Delgado, A. (1998). *Inteligencia artificial y minirobots*. Ecoe Ediciones.
- Dodge, Y. and Commenges, D. (2006). *The Oxford dictionary of statistical terms*. Oxford University Press on Demand.
- ESRI (12/08/2018). Vista general del conjunto de herramientas de Hidrología. <http://desktop.arcgis.com/es/arcmap/10.3/tools/spatial-analyst-toolbox/an-overview-of-the-hydrology-tools.htm>.

- Fattorelli, S. and Fernández, P. C. (2007). *Diseño hidrológico*. Zeta Editores.
- Hosking, J. R. M. and Wallis, J. R. (2005). *Regional frequency analysis: an approach based on L-moments*. Cambridge university press.
- Ibáñez, S., Moreno, H., and Gisbert, J. (2011). Morfología de las cuencas hidrográficas. *Universidad politécnica de Valencia*.
- Jardí Porqueras, M. (1985). Forma de una cuenca de drenaje. análisis de las variables morfométricas que nos la definen. *Revista de Geografía*, 19(1):41–68.
- Jasiewicz, J. (16/08/2018). The module r.stream.order.  
<https://svn.osgeo.org/grass/grass-addons/grass7/raster/r.stream.order/r.stream.order.html>.
- Juan, Á. A., Sedano, M., and Vila, A. (2006). La distribución normal. *Universidad abierta de Cataluña*.
- Koutrouvelis, I. A. and Canavos, G. C. (1999). Estimation in the pearson type 3 distribution. *Water resources research*, 35(9):2693–2704.
- Kvam, P. H. and Vidakovic, B. (2007). *Nonparametric statistics with applications to science and engineering*, volume 653. John Wiley & Sons.
- Machiwal, D. and Jha, M. K. (2012). *Hydrologic time series analysis: theory and practice*. Springer Science & Business Media.
- Massey Jr, F. J. (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78.
- Minnaard, C. L., Condesse, V., Minnaard, V., and Rabino, M. C. (2005). Los gráficos de caja: un recurso innovador. *Revista Iberoamericana de Educación*, 35.
- Moreno Grande, F. A., Esquivel Jimenez, J. R., et al. (2015). Estudio morfométrico de la cuenca del río azul, afluente del río calima, departamento del valle del cauca.
- Ordoñez, J. (2011). Contribuyendo al desarrollo de una cultura del agua y la gestión integral del recurso hídrico.
- Organization, W. M. (1986). *Manual for estimation of probable maximum precipitation*. Secretariat of the World Meteorological Organization.
- Ortiz, O. (2008). *Hidrología de superficie*. Universidad Nacional de Cajamarca.

- Pearson, K. (1895). X. contributions to the mathematical theory of evolution.—ii. skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London.(A.)*, (186):343–414.
- Phine, H. N. and Hira, M. A. (1983). Log pearson type-3 distribution: Parameter estimation. *Journal of Hydrology*, 64(1-4):25–37.
- Rojas, M. (2003). Técnicas estadísticas paramétricas y no paramétricas equivalentes: Resultados comparativos por simulación. *Instituto deficiencias matemáticas. Ingeniería en estadística informática. Escuela Superior Politécnica del Litoral. Ecuador*.
- Sánchez Turcios, R. A. (2015). t-student: Usos y abusos. *Revista mexicana de cardiología*, 26(1):59–61.
- Thom, H. C. (1958). A note on the gamma distribution. *Monthly Weather Review*, 86(4):117–122.
- Upegui, Jorge Julián Vélez, G. A. B. (2011). Estimación del tiempo de concentración y tiempo de rezago en la cuenca experimental urbana de la quebrada san luis, manizales. *Dyna*, 78(165):58–71.
- Villón Béjar, M. (2002). *Hidrología Estadística*. Instituto Tecnológico de Costa Rica.
- Villón Béjar, M. (2004). *Hidrología*. Instituto Tecnológico de Costa Rica.
- Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29(3/4):350–362.
- Zaiontz, C. (10/09/2018). Fligner killeen test.  
<http://www.real-statistics.com/one-way-analysis-of-variance-anova/homogeneity-variances/fligner-killeen-test/>.
- Zogheib, B. and Elsheli, A. (2015). Approximations to the t distribution. *Applied Mathematical Sciences*, 9(49):2445–2449.
- Zogheib, B. and Hlynka, M. (2009). *Approximations of the standard normal distribution*. University of Windsor, Department of Mathematics and Statistics Windsor, Canada.



© Unu Ingenieros

Telf. (+51) 961304698

E-mail 1: [contacto@unuingenieros.com](mailto:contacto@unuingenieros.com)

E-mail 2: [unu.ingenieros@gmail.com](mailto:unu.ingenieros@gmail.com)

[www.unuingenieros.com](http://www.unuingenieros.com)

CAJAMARCA - PERÚ

**Hidráulica - Hidrología - SIG&Teledetección**



ONLINE