

56

**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

CAMPUS ARAGON

290661

“ESTIMACION DE EVENTOS HIDROLOGICOS
EN CASOS ESPECIALES”

TESIS PROFESIONAL

QUE PARA OBTENER EL TITULO DE:

INGENIERO CIVIL

PRESENTA

JOSE MIGUEL VIDAL BARRAGAN

MARZO DE 2001



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



**EL PRESENTE TRABAJO SE IMPRIMIÓ CON EL APOYO DE
LA COMISIÓN NACIONAL DEL AGUA**



**ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
ARAGÓN
DIRECCIÓN**

**UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO**

**JOSÉ MIGUEL VIDAL BARRAGÁN
PRESENTE.**

En contestación a la solicitud de fecha 31 de octubre del año en curso, relativa a la autorización que se le debe conceder para que el señor profesor, Ing. MARTÍN ORTÍZ LEÓN pueda dirigirle el trabajo de tesis denominado "ESTIMACIÓN DE EVENTOS HIDROLÓGICOS EN CASOS ESPECIALES", con fundamento en el punto 6 y siguientes, del Reglamento para Exámenes Profesionales en esta Escuela, y toda vez que la documentación presentada por usted reúne los requisitos que establece el precitado Reglamento; me permito comunicarle que ha sido aprobada su solicitud.

Aprovecho la ocasión para reiterarle mi distinguida consideración.

**Atentamente
"POR MI RAZA HABLARÁ EL ESPÍRITU"
San Juan de Aragón, México, 15 de noviembre del 2000
EL DIRECTOR**

M en R.I. CARLOS EDUARDO LEVY VÁZQUEZ

C p Secretaría Académica.
C p Jefatura de la Carrera de Ingeniería Civil.
C p Asesor de Tesis.

CELV/AIR/BOC'11a.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
CAMPUS ARAGÓN

SECRETARÍA ACADÉMICA

UNIVERSIDAD NACIONAL
AVENIDA II
MÉXICO

Ing. GILBERTO GARCÍA SANTAMARÍA GONZÁLEZ
Jefe de la Carrera de Ingeniería Civil,
Presente.

En atención a la solicitud de fecha 5 de diciembre del año en curso, por la que se comunica que el alumno JOSÉ MIGUEL VIDAL BARRAGÁN, de la carrera de Ingeniero Civil, ha concluido su trabajo de investigación intitulado "ESTIMACIÓN DE EVENTOS HIDROLÓGICOS EN CASOS ESPECIALES", y como el mismo ha sido revisado y aprobado por usted, se autoriza su impresión; así como la iniciación de los trámites correspondientes para la celebración del Examen Profesional.

Sin otro particular, reitero a usted las seguridades de mi atenta consideración.

Atentamente
"POR MI RAZA HABLARÁ EL ESPÍRITU"
San Juan de Aragón, México, 6 de diciembre del 2000
EL SECRETARIO


Lic. ALBERTO IBARRA ROSAS

C p Asesor de Tesis ✓
C p Interesado.

AIR/RCC/vr

AGRADECIMIENTOS

A MIS PADRES

A quienes nunca podré pagar todos sus sacrificios y desvelos por formarme y educarme, ya que gracias a ellos estoy a punto de lograr la ilusión mas grande de mi vida, ser un ingeniero civil. Perdón por todos los malos momentos que les he hecho pasar, espero no haberlos defraudado, hojala se sientan orgullosos de mi como yo lo estoy de ustedes. Gracias por todo.

A MIS HERMANOS

Gracias por el apoyo que me han brindado durante el tiempo que duraron mis estudios y el que me han brindado siempre. Espero que siempre exista confianza entre nosotros y que nos ayudemos y apoyemos en todo momento, como lo que somos una familia.

A MI ABUELA, TIAS Y TIOS

Por la convivencia que siempre ha existido entre nosotros, por que siempre me han ayudado a mi como a mis hermanos tanto en momentos buenos como malos, incluso como mis propios padres, espero que esa convivencia nunca se termine, aunque haya momentos en que parezca que sea así, pero esto no ocurre al contrario se fortalece mas.

A MIS PRIMAS Y PRIMOS

Esperando que esto les sirva como un aliciente para que sigan adelante y sepan que pueden lograr todo lo que se propongan, simplemente basta un pequeño esfuerzo.

A quienes desafortunadamente ya no están presentes con nosotros, y que nunca dejaremos de extrañar, pero se que desde donde se encuentran me estuvieron ayudando en todo momento:

MIGUEL BARRAGAN

FABIOLA SOLIS

SAMANTHA DE LA SANCHA

A EL ING. MARTIN ORTIZ LEON

Mi asesor de tesis, por que gracias a su apoyo, enseñanzas y buenos consejos pude realizar mi trabajo de tesis y puedo desarrollarme dentro de lo que es la ingeniería civil.

A LOS INGENIEROS:

JOSE DOMINGUEZ
JESUS GRANADOS
LUIS ESPINOZA

Gracias por haberme recibido con las puertas abiertas en la CNA, y por haberme apoyado durante la realización de mi servicio social y tesis. Espero que nunca se pierda la amistad que surgió cuando estaba con ustedes.

CONTENIDO

ANTECEDENTES.....	1
INTRODUCCION	4
CAPITULO I. CONCEPTOS BASICOS DE PROBABILIDAD Y ESTADISTICA	
1.1 ESTADISTICA.....	6
1.2 PROBABILIDAD.....	13
1.3 PERIODO DE RETORNO.....	17
1.4 FUNCIONES DE DISTRIBUCION DE PROBABILIDAD MAS EMPLEADAS EN HIDROLOGIA.....	23
CAPITULO II. MODELOS PROBABILISTICOS	
2.1 METODOS DE ESTIMACION DE PARÁMETROS.....	24
2.2 DISTRIBUCION DE FRECUENCIAS PARA UNA SOLA POBLACIÓN.....	25
2.2.1 FUNCION DE DENSIDAD NORMAL O GAUSS.....	26
2.2.2 DISTRIBUCION GUMBEL.....	29
2.2.3 DISTRIBUCION LOGNORMAL DE 2 PARAMETROS.....	30
2.2.4 DISTRIBUCION LOGNORMAL DE 3 PARAMETROS.....	31
2.2.5 DISTRIBUCION GAMMA DE 2 PARAMETROS.....	32
2.2.6 DISTRIBUCION PEARSON 3 O GAMMA DE 3 PARAMETROS.....	33
2.2.7 DISTRIBUCION LOG PEARSON TIPO III.....	33
2.2.8 GENERAL DE VALORES EXTREMOS.....	34
2.3 MINIMO ERROR CUADRÁTICO.....	37
CAPITULO III. CASOS ESPECIALES EN HIDROLOGIA	
3.1 CASOS ESPECIALES EN HIDROLOGIA.....	38
3.2 DETERMINACION DE CURVAS DE FRECUENCIA.....	42
3.2.1 DISTRIBUCION.....	42
3.2.2 APLICACIÓN DE LA DISTRIBUCION.....	42
3.2.3 ESTIMACION DE INCLINACION GENERALIZADA.....	43
3.2.4 COEFICIENTE OBLICUO.....	44
3.2.5 REGISTRO ROTO.....	45
3.2.6 REGISTRO INCOMPLETO.....	46
3.2.7 AÑOS CON AVENIDAS DE MAGNITUD CERO.....	47
3.2.8 POBLACIONES MIXTAS.....	49
3.2.9 OUTLIERS (GASTOS EXTRAÑOS).....	52
3.2.10 LIMITES DE CONFIANZA.....	53
3.2.11 DATOS HISTORICOS.....	56
CAPITULO IV. APLICACIONES	
4.1 APLICACIÓN DE LA DISTRIBUCION LOG PEARSON III.....	60
4.2 AJUSTE PARA OUTLIERS ALTOS.....	68
4.3 PRUEBA Y AJUSTE PARA OUTLIERS BAJOS.....	75
4.4 AJUSTE PARA EVENTOS CON DATOS DE DILUVIO DE MAGNITUD CERO.....	81
4.5 APLICACIÓN DE LA DISTRIBUCION DOBLE GUMBEL.....	86
CAPITULO V. CONCLUSIONES	
5.1 CONCLUSIONES.....	89

APENDICES

1	VALORES DE K PARA COEFICIENTES OBLICUOS.....	91
2	ERROR MEDIO CUADRADO.....	93
3	VALORES DE K PARA OUTLIERS.....	94
4	LIMITES DE CONFIANZA.....	95
5	DIAGRAMAS DE FLUJO.....	99
6	DESCRIPCION DEL ALGORITMO DE ROSEN BROCK Y DEL PROGRAMA GUMP2P.....	101

BIBLIOGRAFIA

ANTECEDENTES



ANTECEDENTES

De acuerdo con el U.S. Council for Science and Technology "Hidrología" es la ciencia que trata de las aguas de la tierra, su ocurrencia, circulación y distribución, sus propiedades químicas y físicas y su reacción con el medio ambiente, incluyendo su relación con los seres vivientes. El dominio de la hidrología abarca la historia de la existencia total del agua sobre la tierra.

De acuerdo con los métodos matemáticos, la hidrología paramétrica se define como "La aproximación a la hidrología en donde el ciclo hidrológico es trazado como un sistema determinado".

Dentro de la hidrología paramétrica el término "Simulación" se define como "El desarrollo y aplicación de modelos matemáticos para representar la iteración de la variación del tiempo en los procesos físicos"

APLICACIONES Y LIMITACIONES

Para plantear las aplicaciones de la hidrología, se debe partir de la premisa de que el hombre no puede existir sin el agua. Por lo anterior, en la época actual el conocimiento de la hidrología y el disponer de gente preparada en este campo es de importancia fundamental para un efectivo manejo del agua, trascendental para nuestra sociedad.

Aunque el uso del agua para actividades domesticas es vital, su uso en la industria, comercio, agricultura y recreación es básico dentro del desarrollo del país.

Dado que el agua es un liquido vital y escaso, es lógico pensar que debemos contar con técnicas adecuadas para la planeación, manejo y desarrollo de los aprovechamientos hidrológicos del país. Una planeación y manejo efectivo de los recursos del agua es el compromiso del hidrólogo y solo es posible alcanzarlo si se entiende claramente los sistemas físicos que deben manejar. Por otra parte, se requieren modelos matemáticos confiables para evaluar y predecir el funcionamiento de los sistemas hidrológicos.

Desde el punto de vista de diseño de una obra hidráulica, los principales objetivos de la hidrología pueden reunirse en dos grandes grupos:

- a) Obtención de la avenida máxima que con una determinada frecuencia puede ocurrir en cierto lugar, lo cual es necesario considerar al diseñar vertedores, puentes y drenajes en general.
- b) Conocimiento de la cantidad, frecuencia y naturaleza de ocurrencia del transporte del agua sobre la superficie terrestre. Esto se requiere en el diseño de sistemas de irrigación, abastecimiento de agua, aprovechamientos hidroeléctricos y navegación de ríos.

Por lo que respecta a las limitaciones de la hidrología, estas se relacionan con la información hidrológica disponible y la complejidad de los procesos naturales con los cuales trata.

Lo anterior involucra a que en general cada problema es único y es difícil analizarlos con un sistema deductivo riguroso. Esto ha involucrado que al tratar de uniformarse los criterios, se plantean cada vez sistemas más complejos en donde interviene un mayor número de parámetros, que en ocasiones son difíciles de cuantificar.

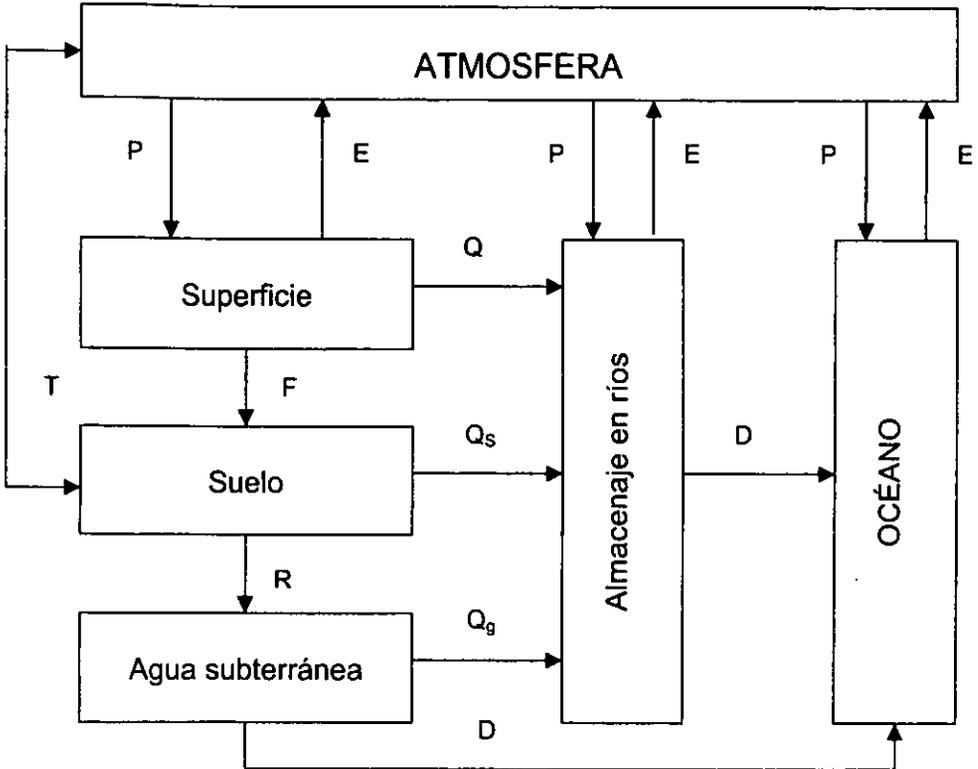
Al aplicar un criterio hidrológico para el análisis de un problema, debe considerarse que la aplicación del mismo no involucra el resultado esperado, a menos que las bases con las que se elaboró dicho criterio sean las mismas a las del problema. En caso contrario deberá tomarse en cuenta este hecho en la interpretación del resultado y darle a este el peso que merece. Conforme se disponga de mayor información hidrológica se podrán ir ajustando los criterios existentes, elaborar otros y tener una mayor seguridad en los resultados hidrológicos que se obtengan.

EL CICLO HIDROLOGICO

El ciclo hidrológico es un proceso continuo por medio del cual el agua es transportada de los océanos a la atmósfera, a la tierra y regresa al mar. El ciclo hidrológico es un término descriptivo aplicable a la circulación general del agua. Este ciclo puede empezar con la evaporación de los océanos. El vapor resultante es transportado por las masas de aire en movimiento. En determinadas condiciones, el vapor se condensa formando nubes que a su vez, pueden ocasionar precipitaciones. De la precipitación sobre el terreno, una parte es retenida por la superficie, otra escurre sobre ella y la restante penetra en el suelo. El agua retenida es devuelta a la atmósfera por evaporación y por la transpiración de las plantas. La parte que escurre sobre la superficie es drenada por arroyos y ríos hasta el océano; aunque parte se pierde por evaporación. El agua se infiltra satisface la humedad del suelo y abastece los depósitos subterráneos, de donde puede fluir hacia las corrientes de los ríos, o bien descarga en los océanos; la que queda detenida en la capa vegetal del suelo es regresada a la atmósfera por transpiración.

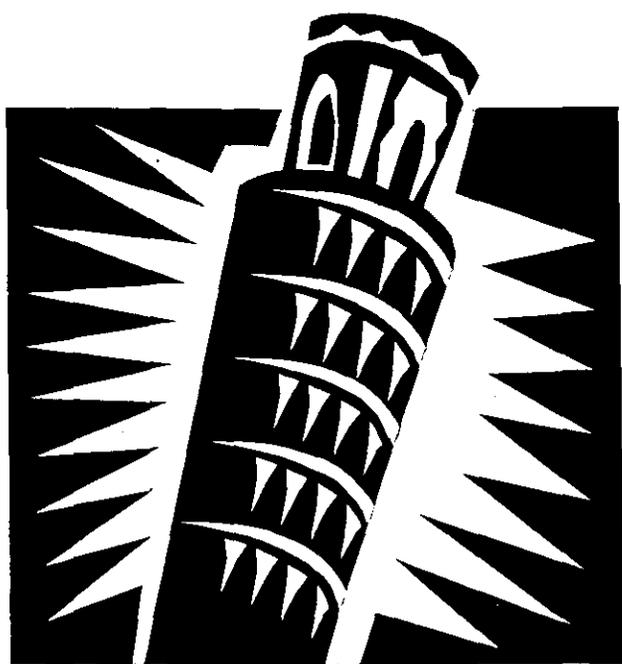
Esta descripción simplificada del ciclo hidrológico es de tipo cualitativo y en ella no se ha incluido el tiempo. Por ejemplo, después de ocurrida una tormenta, el efecto inmediato en un río se deja sentir por el escurrimiento superficial, además de existir recarga del agua subterránea. Puede decirse también que no hay evaporación durante la tormenta, y que toda el agua de lluvia se intercepta, infiltra y escurre superficialmente. Dentro de este ciclo existen diversos subciclos. Un ejemplo de esto es la evaporación del agua de la superficie terrestre y su subsecuente precipitación sobre ella antes de regresar al océano. La fuerza que controla la totalidad del sistema del transporte del agua proviene del sol, el cual provee la energía requerida para la evaporación. Cabe hacer notar que las cualidades del agua también cambian durante su paso a través del ciclo; el agua de mar se convierte en agua fresca por la evaporación.

Ciclo hidrológico
Representación cualitativa



E.- Evaporación	P.- Precipitación	T.- Transpiración
Q.- Escurrimiento superficial	Q _s - Escurrimiento Superficial	Q _g - Escurrimiento subterráneo
F.- Infiltración	D.- Descarga a los océanos	R.- recarga

INTRODUCCION



INTRODUCCION

Para predecir el gasto pico de una avenida asociada a cierto periodo de retorno o frecuencia de ocurrencia se emplean métodos estadísticos. Para aplicar estos métodos se requiere contar con registros de gastos máximos anuales o precipitación máxima en 24 horas, lográndose mayor precisión a medida que se cuente con mas años de registro.

Los métodos para evaluar el gasto máximo a partir de un periodo de retorno consideran que el gasto máximo anual es una variable aleatoria que se distribuye de acuerdo a una función de probabilidad. Por lo general en nuestro país, se cuenta con pocos años de registro por lo que las curvas de distribución de probabilidad se tienen que prolongar en sus extremos a fin de poder inferir eventos mayores o menores a los registrados.

Las características o parámetros estadísticos de las funciones de distribución de probabilidad se pueden describir a partir de los momentos de las funciones de densidad de probabilidad(fdp).

Los principales parámetros son la media (agrupamiento de la variable aleatoria de fdp o de los datos medidos respecto a un valor central); la desviación estándar (dispersión de la variable aleatoria de la fdp o de los datos observados con respecto a la media) y el coeficiente de asimetría (grado de asimetría de la fdp o hacia donde esta localizada la mayoría de los datos medidos, ya sea para los valores máximos o mínimos de estos). Por convención las letras griegas son usadas para representar los parámetros de la fdp de la población.

Casos especiales en hidrología

En ocasiones en algunos textos de hidrología aparece la ecuación siguiente:

$$X = \bar{X} + KS$$

que según Chow (1951), se llama ecuación para el análisis hidrológico de frecuencias ya sea de gastos o lluvias anuales: \bar{X} y S son la media y la desviación estándar de los datos de la muestra, K se llama factor de frecuencia, y cuyo valor depende del tipo de función de distribución de probabilidad que se use.

En la practica la representación gráfica de los datos medidos con las diferentes funciones de ajustes se pueden hacer mediante cambios de escala, es decir para cada caso usando un papel especial llamado papel de probabilidad. Este tiene en las ordenadas el valor de la variable aleatoria y en el otro la probabilidad correspondiente a cada valor.

Para poder extrapolar las estimaciones probabilísticas acerca de la variable en cuestión, se realizan ajustes de las funciones de probabilidad. Así por ejemplo, si se requiere estimar la probabilidad de que en una sección de un río se presente un gasto mayor que 500 m³/s, es necesario "ajustar" a los gastos máximos anuales observados una función de distribución que represente a la población de estos gastos. El proceso de ajuste tiene dos fases: la identificación del tipo de función de distribución adecuada y el cálculo de los parámetros que la determinan.

- Identificación de la función de distribución de probabilidad.

Para identificar el tipo de función adecuada es recomendable seguir algunos de estos procedimientos:

1.- Histograma de las observaciones: se construye el histograma de frecuencias para observar la magnitud de las posibles asimetrías y la forma general del histograma. Se comparan con las gráficas de las funciones de densidad de probabilidad y se escoge la que mejor se ajuste a los datos.

2.- Papeles de probabilidad: se dibujan los datos contra su probabilidad de no excedencia como puntos en cada uno de los papeles de probabilidad de uso común. En aquel en el cual los puntos quedan prácticamente formando una línea recta corresponde el ajuste a la mejor función de distribución.

3.- Suma de los errores cuadráticos: se calcula la suma de los errores cuadráticos de cada función de distribución y luego se escoge aquella para la cual la suma de estos errores es la más pequeña. Es decir se calcula:

$$S = \sum \varepsilon_i^2 = \sum (x_i - \hat{x}_i)^2$$

donde: x_i es el dato medido i -ésimo y \hat{x} es el valor estimado con la función de distribución para el i -ésimo dato

- Estimación de los parámetros de las distribuciones

Para estimar el valor de los parámetros que mejor se ajusten a los datos de la muestra, existen varios criterios uno de los más simples es el de los momentos. Este consiste en igualar los valores de las características estadísticas de la muestra con las de la población; esto es; hacer que la media de los valores muestrados sea igual a la de la función de distribución, y que las variancias, y el coeficiente de asimetría (relacionado con el tercer momento) de la muestra y la población sean iguales. El número de ecuaciones basadas en las características estadísticas es igual al número de parámetros de la función.

CAPITULO I



CAPITULO UNO

CONCEPTOS BASICOS DE PROBABILIDAD Y ESTADISTICA

El diseño y planeación de las obras hidráulicas, implica eventos hidrológicos a futuro, por lo que se debe de optimar su costo y por otro lado se debe de asociar el costo de las perdidas en caso de falla.

La complejidad de los procesos físicos hace imposible una estimación segura de los eventos de diseño con base en las leyes de la mecánica o la física, debido a que requeriría la creación de un modelo matemático grande, complejo y muy difícil de manejar, por lo que en nuestro caso como en otras ciencias, el camino de la probabilidad y estadística juega un papel importante en el análisis hidrológico.

1.1 Estadística

Es la rama de las matemáticas que señala las reglas para recabar, presentar y procesar los datos o eventos obtenidos al repetir varias veces un experimento, definido este, como en el proceso de observación y registro de ciertos fenómenos; es una herramienta para la toma de decisiones cuando existe incertidumbre. Los datos procedentes de un experimento, son determinísticos o aleatorios, los primeros pueden ser previstos y los segundos no, la estadística se encarga de estudiar a los aleatorios.

Al total de los datos o eventos posibles de obtener, al llevar a cabo una secuencia de experimentos, se llama población: la cual puede ser finita o infinita, dependiendo del número de datos, si son limitados o ilimitados, real si sus datos son concretos e hipotética si se consideran todas sus formas imaginables en que pueda ocurrir un dato, en la hidrología la mayoría de las veces se trabaja con poblaciones hipotéticas e infinitas.

Un grupo de datos tomados o extraídos de una población, se le denomina como muestra, cuyo número de datos n se conoce como: tamaño de la muestra.

Métodos estadísticos

La materia prima de la estadística consiste en conjuntos de números obtenidos al contar o medir cosas. Al recopilar datos estadísticos se ha de tener especial cuidado para garantizar que la información sea completa y correcta.

El primer problema para los estadísticos reside en determinar qué información y cuánta se ha de reunir.

En realidad, la dificultad al compilar un censo está en obtener el número de habitantes de forma completa y exacta; de la misma manera que un físico que quiere contar el número de colisiones por segundo entre las moléculas de un gas debe empezar determinando con precisión la naturaleza de los objetos a contar. Los estadísticos se enfrentan a un complejo problema cuando, por ejemplo, toman una muestra para un sondeo de opinión o una encuesta electoral. El seleccionar una muestra capaz de representar con exactitud las preferencias del total de la población no es tarea fácil.

Para establecer una ley física, biológica o social, el estadístico debe comenzar con un conjunto de datos y modificarlo basándose en la experiencia. Por ejemplo, en los primeros estudios sobre crecimiento de la población los cambios en el número de habitantes se predecían calculando la diferencia entre el número de nacimientos y el de fallecimientos en un determinado lapso. Los expertos en estudios de población comprobaron que la tasa de crecimiento depende sólo del número de nacimientos, sin que el número de defunciones tenga importancia. Por tanto, el futuro crecimiento de la población se empezó a calcular basándose en el número anual de nacimientos por cada 1.000 habitantes. Sin embargo, pronto se dieron cuenta que las predicciones obtenidas utilizando este método no daban resultados correctos. Los estadísticos comprobaron que hay otros factores que limitan el crecimiento de la población. Dado que el número de posibles nacimientos depende del número de mujeres, y no del total de la población, y dado que las mujeres sólo tienen hijos durante parte de su vida, el dato más importante que se ha de utilizar para predecir la población es el número de niños nacidos vivos por cada 1.000 mujeres en edad de procrear. El valor obtenido utilizando este dato mejora al combinarlo con el dato del porcentaje de mujeres sin descendencia. Por tanto, la diferencia entre nacimientos y fallecimientos sólo es útil para indicar el crecimiento de población en un determinado periodo de tiempo del *pasado*, el número de nacimientos por cada 1.000 habitantes sólo expresa la tasa de crecimiento en el mismo periodo, y sólo el número de nacimientos por cada 1.000 mujeres en edad de procrear sirve para predecir el número de habitantes en el *futuro*.

Tabulación y presentación de los datos

Los datos recogidos deben ser organizados, tabulados y presentados para que su análisis e interpretación sean rápidos y útiles. Por ejemplo, para estudiar e interpretar la distribución de las notas o calificaciones de un examen en una clase con 30 alumnos, primero se ordenan las notas en orden creciente: 3,0; 3,5; 4,3; 5,2; 6,1; 6,5; 6,5; 6,5; 6,8; 7,0; 7,2; 7,2; 7,3; 7,5; 7,5; 7,6; 7,7; 7,8; 7,8; 8,0; 8,3; 8,5; 8,8; 8,8; 9,0; 9,1; 9,6; 9,7; 10 y 10. Esta secuencia muestra, a primera vista, que la máxima nota es un 10, y la mínima es un 3; el rango, diferencia entre la máxima y la mínima es 7.

En un diagrama de frecuencia acumulada, como el de la figura 1, las notas aparecen en el eje horizontal y el número de alumnos en el eje vertical izquierdo, con el correspondiente porcentaje a la derecha. Cada punto representa el número total de estudiantes que han obtenido una calificación menor o igual que el valor dado.

Por ejemplo, el punto A corresponde a 7,2, y según el eje vertical, hay 12 alumnos, o un 40%, con calificaciones menores o iguales que 7,2.

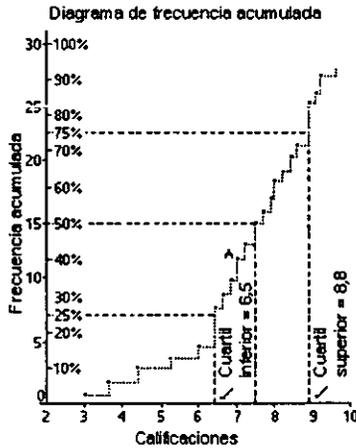


Figura 1

Para analizar las calificaciones obtenidas por 10 clases de 30 alumnos cada una en cuatro exámenes distintos (un total de 1.200 calificaciones), hay que tener en cuenta que la cantidad de datos es demasiado grande para representarlos como en la figura 1. El estadístico tiene que separar los datos en grupos elegidos previamente denominados intervalos. Por ejemplo, se pueden utilizar 10 intervalos para tabular las 1.200 calificaciones, que se muestran en la columna (a) de la tabla de distribución de datos adjunta; el número de calificaciones por cada intervalo, llamado frecuencia del intervalo, se muestra en la columna (c). Los números que definen el rango de un intervalo se denominan límites. Es conveniente elegir los límites de manera que los rangos de todos los intervalos sean iguales y que los puntos medios sean números sencillos. Una calificación de 8,7 se cuenta en el intervalo entre 8 y 9; una calificación igual a un límite de intervalo, como 9, se puede asignar a cualquiera de los dos intervalos, aunque se debe hacer de la misma manera a lo largo de toda la muestra. La frecuencia relativa, columna (d), es la proporción entre la frecuencia de un intervalo y el número total de datos. La frecuencia acumulada, columna (e), es el número de estudiantes con calificaciones iguales o menores que el rango de cada intervalo sucesivo. Así, el número de estudiantes con calificaciones menores o iguales a 3 se calcula sumando las frecuencias de la columna (c) de los tres primeros intervalos, dando 53. La frecuencia acumulada relativa, columna (f), es el cociente entre la frecuencia acumulada y el número total de notas.

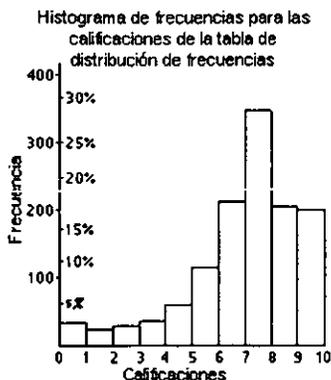


Figura 2

Los datos de una tabla de distribución de frecuencias se pueden representar gráficamente utilizando un histograma o diagrama de barras (como en la figura 2), o como un polígono de frecuencias acumuladas (como en la figura 3). El histograma es una serie de rectángulos con bases iguales al rango de los intervalos y con área proporcional a sus frecuencias. El polígono de la figura 3 se obtiene conectando los puntos medios de cada intervalo de un histograma de frecuencias acumuladas con segmentos rectilíneos.

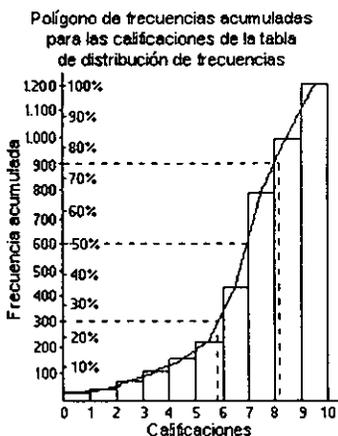


Figura 3

En los periódicos y otros medios de comunicación los datos se representan gráficamente utilizando símbolos de diferente longitud o tamaño que representan las distintas frecuencias.

Valores de la tendencia central

Una vez que los datos han sido reunidos y tabulados, comienza el análisis con el objeto de calcular un número único, que represente o resuma todos los datos. Dado que por lo general la frecuencia de los intervalos centrales es mayor que el resto, este número se suele denominar valor o medida de la *tendencia central*.

Sean x_1, x_2, \dots, x_n los datos de un estudio estadístico. El valor utilizado más a menudo es la media aritmética o promedio aritmético que se escribe \bar{X} , y que es igual a la suma de todos los valores dividida por n :

$$\bar{X} = \frac{\sum X_i}{n}$$

El símbolo Σ , o *sumatorio*, denota la suma de todos los datos. Si las x se agrupan en k intervalos, con puntos medios m_1, m_2, \dots, m_k y frecuencias f_1, f_2, \dots, f_k , la media aritmética viene dada por

$$\frac{\sum f_i m_i}{\sum f_i}$$

donde:

$$i = 1, 2, \dots, k.$$

La mediana y la moda son otros dos valores de la tendencia central. Si las x se ordenan según sus valores numéricos, si n es impar la mediana es la x que ocupa la posición central y si n es par la mediana es la media o promedio de las dos x centrales. La moda es la x que aparece con mayor frecuencia. Si dos o más x aparecen con igual máxima frecuencia, se dice que el conjunto de las x no tiene moda, o es *bimodal*, siendo la moda las dos x que aparecen con más frecuencia, o es *trimodal*, con modas las tres x más frecuentes.

Medidas de la dispersión

Normalmente la estadística también se ocupa de la dispersión de la distribución, es decir, si los datos aparecen sobre todo alrededor de la media o si están distribuidos por todo el rango. Una medida de la dispersión es la diferencia entre dos percentiles, por lo general entre el 25 y el 75. El percentil p es un número tal que un p por ciento de los datos son menores o iguales que p . En particular, los percentiles 25 y 75 se denominan cuartiles inferior y superior respectivamente. La desviación típica es otra medida de la dispersión, pero más útil que los percentiles, pues está definida en términos aritméticos como se explica a continuación.

La desviación de un elemento del conjunto es su diferencia con respecto a la media; por ejemplo, en la sucesión x_1, x_2, \dots, x_n la desviación de x_1 es $x_1 - \bar{x}$, y el cuadrado de la desviación es $(x_1 - \bar{x})^2$. La *varianza* es la media del cuadrado de las desviaciones.

Por último, la desviación típica, representada por la letra griega sigma (σ), es la raíz cuadrada de la varianza, y se calcula de la siguiente manera:

$$\sigma = \sqrt{\frac{1}{n} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2]} = \frac{1}{n} \sqrt{\sum (x_i - \bar{x})^2}$$

Si la desviación típica es pequeña, los datos están agrupados cerca de la media; si es grande, están muy dispersos.

Correlación

Cuando dos fenómenos sociales, físicos o biológicos crecen o decrecen de forma simultánea y proporcional debido a factores externos, se dice que los dos fenómenos están positivamente correlados. Si uno crece en la misma proporción que el otro decrece, los dos fenómenos están negativamente correlados. El grado de correlación se calcula aplicando un coeficiente de correlación a los datos de ambos fenómenos. El coeficiente de correlación más utilizado es:

$$\frac{\sum \left(\frac{x}{\sigma^x} \cdot \frac{y}{\sigma^y} \right)}{N}$$

donde:

x es la desviación de una variable con respecto a su media, y es la desviación de la otra variable con su media; N es el número total de casos en las series. Una correlación positiva perfecta tiene un coeficiente +1, y para una correlación negativa perfecta es -1. La ausencia de correlación da como coeficiente 0. Por ejemplo, el coeficiente 0,89 indica una correlación positiva grande, -0,76 es una correlación negativa grande y 0,13 es una correlación positiva pequeña.

Modelos matemáticos

Un modelo matemático es una representación ideal (en la forma de un sistema, proposición, fórmula o ecuación) de un fenómeno físico, biológico o social.

Así, un dado teórico perfectamente equilibrado, que se puede lanzar de forma aleatoria, es un modelo matemático de un dado real. La probabilidad de que en n lanzamientos de un dado matemático se obtenga k veces un 6 es

$$P(k) = \binom{n}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}$$

donde:

$\binom{n}{k}$ es la representación de un número combinatorio

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{1 \cdot 2 \cdot \dots \cdot k}$$

$$\left(\binom{n}{0} = 1\right)$$

El estadístico que utiliza un dado real debe diseñar un experimento, como lanzar el dado un gran número de veces, para determinar, a partir de los resultados obtenidos, la posibilidad de que el dado esté perfectamente equilibrado y de que el lanzamiento sea aleatorio.

Muchos conjuntos de medidas experimentales presentan el mismo tipo de distribución de frecuencias que se pueden representar con un modelo matemático único. Por ejemplo, el número de veces que sale un 6 al lanzar un dado n veces, el peso de N garbanzos tomados al azar de una bolsa o las presiones atmosféricas medidas por distintos estudiantes sucesivamente en el mismo barómetro. En todos los casos los valores presentan patrones de frecuencia muy similares. Los estadísticos adoptan un modelo que es un prototipo o idealización matemática de todos esos patrones o distribuciones. Una forma de modelo matemático puede ser una ecuación de la distribución de frecuencias, en la que el número de medidas o valores se considera infinito:

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

donde: e es la base de los logaritmos neperianos, e y representa la frecuencia del valor x . La gráfica de esta fórmula (figura 4) es una curva en forma de campana llamada curva de probabilidad normal o *gaussiana*:

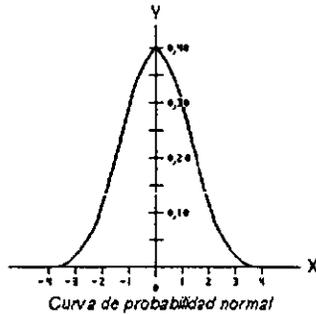


Figura 4

1.2. Probabilidad

La probabilidad de un resultado se representa con un número entre 0 y 1, ambos inclusive. La probabilidad 0 indica que el resultado no ocurrirá nunca, y la probabilidad 1 que el resultado ocurrirá siempre. Los problemas más sencillos estudian la probabilidad de un suceso favorable en un experimento o acontecimiento con un número finito de resultados, todos ellos con igual probabilidad de ocurrir. Si un experimento tiene n posibles resultados, y f de ellos se consideran favorables, la probabilidad de un suceso favorable es f/n . Por ejemplo, un dado no trucado se puede lanzar de seis formas posibles, por tanto, la probabilidad de que salga un 5 o un 6 es $2/6$.

Problemas más complicados estudian acontecimientos en que los distintos resultados tienen distintas probabilidades de ocurrir. Por ejemplo, encontrar la probabilidad de que salga 5 o 6 al lanzar un par de dados: los distintos resultados (2, 3, ..., 12) tienen distintas probabilidades. Algunos experimentos pueden incluso tener un número infinito de posibles resultados, como la probabilidad de que una cuerda de circunferencia dibujada aleatoriamente sea de longitud mayor que el radio.

Los problemas que estudian experimentos repetitivos relacionan la probabilidad y la estadística. Algunos ejemplos: encontrar la probabilidad de obtener 5 veces un 3 y al menos 4 veces un 6 al lanzar un dado sin hacer trampas 50 veces; si una persona lanza una moneda al aire y da un paso hacia delante si sale cara y un paso hacia atrás si sale cruz, calcular la probabilidad de que, después de 50 pasos, la persona esté a menos de 10 pasos del origen.

En términos probabilísticos, dos sucesos de un experimento son mutuamente excluyentes si la probabilidad de que los dos ocurran al mismo tiempo es cero; dos sucesos son independientes si la probabilidad de que ocurran al mismo tiempo es igual al producto de sus probabilidades individuales. Es decir, dos sucesos son excluyentes si la ocurrencia de uno prohíbe la ocurrencia del otro; dos sucesos son independientes si la ocurrencia o no de uno no afecta a la probabilidad de que el otro ocurra o no.

Probabilidad compuesta es la probabilidad de que todos los casos de un conjunto dado de sucesos ocurran a la vez; probabilidad total es la de que al menos uno de los casos de un conjunto dado de sucesos ocurra. Probabilidad condicional es la probabilidad de que un suceso ocurra cuando se sabe que otro suceso ha ocurrido o va a ocurrir.

Si la probabilidad de que un suceso ocurra es p , la probabilidad de que no ocurra es $q = 1 - p$. Por tanto, la confianza en que el suceso ocurra es p contra q y la de que no ocurra es q contra p . Si las probabilidades de dos sucesos mutuamente excluyentes X e Y son p y P respectivamente, la confianza en que X ocurra y que Y no ocurra es p contra P . Si un experimento debe dar como resultado uno de los sucesos O_1, O_2, \dots, O_n , mutuamente excluyentes, cuyas probabilidades son p_1, p_2, \dots, p_n , respectivamente, y si a cada uno de los posibles resultados se le asigna un valor numérico v_1, v_2, \dots, v_n , el resultado esperado del experimento es $E = p_1v_1 + p_2v_2 + \dots + p_nv_n$. Por ejemplo, una persona lanza un dado, ganando 4 pasteles si saca 1, 2 o 3 y 3 pasteles si saca 4 o 5; pierde 12 pasteles si saca un 6. El resultado esperado con un solo lanzamiento es $3/6 \times 4 + 2/6 \times 3 - 1/6 \times 12 = 1$, o lo que es lo mismo, un pastel.

El uso más generalizado de la probabilidad es su utilización en el análisis estadístico. Por ejemplo, la probabilidad de sacar 7 al lanzar dos dados es $1/6$, lo que significa (se interpreta como) que al lanzar dos dados aleatoriamente, un gran número de veces, alrededor de un sexto de los lanzamientos darán 7. Este concepto se utiliza a menudo para calcular estadísticamente la probabilidad de un suceso que no se puede medir o es imposible de obtener.

Así, si la estadística a largo plazo muestra que por cada 100 personas entre 20 y 30 años sólo habrá 42 vivos cuando tengan 70, lo que quiere decir que la probabilidad de que una de esas personas llegue a los 70 años es de un 42 por ciento.

Las leyes naturales más simples son aquellas que expresan las condiciones bajo las cuales un evento de interés ocurre o no ocurre con certeza, y se puede expresar como:

- a) Si se presenta un conjunto de condiciones S , entonces el evento A , seguramente ocurrirá; o bien
- b) Si se presenta un conjunto de condiciones S , entonces el evento A no puede ocurrir

Cuando un evento A , en presencia de un conjunto de condiciones S , a veces ocurre y a veces no, se llama aleatorio con respecto al conjunto S . Cuando esto sucede, no se han tomado en cuenta en el conjunto S todas las condiciones necesarias para la ocurrencia o no ocurrencia del evento, y no, que exista una ley física que conduzca a esta ocurrencia o no ocurrencia.

Estas condiciones o leyes no se incluyen en el evento A se suplen por una ley de probabilidades, la cual indica con que frecuencia se presenta el evento dadas las condiciones S .

Tenemos un ejemplo a forar un río, (medir el escurrimiento que pasa en su sección), nos damos cuenta de que el volumen mensual escurrido es a veces mayor que 300 000 m³, pero a veces es menor a esta cantidad, si quisiéramos aprovechar el agua de ese río para regar, pero sin hacer una presa, obviamente el dato que se nos proporciona es demasiado vago, y no nos serviría. Podríamos buscar un dato mínimo y un máximo durante los n meses que se han hecho las observaciones y proporcionar un dato de este tipo:

$$20\ 000 \leq V \leq 1\ 000\ 000 \quad (1.1)$$

donde V es el volumen escurrido en m³

Este dato es de mayor utilidad pero aun es muy deficiente, un proyectista recibirá una información más valiosa si se le indica que aproximadamente el 70% de los meses el volumen es mayor de 300 000 m³.

Axiomas de la teoría de probabilidad

Primer axioma de probabilidad

$$0 \leq P(A) \leq 1 \quad (1.2)$$

donde:

A = es un evento

Establece que las probabilidades son números reales que varían entre cero y uno.

Segundo axioma de la probabilidad

$$P(E) = 1 \quad (1.3)$$

donde: E es el espacio muestral

El espacio muestral completo tiene una probabilidad de 1, y esto, expresa la idea de que la probabilidad de un evento seguro, o sea, un evento que deba suceder es igual a uno.

Se dice que dos eventos son mutuamente excluyentes cuando no pueden ocurrir simultáneamente, por ejemplo:

$$A: V > 500\ 000\ m^3$$

Y el evento

$$B: V < 10\,000\text{ m}^3$$

Este par de eventos son mutuamente excluyentes, ya que puede suceder uno u otro, pero no los dos al mismo tiempo. Si la frecuencia del evento A es v_1 y la del evento B es v_2 , entonces la frecuencia de la unión de los dos es $v = v_1 + v_2$; por lo tanto la probabilidad de $C = A \cup B$ es:

Tercer axioma de la probabilidad para eventos mutuamente excluyentes:

$$P(c) = P(A \cup B) = P(A) + P(B) \quad (1.4)$$

Este axioma establece que las funciones de probabilidad deben ser aditivas.

Distribuciones de probabilidad

Las variables aleatorias, ya sean discretas o continuas, se caracterizan por la distribución de probabilidades junto con el valor específico que la variable puede asumir. Una variable aleatoria a través de su rango de frecuencia es generalmente designada con una letra mayúscula, y el valor específico que toma la variable se denota con las letras minúsculas. Por ejemplo, $P(X = x_1)$ es la probabilidad que la variable aleatoria X se tome el valor de x_1 .

Funciones discretas de probabilidad

Cuando el número de valores x que puede tomar una variable aleatoria X es finito, se dice que la variable aleatoria X es discreta.

Hay que hacer la observación que cada uno de los siete eventos tiene una probabilidad finita y la suma es igual a 1, esto es: $\sum P(x_i) = 1$

$P(0) = 0.05$	$P(4) = 0.15$
$P(1) = 0.15$	$P(5) = 0.10$
$P(2) = 0.25$	$P(6) = 0.08$
$P(3) = 0.20$	$P(7) = 0.02$

Funciones continuas de probabilidad

Se asigna este nombre cuando el número n de valores que puede tomar una variable aleatoria X es infinito, un ejemplo puede ser los volúmenes de escurrimiento mensual de un río.

1.3. Periodo de retorno (Tr)

Es el objetivo primario del análisis estadístico de datos hidrológicos, se define como el lapso promedio (numero promedio de años), en que un evento puede ser igualado o excedido. Por ejemplo, si se diseña una obra de gasto de $700 \text{ m}^3/\text{seg}$, asignado a un periodo de retorno de 10 años, significa que el tiempo que transcurre para que vuelva a presentarse un gasto mayor o igual a el puede ser de 8, 10, 13, o 7 años; hay que observar que el promedio de ellos es precisamente 10 años.

Por otra parte la probabilidad de que un evento con un periodo de retorno Tr ocurra en cualquier año es igual a

$$q = \frac{1}{Tr} \quad (1.5)$$

En un conjunto de eventos máximos anuales, ya sean gastos o lluvias, el periodo de retorno que se asocia a cada uno de ellos puede ser estimado con la formula de Weibull

$$Tr = \frac{n+1}{m} \quad (1.6)$$

donde m es numero de orden y n es el numero total de años de registro. Esta expresión no toma en cuenta el numero de eventos del conjunto, es decir, la longitud del registro. Una formula que si lo hace es la propuesta por Gringorten, que se escribe como:

$$Tr = \frac{n+1-2a}{m-a} \quad (1.7)$$

El valor de a , esta comprendido entre 0 y 1, depende del valor de n y se encuentra en la tabla siguiente:

Tabla 1.1. Valores de a , formula de Gringorten

n	10	20	30	40	50
a	0.448	0.443	0.442	0.441	0.440
n	60	70	80	90	100
a	0.440	0.440	0.440	0.439	0.439

La ecuación (1.6) es la mas eficiente para calcular la posición grafica de los datos de cualquier función de distribución de probabilidad y es la mas comúnmente usada.

Para eventos máximos anuales los datos se ordenan de mayor a menor y el lugar que ocupan dentro de la lista, cada uno de ellos corresponde al valor de m ; así, por ejemplo, el valor mayor tiene una $m=1$, mientras que el mas pequeño $m = n$.

Cuando se tiene un conjunto de eventos mínimos los valores que lo forman se ordenan del mas pequeño hasta el mas grande, en este caso el periodo de retorno debe establecerse como el numero promedio de años en que el evento tiene valores menores o igual a el.

Para conocer el riesgo en eventos máximos anuales para una obra hidráulica se puede utilizar la ecuación siguiente:

$$r = 1 - (1-q)^L$$

donde r es la probabilidad o riesgo de que ocurra el evento máximo al menos en L años de vida útil de la obra.

Eligiendo un valor del periodo de retorno y conociendo L se puede calcular el riesgo. En la siguiente tabla se mencionan algunos valores de riesgo para diferentes condiciones de vida útil y periodo de retorno.

Tabla 1.2. Periodo de retorno asociado a varios niveles de riesgo y vida útil de la obra

Riesgo en %	Vida útil en años				
	5	30	50	70	100
	Periodo de retorno en años				
40	10.3	59.2	98.4	137.5	196.3
50	7.7	43.8	72.6	101.5	144.8
60	6.0	33.2	55.1	76.9	119.6
80	3.6	19.1	31.6	44.0	62.6
99	1.7	7.0	11.4	15.7	22.2

Algunas veces es necesario estimar la probabilidad de que un evento, con un determinado periodo de retorno, sea igualado o excedido en un cierto intervalo, vamos a suponer que para el diseño de una presa cuya vida útil sea de 100 años de periodo de retorno pueda ocurrir. Esto es el llamado concepto de riesgo y se define como la probabilidad de que ocurra un evento en cualquier año.

Ejemplo 1.1:

Veamos el siguiente ejemplo en el que determinaremos el riesgo de falla de una obra que tiene una vida útil de 10 años si se diseña para un periodo de retorno de 10 años. La ecuación correspondiente es:

$$R = 1 - \left(1 - \frac{1}{Tr}\right)^n \tag{1.8}$$

donde:

- R = es el riesgo de falla
- Tr = periodo de retorno
- n = vida útil de la obra

sustituyendo $T_r = 10$ años y $n = 10$ años

$$R = 1 - \left(1 - \frac{1}{10}\right)^{10} = 0.651$$

El riesgo es de 65.1 %, es decir, se tiene una probabilidad del 65.1 % de que la obra falle en su vida útil.

Selección del periodo de retorno

Para la selección de la avenida de diseño se requiere por una parte un análisis hidrológico que lo fundamente y con base a este hacer el análisis económico del problema. Conforme se incrementa el tamaño de la avenida de diseño, el costo de la obra aumenta y al mismo tiempo, la probabilidad de exceder la avenida disminuye.

No obstante, de que la selección final del periodo de retorno para una estructura en particular interviene el criterio del ingeniero, es recomendable seguir los siguiente:

- 1.- Importancia de la obra.
- 2.- Magnitud de la estructura, por ejemplo, la altura de los bordes, la capacidad del cauce de alivio, etc.
- 3.- Posible daño a propiedades adyacentes.
- 4.- Costo de mantenimiento.
- 5.- Costo de reparaciones de la obra.
- 6.- Amortización del costo de la estructura de la obra.
- 7.- Inconvenientes por suspensión de su operación.
- 8.- Riesgo por pérdidas de vidas humanas.

La Comisión Nacional del Agua (CNA), como normatividad, recomienda adoptar la siguiente tabla para seleccionar el periodo de retorno.

Tabla 1.3. Selección del periodo de retorno

TIPO DE OBRA HIDRAULICA	Tr (en años)
DRENAJE PLUVIAL	
Lateral libre en calles de poblados donde se tolera encharcamiento de corta duración.	2
Lateral libre en calles de poblados donde no se tolera encharcamiento lateral.	2
Zonas agrícolas.	5
Zonas urbanas:	
a) Poblados pequeños con menos de 100 000 habitantes	2 – 5
b) Poblados medianos entre 100 000 y 1 000 000 de habitantes.	5 – 10
c) Poblados grandes con mas de 1 000 000 de habitantes.	10 – 25
Aeropuertos, estaciones de ferrocarril y autobuses.	10
Cunetas y contracunetas en caminos y carreteras.	5
ESTRUCTURAS DE CRUCE	
Puentes carreteros en:	
a) Caminos locales que comunican poblados pequeños.	25 – 50
b) Caminos regionales que comunican poblados medianos.	50 – 100
c) Carreteras que comunican grandes poblados (ciudades).	500 – 1000
Puentes ferrocarrileros en:	
a) Vías locales aisladas (desvíos)	50 – 100
b) Vías secundarias regionales.	100 – 500
c) Vías primarias del país.	500 – 1000
Puentes, canales o tuberías en conducción de agua:	
a) Para riego, área menor de 1000 has.	10 – 25
b) Para riego, área de 1000 a 10 000 has.	25 – 50
c) Para riego, área mayor a 10 000 has.	50 – 100
d) Abastecimiento industrial.	100 – 500
e) Abastecimiento de agua potable.	
Puentes para tuberías de petróleo y gas.	
a) Abastecimiento secundario local.	25 – 50
b) Abastecimiento regional.	50 – 100
c) Abastecimiento primario.	100 – 500
ALCANTARILLAS PARA PASO DE CORRIENTES PEQUEÑAS	Tr (en años)
En caminos locales que comunican poblados pequeños.	10 – 25
En caminos regionales que comunican poblados medianos	25 – 50
En caminos primarios que comunican poblados grandes.	50 – 100

DELIMITACIONES DE ZONAS FEDERALES	Tr (en años)
Corrientes libres en: a) Zonas semiáridas a húmedas. b) Zonas áridas con régimen de escurrimiento errático. c) Zonas de desbordamiento. Corrientes con obras de control. Además del tramo libre debe tenerse en cuenta el gasto regulado.	5 10 o mayor, con base en la capacidad del cauce natural cavado. 5 o 10 en ambos, o el reguiado de diseño de la obra si es superior.
DELIMITACION DE ZONAS DE PROTECCION EN OBRAS HIDRAULICAS	A juicio de la CNA
ENCAUSAMIENTO DE CORRIENTES	
Corrientes libres en zona: a) Agrícola de pequeña extensión, menor a 1000 ha. b) Agrícola de extensión mediana de 1000 a 10 000 ha. c) Agrícola de extensión grande de 10 000 ha en adelante. d) Para protección a poblaciones pequeñas. e) Para protección a poblaciones medianas. f) Para protección a poblaciones grandes. Corrientes controladas: a) Existente tramo libre. b) No existe un tramo libre.	10 - 25 25 - 50 50 - 100 100 - 500 100 - 500 Tramo libre, idéntico a corrientes libres en zonas mas gasto regulado para ese periodo de retorno o gasto de diseño del control si es superior. Igual a gasto de diseño del control.
PRESAS DERIVADORAS	
Zona de riego pequeña (menor de 1000 ha)	50 - 100
Zona de riego mediana (1000 a 10 000 ha)	100 - 500
Zona de riego grande (mas de 10 000 ha)	500 - 1000
OBRAS DE DESVIO TEMPORAL	
Presas pequeñas	10 - 25
Presas medianas	25 - 50
Presas grandes	50 - 100
Causes de alivio en corriente	25 - 50 (o mayor, según la importancia)

PRESAS DE ALMACENAMIENTO	Tr (en años)
De jales (lodo del procesamiento de minerales en minas).	500 – 1000
Azolve del acarreo del suelo en cuenca.	500 – 1000
De agua para abastecimiento a poblaciones, riego, energía, etc.	500 – 10 000 (dependiendo del grado de daños materiales y vidas)

Datos para el diseño

En el diseño de cualquier estructura se requiere de información, la cual es de dos tipos , una corresponde a la fisiográfica y la otra a la hidrológica.

Tipos de datos:

Los datos hidrológicos ya sean lluvias o gastos, se presentan en orden cronológico, en ocasiones solo algunos de los valores originales tienen aplicación ya que el análisis de los mismos es regido por una condición crítica, es decir, frecuentemente se basa en usar dos tipos de datos, a unos se les llama serie de valores máximos extremos y el otro serie de excedente o de duración parcial.

La serie de valores máximos solo toma en cuenta el valor mas grande o mas pequeño que estén en el registro para un determinado intervalo constante de tiempo, así por ejemplo, si este intervalo es de una año y contiene los valores mas grandes o pequeños se le designa como una serie de máximos o mínimos anuales respectivamente.

La serie de excedentes esta formada por un conjunto de datos los cuales se seleccionan de tal forma que su magnitud sea mayor a una cierto valor de referencia, es decir, el numero de datos de la serie debe ser igual al numero de años de registro.

La serie de máximos anuales se utiliza cuando el diseño debe estar regido por las condiciones mas desfavorables y la de excedentes cuando el segundo valor mas grande en el año puede influir en el diseño.

El periodo de retorno para los máximos anuales se obtiene con la ecuación (1.6) y para los excedentes con la expresión siguiente:

$$T_e = \frac{n}{m} \tag{1.9}$$

Las ecuaciones (1.6) y (1.9) están relacionadas por:

$$T_e = \frac{1}{\log_e \frac{T_r}{T_r - 1}} \quad (1.10)$$

Del análisis de esta ecuación se tiene que para valores de T_r mayores de 10 años este es semejante a T_e , por lo que se puede decir que al contar con un registro mayor de 10 años el análisis hidrológico de los datos si se usa una serie de máximos o una de excedentes.

Los datos fisiográficos son: el área, tipo de terreno y cobertura vegetal de la cuenca, pendiente y longitud del cauce principal, topografía y geología de la zona, etc.

Los datos hidrológicos son la información de lluvia, escurrimiento, evaporación, temperatura, etc.

1.4. FUNCIONES DE DISTRIBUCIONES DE PROBABILIDAD MAS EMPLEADAS EN HIDROLOGIA

Una vez que adoptamos el periodo de retorno es necesario aplicar una distribución de probabilidad para conocer el gasto de diseño, por lo que debemos de hacer extrapolaciones a partir de los gastos o lluvias registrados, pues rara vez este periodo es menor a la longitud de registros.

Por ejemplo, puede ser necesario determinar un gasto de diseño con periodo de retorno de 1000 años a partir de 25 años de registro. Si los gastos máximos anuales (o lluvias) registrados se dibujan con sus respectivos periodos de retorno, generalmente se observa alguna tendencia mas o menos definida. El problema radica en extender esa tendencia hasta el periodo de retorno deseado. Una posibilidad es extrapolar los datos gráficamente. Este método puede dar buenos resultados si se aplica por una persona que tiene gran experiencia, pero tiene la desventaja de la subjetividad, esto es, si veinte personas diferentes lo aplican, se obtendrían veinte resultados diferentes.

Un método adecuado para elegir la lluvia de diseño es la aplicación de las diferentes funciones de distribución de probabilidad y a partir de la función que presente un mejor ajuste de los datos medidos, se podrá seleccionar esta para la extrapolación.

En la estadística existen decenas de funciones de distribución de probabilidad teóricas, obviamente no es posible probar todas para un problema particular, por lo tanto es necesario escoger de esas funciones, las que mejor se adopten al problema de análisis. De las funciones de distribución de probabilidad de uso común en hidrología, se tienen las siguientes:

Normal, Lognormal de dos parámetros, Lognormal de tres parámetros, Gamma de dos parámetros, Pearson Tipo III, Gumbel (valores extremos I), General de valores extremos, Gumbel mixta o de dos poblaciones.

CAPITULO II



CAPITULO DOS

MODELOS PROBABILISTICOS

2.1. METODOS DE ESTIMACIÓN DE PARAMETROS

En general, los problemas hidrológicos son analizados a través del uso del modelo. El modelo puede ser una conceptualización, una relación empírica, un dispositivo físico; o un conjunto de expresiones matemáticas, estadísticas y/o empíricas. La mayoría de los modelos hidrológicos pueden clasificarse como determinísticos, paramétricos, estocásticos, o una combinación de ellos. La diferencia entre estos tres tipos básicos de modelos no está del todo bien delimitada. Se puede pensar que los modelos están formados de alguna combinación de cada componente de los cuales representan un espectro continuo de modelos "tipo" variando desde completamente determinístico por un lado hasta completamente estocástico por el otro.

Un modelo completamente determinístico sería aquel que es obtenido a través de relaciones físicas y no requiere de datos experimentales en su aplicación. Un modelo paramétrico puede ser considerado como un determinístico en el sentido de que una vez que los parámetros del modelo son determinados, el modelo siempre genera la misma salida a partir de la información de entrada. Por otro lado, un modelo paramétrico es estocástico en el sentido de que los parámetros estimados dependen de los datos observados y aquellos cambiarían cuando los datos observados cambien. Un modelo estocástico es aquel en el cual las salidas son predecibles solo en un sentido estadístico. En un modelo estocástico, el empleo repetido de un grupo dado de entradas del modelo genera salidas que no son las mismas pero siguen cierto modelo estadístico.

Uno de los problemas más comunes en hidrología es la estimación de una avenida o sequía de diseño a partir de un registro de escurrimiento bastante corto. Dibujando la magnitud de los eventos medidos, digamos gastos máximos anuales, algún tipo de modelo es generalmente aparente. La pregunta es como emplear este modelo para extender los datos disponibles y estimar el evento de diseño.

Si se dispone de un gran número de eventos observados o medidos a partir de un periodo de registro al menos tan largo como periodo de retorno del evento de diseño a estimarse, entonces el problema se simplifica. En el caso extremo, si se contara con una muestra bastante grande, digamos un millón de eventos, entonces el evento de diseño y su intervalo de Confianza podría derivarse directamente de los datos de la muestra. Dado que tal cantidad de datos hidrológicos no son posibles de obtener, la muestra de datos disponible deberá utilizarse para ajustarles una distribución de frecuencias, la cual a su vez será empleada para extrapolar los eventos de diseño, ya sea gráficamente o por medio de la estimación de sus parámetros.

Como ya se ha mencionado, los parámetros son estimados para una función de distribución a partir de los datos de una muestra. Existen, por supuesto, muchas técnicas por las cuales se obtiene la estimación de parámetros. Por lo que es posible preguntarse ¿Cuál es el camino mas adecuado para su estimación?, ¿Qué significa el mejor estimador? Y ¿Estos son los únicos estimadores?.

Cada estimador de un parámetro es una función de los valores de la muestra, los cuales son observaciones de una variable aleatoria. Así, el propio parámetro estimado de una variable aleatoria que tiene su propia distribución muestral. Un estimador obtenido a partir de un grupo de valores puede ser considerado como un valor observado de una variable aleatoria. Por lo cual, la bondad de un estimador puede ser juzgado a partir de su distribución.

Existen diferentes técnicas de estimación de los parámetros de los modelos hidrológicos. Los Métodos de Momentos (Nash, 1959; Dogge, 1973) y Cumulantes (Doogge, 1973; Harley, 1967; O Meara, 1968; Van de Nes y Hendriks, 1971). Algunos otros métodos populares lo son: el Gráfico (Yevjevich, 1972); el Método de Medias Incompletas (Houghton, 1978); el Método de Momentos de Probabilidad Pesada (Greenwood et al ., 1979); el Método de Momentos Mezclados (Rao, 1980, 1983; Shrader et al ., 1981); el de Estimación de la Máxima Verosimilitud (Douglas, Clerke, y Newton, 1976; Sorooshian, Gupta, y Fulton, 1983; Phien y Jivajirajah, 1984); el Método de Mínimos Cuadrados (Jones, 1971; Snyder, 1972; Bree, 1978a, 1978b); el de Estimación de Parámetros Restringidos (Natale y Todini, 1974; Todini y Wallis, 1977); el de Estimación Bayesiana (Natale y Todini, 1974; Kuczera, 1983a, 1983b); los Métodos de Optimización (Himmelblau, 1972; Rosenbrock, 1960; Palmer, 1969; DeCoursey y Snyder, 1969; Sorooshian, 1980; Sorooshian y Gupta, 1983; Sehitoglu, 1983); el de Regresión y Correlación (Yevjevich, 1972; Clarke, 1973); y el Método de Principio de Máxima Entropía (Tribus, 1969).

2.2. Distribución de frecuencias para una sola población.

Para predecir el gasto pico de una avenida asociada a cierto periodo de retorno o frecuencia de ocurrencia se emplean métodos estadísticos. Para aplicar estos métodos se requiere contar con registros de gastos máximos anuales o lluvias máximas en 24 horas, lográndose mayor precisión a medida que se cuente con mas años de registros.

Los métodos para evaluar el gasto máximo a partir de un periodo de retorno consideran que el gasto máximo anual es una variable aleatoria que se distribuye de acuerdo a una función de probabilidad. Por lo general, en nuestro país, se cuenta con pocos años de registro por lo que las curvas de distribución de probabilidad se tienen que prolongar en sus extremos a fin de poder inferir eventos mayores o menores a los registrados.

2.2.1. Función de Densidad Normal O Gauss

La función de densidad normal de la variable aleatoria X esta definida por:

$$F(x) = F(z) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dz \quad (2.1)$$

donde:

μ y σ son los parámetros de la función Normal.

μ = parámetro de ubicación

σ = parámetro de escala

μ = media

σ^2 = desviación estándar

$$\text{Media} = \frac{\sum_{i=1}^n X_i}{n}$$

$$\text{Desviación estándar} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

Estos parámetros determinan la forma de la distribución y su posición en el eje de las abscisas. Se puede demostrar que estos parámetros son la media y la desviación estándar de la población, hasta la fecha no se conoce analíticamente la integral de la ecuación 2.1, por lo que es necesario de métodos numéricos para valuarla, por lo que se ha definido la variable estandarizada.

$$z = \frac{x - \mu}{\sigma} \quad (2.2)$$

así la ecuación 2.1 se puede escribir como:

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2}} dz \quad (2.3)$$

Es decir, una normal con parámetros $a = 0$ y $b = 1$.

Tabla 2.1.1. Áreas bajo la curva Normal.

Z	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5754
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.9103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6885	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549
0.7	0.7580	0.7612	0.7624	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7996	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9218	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9998
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Partiendo de los datos de lluvia máxima en 24 horas y obteniendo su probabilidad de no - excedencia se entra a la tabla 2.1.1 de donde se obtiene la variable estandarizada z ecuación 2.2 y conocidas σ y μ , se tiene la incógnita que es la variable x , la cual representa a la lluvia ajustada, por lo que se despeja y se tiene:

$$H_p_{ajus} = x = z\sigma + \mu \tag{2.4}$$

Estimadores por Momentos

$$X_T = \mu + \sigma u_T \quad (2.5)$$

$$\hat{\mu} = \bar{x} \quad (2.6)$$

$$\hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{N - 1} \quad (2.7)$$

donde:

$$U_T = \omega - \frac{C_0 + C_1\omega + C_2\omega^2}{1 + d_1\omega + d_2\omega^2 + d_3\omega^3} \quad (2.8)$$

$$C_0 = 2.515517$$

$$C_1 = 0.802853$$

$$C_2 = 0.010328$$

$$d_1 = 1.432788$$

$$d_2 = 0.189269$$

$$d_3 = 0.001308$$

$$\text{Sea } P(t) = 1 - \frac{1}{tr} \quad (2.9)$$

Para $0 < P(t) \leq 0.5$

$$\omega = \sqrt{\text{Ln} \frac{1}{P(t)^2}} \quad (2.10)$$

Si $P(t) > 0.5$

$$\text{Sea } P(t) = \frac{1}{tr} \quad (2.11)$$

$$\omega = \sqrt{\text{Ln} \frac{1}{P(t)^2}} \quad (2.12)$$

2.2.2. Distribución Gumbel

Supongamos que se tienen N muestras, cada una contiene n eventos. Si se selecciona el máximo de los n eventos de cada muestra, es posible demostrar que a medida que n aumenta, la función de distribución de probabilidad de X tiende a:

$$F(x) = e^{-e^{-\alpha(x-\beta)}} \tag{2.13}$$

Por lo tanto la función de distribución de probabilidad es:

$$f(x) = \alpha e^{-\alpha(x-\beta)-e^{-\alpha(x-\beta)}} \tag{2.14}$$

donde:

α y β son los parámetros de la función.

Estos parámetros los podemos estimar como:

$$\alpha = \frac{1.2825}{S} \tag{2.15}$$

$$\beta = \bar{X} - 0.45S \tag{2.16}$$

donde:

\bar{X} : es la media de los datos y S la desviación estándar.

$$\text{Media} = E[X] = \hat{\mu} + 0.5772\hat{\alpha} \tag{2.17}$$

$$\text{Desviación estándar} = E[X - E(X)]^2 = \frac{\pi^2 \hat{\alpha}^2}{6} \tag{2.18}$$

Para muestras pequeñas los parámetros α y β se obtienen como:

$$\alpha = \frac{\sigma_y}{S} \tag{2.19}$$

$$\beta = \bar{x} - \frac{\mu_y}{\alpha} \tag{2.20}$$

μ_y y σ_y se obtienen de la tabla 2.1.2 donde N es el número de registros.

Tabla 2.1.2. Constantes del tamaño de la muestra N , en la función de distribución Gumbel.

N	μ_y	σ_y	N	μ_y	σ_y	N	μ_y	σ_y	N	μ_y	σ_y
8	0.484	0.904	29	0.534	1.104	48	0.547	1.157	76	0.556	1.190
9	0.490	0.928	29	0.535	1.108	49	0.548	1.159	78	0.556	1.192
10	0.495	0.949	30	0.536	1.112	50	0.548	1.160	80	0.556	1.193
11	0.499	0.967	31	0.537	1.115	51	0.548	1.162	82	0.557	1.195
12	0.503	0.983	32	0.538	1.119	52	0.549	1.163	84	0.557	1.196
13	0.507	0.997	33	0.538	1.122	53	0.549	1.165	86	0.558	1.198
14	0.510	1.009	34	0.539	1.125	54	0.550	1.166	88	0.558	1.199
15	0.512	1.020	35	0.540	1.128	55	0.550	1.168	90	0.558	1.200
16	0.515	1.031	36	0.541	1.131	56	0.550	1.169	92	0.558	1.202
17	0.518	1.041	37	0.541	1.133	57	0.551	1.170	94	0.559	1.203
18	0.520	1.049	38	0.542	1.136	58	0.551	1.172	96	0.559	1.204
19	0.522	1.056	39	0.543	1.138	59	0.551	1.173	98	0.559	1.205
20	0.523	1.062	40	0.543	1.141	60	0.552	1.174	100	0.560	1.206
21	0.525	1.069	41	0.544	1.143	62	0.552	1.177	150	0.564	1.225
22	0.526	1.075	42	0.544	1.145	64	0.553	1.179	200	0.567	1.236
23	0.528	1.081	43	0.545	1.148	66	0.553	1.181	250	0.568	1.242
24	0.529	1.086	44	0.545	1.149	68	0.554	1.183	300	0.68	1.247
25	0.530	1.091	45	0.546	1.151	70	0.554	1.185	500	0.572	1.258
26	0.532	1.096	46	0.546	1.153	72	0.555	1.187	1000	0.574	1.268
27	0.533	1.100	47	0.547	1.157	74	0.556	1.189			

Estimadores por Momentos

$$\hat{\mu} = \bar{X} - 0.5772\hat{\alpha} = \bar{X} - 0.45S \tag{2.21}$$

$$\hat{\alpha} = \frac{\sqrt{6}}{\pi} S = 0.78S \tag{2.22}$$

donde:

S: es la desviación estándar

2.2.3. Distribución Lognormal de 2 parametros

$$f(x) = \frac{1}{\sqrt{2\pi x\sigma_y}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu_y}{\sigma_y}\right)^2}; x > 0 \tag{2.23}$$

donde:

μ_y y σ_y ; son parámetros de ubicación y escala respectivamente.

Estimadores por Momentos

$$\hat{\mu}_y = \frac{1}{N} \sum \text{Ln}X \quad (2.24)$$

$$\hat{\sigma}_y = \sqrt{\frac{\sum (\text{Ln}x_i - \hat{\mu}_y)^2}{N-1}} \quad (2.25)$$

2.2.4. Lognormal de 3 parámetros.

$$f(x) = \frac{1}{\sqrt{2\pi}(x-x_0)\sigma_y} e^{-\frac{1}{2}\left(\frac{\text{Ln}(x-x_0)-\mu_y}{\sigma_y}\right)^2} \quad (2.26)$$

donde:

x_0 , μ y σ son parámetros de ubicación, forma y escala, respectivamente.

Estimadores por Momentos

$$\hat{X} = \bar{X} \left[1 - \frac{n_x}{n_z} \right] \quad (2.27)$$

donde:

n_x = coeficiente de variación de la serie X.

$$n_x = \frac{S}{\bar{X}} \quad (2.28)$$

n_z = coeficiente de variación de la serie $(X - \bar{X})$

$$n_z = \frac{1 - \omega^{3/2}}{\omega^{3/2}} \quad (2.29)$$

$$\omega = \frac{(\hat{y}_x^2 + 4)^{1/2} - \hat{y}_x}{2} \quad (2.30)$$

Sea:

$$\hat{\mu}_x = \frac{1}{N} \sum X_i \quad (2.31)$$

$$\hat{\sigma}_x = \left[\frac{\sum (X_i - \mu_x)^2}{N-1} \right]^{1/2} \quad (2.32)$$

Por lo que:

$$\hat{\mu}_y = Ln\left(\frac{\sigma_x}{n_z}\right) - 0.5Ln(n_z^2 + 1) \quad (2.33)$$

$$\hat{\sigma} = [Ln(n_z^2 + 1)]^{1/2} \quad (2.34)$$

donde:

x_0, μ_y, σ_y son parámetros de ubicación, forma y escala, respectivamente.

2.2.5. Distribución Gamma de 2 parámetros.

$$f(x) = \frac{x^{\beta-1} e^{-x/\alpha}}{\alpha^\beta \Gamma(\beta)} \quad (2.35)$$

donde:

α = parámetro de escala;

β = parámetro de forma

Γ = es el coeficiente de asimetría poblacional.

Estimadores por Momentos

La media y la desviación estándar en términos de α y β son:

$$\mu = \alpha\beta \quad (2.36)$$

$$\sigma = \alpha\sqrt{\beta} \quad (2.37)$$

Donde en forma explícita, los parámetros se expresan como:

$$\beta = \frac{\mu^2}{\sigma^2} \quad (2.38)$$

$$\alpha = \frac{\mu}{\beta} = \frac{\sigma^2}{\mu} \quad (2.39)$$

Por lo que los estimadores por momentos serán:

$$\hat{n} = \hat{\beta} = \left(\frac{\bar{X}}{S} \right)^2 \quad (2.40)$$

$$\frac{1}{\hat{\lambda}} = \hat{\alpha} = \frac{S^2}{\bar{X}} \quad (2.41)$$

donde:

$$\hat{\lambda} = \frac{1}{\hat{\alpha}} \quad (2.42)$$

$$\hat{n} = \hat{\beta} \quad (2.43)$$

2.2.6. Distribución Pearson III o Gamma de tres parámetros

$$f(x) = \frac{1}{\alpha_1 \Gamma(\beta_1)} \left\{ \frac{x - \delta_1}{\alpha_1} \right\}^{\beta_1 - 1} e^{-\frac{x - \delta_1}{\alpha_1}} \quad (2.44)$$

donde:

α_1 , β_1 y δ_1 , son los parámetros de la función y $\Gamma(\beta_1)$ es la función Gamma.

Media = $E[X] = X_0 + \alpha\beta$

Varianza = $E[X - E(X)]^2 = \alpha^2\beta$

2.2.7. Distribución Log Pearson Tipo III

$$F(X) = \frac{1}{\alpha\lambda\Gamma\beta} \left\{ \frac{LnX - Y_0}{\alpha} \right\}^{\beta - 1} e^{-\left[\frac{LnX - Y_0}{\alpha} \right]} \quad (2.45)$$

$$Y_0 = Ln X_0 \quad (2.46)$$

donde:

Y_0 , α y β son los parámetros de ubicación, escala y forma respectivamente.

Estimadores por Momentos

$$\hat{\beta} = \frac{4}{g^2} \quad (2.47)$$

donde:

g = coeficiente de asimetría.

$$\hat{\alpha} = \frac{gS}{2} = \frac{\sigma}{\sqrt{\beta}} = \frac{S}{\sqrt{\beta}} \quad (2.48)$$

$$\hat{X}_0 = \bar{X} - \hat{\alpha}\hat{\beta} = \bar{X} - S\sqrt{\beta} \quad (2.49)$$

2.2.8. General de Valores Extremos (GVE)

$$f(x) = \frac{1}{\alpha} \left[1 - \left(\frac{x - \mu}{\alpha} \right) \beta \right] e^{-\left(1 - \left[\frac{x - \mu}{\alpha} \right] \beta \right)^{1-\beta}} \quad (2.50)$$

donde:

u , α y β son los parámetros de ubicación, escala y forma, respectivamente.

Sea g = coeficiente de asimetría muestral

Si $-11.35 < g < 1.1396$

entonces:

$$\hat{\beta} = 0.279434 - 0.333535g + 0.048306g^2 + 0.023314g^3 + 0.00376g^4 + 0.000263g^5 \quad (2.51)$$

Si $1.14 < g < 18.95$

entonces:

$$\hat{\beta} = 0.25031 - 0.29219g + 0.075357g^2 - 0.010883g^3 + 0.000904g^4 - 0.000043g^5 \quad (2.52)$$

Si $\beta < 0$ la variable es de tipo II

$$X_2 = \mu + \frac{\alpha}{\beta} - \frac{\alpha}{\beta} y_2 \quad (2.53)$$

Relación entre X y la variable reducida y .

$$\text{Sea } A = \mu + \frac{\alpha}{\beta} \quad (2.54)$$

Nuevo parámetro de ubicación

$$\text{Sea } B = -\frac{\alpha}{\beta} > 0 \quad (2.55)$$

Nuevo parámetro de escala

Por lo tanto:

$$X_2 = A + By_2 \quad (2.56)$$

$$E[X_2] = \hat{A} + \hat{B}E[y_2] \quad (2.57)$$

$$\text{Var}[X_2] = \hat{\beta}^2 \text{Var}(y_2) \quad (2.58)$$

$$E[X_2] = S^2 \quad (2.59)$$

$$\text{Var}[X_2] = \bar{X} \quad (2.60)$$

Estimación de parámetros

$$\hat{\beta} = \left[\frac{\text{Var}(X_2)}{\text{Var}(Y_2)} \right]^{1/2} \quad (2.61)$$

$$\hat{A} = E[X_2] - \hat{B}E[Y_2] = \bar{X} - \hat{B}E[Y_2] \quad (2.62)$$

donde:

$$E[Y_2] = \Gamma(1 + \hat{\beta}) \quad (2.63)$$

$$\text{Var}[Y_2] = \Gamma(1 + 2\hat{\beta}) - \Gamma^2(1 + \hat{\beta}) \quad (2.64)$$

Por lo que los estimadores de los parámetros están definidos por:

$$\hat{\alpha} = -\hat{\beta}\hat{\beta} \quad (2.65)$$

$$\hat{\mu} = \hat{\lambda} - \frac{\hat{\alpha}}{\hat{\beta}} = \hat{\lambda} + \hat{\beta} \quad (2.66)$$

donde:

$\hat{\beta}$ ya fue presentado previamente.

Por otra parte, si $\hat{\beta} > 0$, la variable es de tipo III, donde:

$$X_3 = \mu \frac{\alpha}{\beta} + \frac{\alpha}{\beta} Y_3 \quad (2.67)$$

donde:

$$\hat{\alpha} = \hat{\beta} \hat{\beta} \quad (2.68)$$

$$\hat{\mu} = \hat{\lambda} - \hat{\beta} \quad (2.69)$$

$$\hat{\lambda} = \bar{X} - \hat{\beta} E[Y_3] \quad (2.70)$$

$$\hat{\beta} = \left[\frac{Var(X_3)}{Var(Y_3)} \right]^{1/2} \quad (2.71)$$

$$E[Y_3] = -\Gamma(1 + \hat{\beta}) \quad (2.72)$$

$$Var(Y_3) = \Gamma(1 + 2\hat{\beta}) - \Gamma^2(1 + \hat{\beta}) \quad (2.73)$$

2.3. Mínimo error cuadrático

Para cada valor que tenemos de lluvia (X_i), se obtiene su respectivo ajuste como se acaba de describir. Después, con el propósito de poder elegir los valores de alguna función en particular, conviene comparar uno y otro método de ajuste de los datos a las distintas funciones de distribución, uno de estos métodos es el de mínimo error cuadrático, que consiste en obtener la raíz cuadrada del cuadrado de la suma de las diferencias existentes entre el dato X_i menos el valor ajustado X , esto es:

$$E^2 = \sqrt{\sum_{i=1}^n (x_i - Hp_{ajus})^2}$$

A cada valor ajustado le corresponde un determinado período de retorno, pero para cada obra en particular, esos valores ajustados no tienen el mismo valor de T_r con el que pretende diseñar dicha obra, por lo que el objetivo del ajuste probabilístico es determinar el valor de dicho evento (lluvia o gasto) en el período de años para el cual se pretenda diseñar determinada obra.

CAPITULO III



CAPITULO TRES

3.1. CASOS ESPECIALES EN HIDROLOGÍA

Los casos especiales en hidrología son representados por la siguiente ecuación:

$$X = \bar{X} + KS$$

donde:

\bar{X} = media.

S = desviación estándar.

K = factor de frecuencia cuyo valor depende del tipo de función de distribución de probabilidad que se use.

Así, por ejemplo se tiene que para una distribución normal se expresa como:

$$X = \bar{X} + zS$$

donde z es el valor de la variable estandarizada y se obtiene de la tabla 2.1.1., en función de la probabilidad que se le asigna al evento x.

Para la función de distribución de probabilidad Gumbel el factor de frecuencia esta dado por:

$$K = -\frac{\sqrt{6}}{\pi} \left[0.5772 + \text{Ln} \left(\text{Ln} \left(\frac{Tr}{Tr-1} \right) \right) \right]$$

Sin embargo, esta ecuación es aplicable cuando la muestra es de tamaño infinito, esto se cumple cuando se dispone de una muestra mayor de 100 datos, lo cual no ocurre en nuestro medio, por lo que el valor de K se obtiene de la siguiente tabla:

Tabla 3.1. Valores del factor de frecuencia K, para la distribución Gumbel.

Tamaño de la muestra	Periodo de retorno, en años						
	10	20	25	50	75	100	1000
15	1.703	2.410	2.632	3.321	3.721	4.005	6.265
20	1.625	2.302	2.517	3.179	3.563	3.836	6.006
25	1.575	2.235	2.444	3.088	3.463	3.729	5.842
30	1.541	2.188	2.393	3.026	3.393	3.653	5.727
40	1.495	2.126	2.326	2.943	3.301	3.554	5.476
50	1.466	2.086	2.283	2.889	3.241	3.491	5.478
60	1.466	2.059	2.253	2.852	3.200	3.446	5.419
70	1.430	2.030	2.230	2.824	3.169	3.413	5.359
75	1.423	2.029	2.220	2.812	3.155	3.400	5.339
100	1.401	1.998	2.187	2.770	3.109	3.349	5.261

Función de distribución Pearson tipo III.

Esta función es un caso especial de la Gamma y se le conoce como Log Pearson en donde la transformación $y = \log x$ se usa para reducir la asimetría; en caso de que la asimetría para esta situación valga cero la distribución Log Pearson III se reduce a un Lognormal y la Pearson tipo III a una normal. La distribución Log Pearson se describe:

$$\log x = \bar{y} + KS_y$$

$$\bar{y} = \sum_{i=1}^n \frac{\log x_i}{n}$$

$$S_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}$$

$$y_i = \log x_i$$

K se obtiene de la siguiente tabla.

Tabla 3.2. Valores de K para la función de distribución Pearson tipo III

Coeficiente de asimetría g	Periodo de retorno, en años										
	1.0101	1.0526	1.1111	1.2500	2	5	10	25	50	100	200
<i>Asimetría positiva</i>											
3.0	-0.667	-0.665	-0.660	-0.636	-0.396	0.420	1.180	2.278	3.152	4.051	4.970
2.9	-0.690	-0.668	-0.681	-0.651	-0.390	0.440	1.195	2.277	3.134	4.013	4.909
2.8	-0.714	-0.711	-0.702	-0.666	-0.384	0.460	1.210	2.775	3.114	3.973	4.847
2.7	-0.740	-0.736	-0.724	-0.681	-0.376	0.479	1.224	2.272	3.093	3.932	4.783
2.6	-0.769	-0.762	-0.747	-0.696	-0.368	0.499	1.238	2.267	0.071	3.889	4.718
2.5	-0.799	-0.790	-0.771	-0.711	-0.360	0.518	1.250	2.262	3.048	3.845	4.652
2.4	-0.832	-0.819	-0.795	-0.725	-0.351	0.537	1.262	2.256	3.023	3.800	4.584
2.3	-0.867	-0.850	-0.819	-0.739	-0.341	0.555	1.274	2.248	2.997	3.753	4.515
2.2	-0.905	-0.882	-0.844	-0.752	-0.330	0.574	1.284	2.240	2.970	3.705	4.444
2.1	-0.946	-0.914	-0.869	-0.765	-0.319	0.592	1.294	2.230	2.942	3.656	4.372
2.0	-0.990	-0.949	-0.895	-0.777	-0.307	0.609	1.302	2.219	2.912	3.605	4.398
1.9	-1.037	-0.984	-0.920	-0.788	-0.294	0.627	1.310	2.207	2.881	3.553	4.223
1.8	-1.087	-1.020	-0.945	-0.799	-0.282	0.643	1.318	2.193	2.848	3.499	4.147
1.7	-1.140	-1.056	-0.970	-0.808	-0.268	0.660	1.324	2.179	2.815	3.444	4.069
1.6	-1.197	-1.093	-0.994	-0.817	-0.254	0.675	1.329	2.163	2.780	3.388	3.990
1.5	-1.256	-1.131	-1.018	-0.825	-0.240	0.690	1.333	2.146	2.743	3.330	3.910
1.4	-1.318	-1.168	-1.041	-0.832	-0.225	0.705	1.337	2.128	2.706	3.271	3.828
1.3	-1.383	-1.206	-1.064	-0.838	-0.210	0.719	1.339	2.108	2.666	3.211	3.745
1.2	-1.449	-1.243	-1.086	-0.844	-0.195	0.732	1.340	2.087	2.626	3.149	3.661
1.1	-1.518	-1.280	-1.107	-0.848	-0.180	0.745	1.341	2.066	2.585	3.087	3.575
1.0	-1.588	-1.317	-1.128	-0.852	-0.164	0.758	1.340	2.043	2.542	3.022	3.489
0.9	-1.660	-1.353	-1.147	-0.854	-0.148	0.769	1.339	2.016	2.498	2.957	3.401
0.8	-1.733	-1.388	-1.166	-0.856	-0.132	0.780	1.336	1.993	2.453	2.891	3.312
0.7	-1.806	-1.423	-1.183	-0.857	-0.116	0.790	1.333	1.967	2.407	2.824	3.223
0.6	-1.880	-1.458	-1.200	-0.857	-0.099	0.800	1.328	1.939	2.359	2.755	3.132
0.5	-1.955	-1.491	-1.216	-0.856	-0.083	0.808	1.323	1.910	2.311	2.686	3.041
0.4	-2.029	-1.524	-1.231	-0.855	-0.066	0.816	1.317	1.880	2.261	2.615	2.949
0.3	-2.104	-1.555	-1.245	-0.853	-0.050	0.824	1.309	1.849	2.211	2.544	2.856
0.2	-2.178	-1.586	-1.258	-0.850	-0.033	0.830	1.301	1.818	2.159	2.472	2.763
0.1	-2.252	-1.616	-1.270	-0.846	-0.017	0.836	1.292	1.785	2.107	2.400	2.670
0.0	-2.326	-1.645	-1.820	-0.842	0.000	0.842	1.282	1.751	2.054	2.326	2.576

Coeficiente de asimetría g	Periodo de retorno, en años										
	1.0101	1.0526	1.1111	1.2500	2	5	10	25	50	100	200
<i>Asimetría negativa</i>											
-0.1	-2.400	-1.673	-1.292	-0.836	0.017	0.846	1.270	1.716	2.000	2.252	2.482
-0.2	-2.472	-1.700	-1.301	-0.830	0.033	0.850	1.258	1.680	1.945	2.178	2.388
-0.3	-2.544	-1.726	-1.309	-0.824	0.050	0.853	1.245	1.643	1.890	2.104	2.294
-0.4	-2.615	-1.750	-1.317	-0.816	0.066	0.855	1.231	1.606	1.834	2.029	2.201
-0.5	-2.686	-1.774	-1.323	-0.808	0.083	0.856	1.216	1.567	1.777	1.955	2.108
-0.6	-2.755	-1.797	-1.328	-0.800	0.099	0.857	1.200	1.528	1.720	1.880	2.016
-0.7	-2.824	-1.819	-1.333	-0.790	0.116	0.857	1.183	1.488	1.663	1.806	1.926
-0.8	-2.891	-1.839	-1.336	-0.780	0.132	0.856	1.166	1.448	1.606	1.733	1.837
-0.9	-2.957	-1.858	-1.339	-0.769	0.148	0.854	1.147	1.407	1.549	1.660	1.749
-1.0	-3.022	-1.877	-1.340	-0.758	0.164	0.852	1.128	1.366	1.492	1.588	1.664
-1.1	-3.090	-1.894	-1.341	-0.745	0.180	0.846	1.107	1.324	1.435	1.518	1.581
-1.2	-3.149	-1.910	-1.340	-0.732	0.195	0.844	1.086	1.282	1.379	1.449	1.501
-1.3	-3.211	-1.925	-1.339	-0.719	0.210	0.838	1.064	1.240	1.324	1.383	1.424
-1.4	-3.271	-1.938	-1.337	-0.705	0.225	0.832	1.041	1.198	1.270	1.318	1.351
-1.5	-3.330	-1.951	-1.333	-0.690	0.240	0.825	1.018	1.157	1.217	1.256	1.282
-1.6	-3.388	-1.962	-1.329	-0.675	0.254	0.817	0.994	1.116	1.166	1.197	1.216
-1.7	-3.444	-1.972	-1.324	-0.660	0.268	0.808	0.970	1.075	1.116	1.140	1.155
-1.8	-3.499	-1.981	-1.318	-0.643	0.282	0.799	0.945	1.035	1.059	1.087	1.097
-1.9	-3.553	-1.989	-1.310	-0.627	0.294	0.788	0.920	0.996	1.023	1.037	1.044
-2.0	-3.605	-1.996	-1.302	-0.609	0.307	0.777	0.895	0.959	0.980	0.990	0.995
-2.1	-3.656	-2.001	-1.294	-0.592	0.319	0.765	0.869	0.923	0.939	0.946	0.949
-2.2	-3.705	-2.006	-1.284	-0.574	0.330	0.752	0.844	0.888	0.900	0.905	0.907
-2.3	-3.753	-2.009	-1.274	-0.555	0.341	0.739	0.819	0.855	0.864	0.867	0.869
-2.4	-3.800	-2.001	-1.262	-0.537	0.351	0.725	0.795	0.823	0.830	0.832	0.833
-2.5	-3.845	-2.012	-1.250	-0.518	0.360	0.711	0.771	0.793	0.798	0.799	0.800
-2.6	-3.889	-2.013	-1.238	-0.499	0.368	0.696	0.747	0.764	0.768	0.769	0.769
-2.7	-3.932	-2.012	-1.224	-0.479	0.376	0.681	0.724	0.738	0.740	0.740	0.741
-2.8	-3.973	-2.010	-1.210	-0.460	0.384	0.666	0.702	0.712	0.714	0.714	0.714
-2.9	-4.013	-2.007	-1.195	-0.440	0.390	0.651	0.681	0.683	0.689	0.690	0.690
-3.0	-4.051	-2.003	-1.180	-0.420	0.396	0.636	0.660	0.666	0.666	0.667	0.667

3.2. DETERMINACIÓN DE CURVAS DE FRECUENCIA

3.2.1.- Distribución

Los eventos de diluvio son una sucesión de fenómenos naturales que es muy complicado determinarlos debido a que no se puede aplicar ninguna distribución estadística conocida. Con diversos estudios que se han llevado a cabo se ha demostrado que para resolver este tipo de problemas en los cuales se deben definir las probabilidades díciles de diluvio, la Distribución Log Pearson tipo III es el método más exacto para el análisis de datos de la serie anual que utilizan un coeficiente oblicuo generalizado como se describe a continuación.

3.2.2.- Aplicación de la distribución

La forma recomendada para aplicar la distribución Log Pearson III, es calcular los logaritmos de base 10 de la descarga Q, a una probabilidad de excedencia seleccionada P, por la siguiente ecuación:

$$\text{Log } Q = \bar{X} + KS \quad (3.1)$$

donde:

\bar{X} y S: son la media y desviación estándar y K es un factor de la función, el coeficiente oblicuo y la probabilidad de excedencia seleccionada.

Los valores de K se obtiene en el apéndice 1.

La media, la desviación estándar y coeficiente de inclinación de datos de la estación se obtienen con las siguientes ecuaciones:

$$\bar{X} = \frac{\sum X}{N} \quad (3.2)$$

$$S = \left[\frac{\sum (X - \bar{X})^2}{(N-1)} \right]^{0.5} \quad (3.3)$$

$$S = \left[\frac{(\sum X^2) - \frac{(\sum X)^2}{N}}{(N-1)} \right]^{0.5} \quad (3.4)$$

$$G = \frac{N \sum (X - \bar{X})^3}{(N-1)(N-2)S^3} \quad (3.5)$$

$$G = \frac{N^2(\sum X^3) - 3N(\sum X)(\sum X^2) + 2(\sum X)^3}{N(N-1)(N-2)S^3} \quad (3.6)$$

donde:

- X = logaritmo de flujo de la cresta anual
 N = el número de artículos en juego de los datos
 \bar{X} = logaritmo medio
 S = desviación normal de logaritmos
 G = coeficiente oblicuo de logaritmos

La precisión de los valores calculados con las ecuaciones (3.4) y (3.6) es más sensible que los obtenidos con las ecuaciones (3.3) y (3.5) al número de dígitos significantes usados en su cálculo. Cuando los medios del calculo disponibles sólo mantienen un número limitado de dígitos significantes, las ecuaciones (3.4) y (3.6) son preferibles.

3.2.3. Estimación de Inclinación Generalizada

El coeficiente oblicuo del registro de la estación (inclinación de la estación) es sensible a los eventos extremos; así es difícil de obtener estimaciones de inclinaciones exactas de las muestras pequeñas. La exactitud del coeficiente oblicuo puede ser mejorada por la inclinación de la estación con inclinación generalizada estimada, agrupando información de los sitios cercanos. Las pautas siguientes se recomiendan para estimar la inclinación generalizada.

El procedimiento recomendado por desarrollar coeficientes oblicuos generalizados requiere el uso de por lo menos 40 estaciones, o todas las estaciones dentro de un radio de 170 km. El procedimiento incluye análisis a través de tres métodos:

- 1) Isolneas oblicuas utilizando un mapa
- 2) La ecuación de la predicción oblicua
- 3) La media de los valores oblicuos de la estación.

Cada uno de los métodos se discute separadamente.

En México no se cuenta con mapa de isolneas, solo se cuenta con ellas en los Estados Unidos, pero aun así se explicara el procedimiento para utilizarlas.

Para desarrollar las isolneas, se trazan líneas de los valores oblicuos de cada estación al centro de su cubeta de desagüe y se examinan los datos trazados para cualquier tendencia geográfica o topográfica. Si un modelo es evidente, entonces las isolneas son arrastradas y el promedio de las diferencias cuadradas valoraran el error medio cuadrado (*MSE*).

El *MSE* se usará estimando la exactitud del mapa del isoclinas. Si ningún modelo es evidente, entonces un mapa del isoclinas no puede ser dibujado y por consiguiente no será considerado.

La ecuación de la predicción debe usarse preferentemente por estimar el coeficiente oblicuo en las estaciones con variables que están dentro del rango de datos. El *MSE* se usará para evaluar la exactitud de la ecuación.

Se determina la media aritmética y la variación del régimen de escurrimiento para 40 estaciones, pero si no se cuenta con estas se puede realizar con 20 estaciones para estimar el coeficiente oblicuo generalizado. Las áreas del desagüe y meteorológicas, las características topográficas, y geológicas deben ser representativas de la región alrededor de la estación de interés.

Se selecciona el método que proporciona las estimaciones de coeficiente de inclinación más exacto. Se compara el *MSE* obtenido por las isoclinas y el obtenido por la ecuación de la predicción. El *MSE* más pequeño debe compararse a la variación de los datos. Si el *MSE* no es significativamente más pequeño que la variación, la isocline tampoco, ni la ecuación de la predicción proporciona una estimación más exacta del coeficiente oblicuo que hace el valor medio. El coeficiente oblicuo medio debe usarse cuando se proporciona la estimación más exacta y la variación debe usarse en la ecuación (3.7) para $MSE_{\bar{G}}$.

3.2.4. Coeficiente Oblicuo

La estación y el coeficiente oblicuo pueden combinarse para formar una mejor estimación de la inclinación. Bajo la asunción que la inclinación generalizada es imparcial e independiente de la inclinación de la estación, el error medio cuadrado (*MSE*) de la estimación es minimizado y se generaliza una inclinación en proporción del inverso a su ecuación. La ecuación siguiente debe utilizarse para calcular el coeficiente oblicuo:

$$G_w = \frac{MSE_{\bar{G}}(G) + MSE_G(\bar{G})}{MSE_{\bar{G}} + MSE_G} \quad (3.7)$$

donde:

G_w = Coeficiente oblicuo

G = Inclinación de la estación

\bar{G} = Inclinación generalizada

$MSE_{\bar{G}}$ = Error medio cuadrado de la inclinación generalizada

MSE_G = Error medio cuadrado de la inclinación de la estación

La ecuación (3.7) puede usarse para calcular una estimación de inclinación sin tener en cuenta la fuente de inclinación generalizada, con tal de que el MSE de la inclinación generalizada puede estimarse. En este caso se tomara $MSE_{\tilde{G}} = 0.302$, para utilizarse en la ecuación (3.7). Para el uso en G_w , esta función (MSE_G) puede aproximarse con exactitud suficiente por la ecuación:

$$MSE_G \approx 10^{\left[A - B \left[\log_{10} \left(\frac{N}{10} \right) \right] \right]} \quad (3.8)$$

donde:

$$A = -0.33 + 0.08 |G| \text{ si } |G| \leq 0.90$$

$$A = -0.52 + 0.30 |G| \text{ si } |G| > 0.90$$

$$B = 0.94 - 0.26 |G| \text{ si } |G| \leq 1.50$$

$$B = 0.55 \text{ si } |G| > 1.50$$

en que $|G|$ es el valor absoluto de la inclinación de la estación (usado como una estimación de inclinación de la población) y N es la longitud del registro en años. Si el ajuste histórico descrito en el punto 11 se ha aplicado, la inclinación históricamente ajustada \tilde{G} , y el periodo histórico H , serán usados para G y N respectivamente, en la ecuación (3.8).

La aplicación de la ecuación (3.8) y el apéndice 2, a las estaciones con valor de inclinación absoluto mayor que 2 y los periodos largos de registro dan peso relativamente pequeño al valor de la estación.

3.2.5. Registro roto

Un registro roto es aquel en el cual, los datos de una estación no presentan una homogeneidad, esto es que contenga datos ya sea extremadamente altos o bajos, a comparación de los demás datos, esto hace que se presenten outliers altos o bajos que son descritos mas adelante, y por lo regular son los registros que son tratados con funciones de probabilidad mixtas.

3.2.6. Registro incompleto

Un registro incompleto se refiere a un registro de una corriente en el que el poco de flujo del agua o el agua en exceso, provocan que no se puedan medir sus gastos, y no se registren sus datos en las estaciones. Los datos altos o bajos requieren de tratamiento diferente.

Cuando no se han registrado los gastos mas altos durante el periodo de registro sistemático, hay normalmente información disponible de que la descarga de la cresta pueda estimarse.

3.2.7. Años con Avenidas de Magnitud Cero

Algunas corrientes en zonas áridas no presentan escurrimiento a lo largo de todo el año. Así la serie de avenidas anuales en estas corrientes tendrán uno o más valores de avenidas con magnitud cero. Esto imposibilita el análisis estadístico normal de los datos al aplicar algunas funciones de distribución de probabilidad recomendada de la familia Log Pearson, en caso especial se hará referencia a la distribución Log Pearson tipo III, debido a que el logaritmo de cero es menos infinito. En este caso se recomienda el ajuste de la probabilidad condicional para determinar la curva de frecuencias de registros con avenidas máximas anuales de magnitud cero.

Ajuste de probabilidad condicional

En estaciones donde el registro de gastos máximos anuales se trunca por la omisión de gastos pico inferiores a un nivel de referencia, años con gastos de magnitud cero y gastos extraños de baja magnitud, se recomienda el ajuste de la probabilidad condicional para obtener la curva de frecuencias. Este procedimiento es aplicable cuando no mas del 25 % del registro total ha sido truncado. Se define un nivel de truncacion como el gasto mínimo que excluirá los gastos pico por abajo del nivel base de la estación, gastos cero, todos los gastos extraños de baja magnitud y ningún otro gasto, debido a que los datos de las estaciones tratadas por este procedimiento no pueden ajustarse a la distribución Log Pearson tipo III ninguna curva de frecuencias calculada deberá compararse con una gráfica de los valores observados.

Los datos deberán revisarse para calcular los estadísticos de los gastos pico superiores al nivel de referencia, antes de aplicar el ajuste de probabilidad condicional. Esto es, que se hayan concluido los procedimientos para detectar los gastos extraños, recalculando los estadísticos de los gastos pico superiores al nivel del truncamiento e incorporando información histórica aplicable. Todos, excepto la última etapa de cálculo descrita en el siguiente diagrama de flujo deben haberse concluido.

Las etapas para el ajuste de probabilidad condicional se describen a continuación:

1. – Estime la probabilidad \tilde{P} de que cualquier gasto pico anual excederá el nivel de truncamiento con la fórmula:

$$\tilde{P} = \frac{N}{n} \quad (3.9)$$

donde:

N : es el número de gastos pico anuales superiores al nivel de truncamiento

n : es el número total de años de registro.

Si se incluye la información histórica, entonces deberá aplicarse la ecuación (3.10) en lugar de la (3.11)

$$\tilde{P} = \frac{H - WL}{H} \quad (3.11)$$

donde:

H : es la longitud histórica del registro

L : es el número de gastos pico truncados, y

W : es la ponderación del registro sistemático, que se calcula con la ecuación (3.3)

$$W = \frac{H - Z}{N + L} \quad (3.12)$$

donde:

H : Longitud histórica del registro

Z : Número de gastos pico históricos incluyendo gastos extraños (outliers) con información histórica

N : Número de gastos pico anuales superiores al nivel de truncamiento

L : Número de gastos pico truncados

2. – Vuelva a calcular las probabilidades de excedencia P , para puntos seleccionados, P_d , en la curva de frecuencia al aplicar la siguiente expresión:

$$P = \tilde{P} \cdot P_d \quad (3.13)$$

La ecuación anterior considera la omisión de gastos pico inferiores al nivel de truncamiento.

3. - Las probabilidades de excedencia P , calculadas con la ecuación (3.13), comúnmente no son aquellas requeridas para calcular los estadísticos de la muestra sintética. Por consiguiente, es necesario interpolar gráfica o analíticamente para obtener log - valores de descarga para las probabilidades de excedencia 0.01, 0.10, y 0.50 probabilidades de excedencia.

4. - Debido a que la probabilidad condicional se ajustó la curva de frecuencia sin conocer los estadísticos, serán entonces calculados los sintéticos. Estos estadísticos sintéticos se determinan con base en los valores de las tres probabilidades de excedencia definidas en el paso 3, al aplicar las ecuaciones siguientes:

$$G_s = -2.50 + 3.12 \frac{\text{Log}\left(\frac{Q_{0.01}}{Q_{0.10}}\right)}{\text{Log}\left(\frac{Q_{0.10}}{Q_{0.50}}\right)} \quad (3.14)$$

$$S_s = \frac{\text{Log}\left(\frac{Q_{0.01}}{Q_{0.50}}\right)}{K_{0.01} - K_{0.50}} \quad (3.15)$$

$$\bar{X}_s = \text{Log}(Q_{0.50}) - K_{0.50}(S_s) \quad (3.16)$$

donde:

G_s , S_s , y X_s son el coeficiente de asimetría sintético logarítmico, desviación estándar y media, respectivamente; $Q_{0.01}$, $Q_{0.10}$, y $Q_{0.50}$ son los gastos pico con probabilidades de excedencia de 0.01, y 0.10, y 0.50, respectivamente, y $K_{0.01}$ y $K_{0.50}$ son las desviaciones Pearson Tipo III para las probabilidades de excedencia de 0.01 y 0.50, respectivamente, y G_s , es el coeficiente de asimetría. La ecuación (3.14) es una aproximación apropiada para aplicarse en el rango de -2 a $+2.5$ del coeficiente de asimetría.

5. - La curva de frecuencia desarrollada con los estadísticos sintéticos debe compararse con gastos pico anuales observados. La localización en el gráfico deberá considerar el número total de años de registro, n o H , como es lo apropiado.

El requisito adicional mínimo para llegar a una curva de frecuencia final es la determinación de la asimetría ponderada. Se desarrollarán algunos ejemplos numéricos para ilustrar los pasos básicos en el proceso de cálculo la curva de frecuencias al aplicar el ajuste de la probabilidad condicional.

Otras consideraciones en un análisis completo podrían incluir la comparación entre dos estaciones, uso de datos de lluvia, u otras técnicas que se describen en este trabajo.

3.2.8. Poblaciones Mixtas

En algunas regiones las avenidas son generadas por diferentes tipos de eventos. Por ejemplo, la avenida puede generarse por deshielo, tormenta o por una combinación de ambos fenómenos. Por lo que resultan curvas de frecuencias con grandes coeficientes de asimetría anormal reflejado en cambios de pendiente anormal cuando son graficados en papel de probabilidad Log Normal. Estos registros no son homogéneos y requieren un tratamiento especial. En algunas situaciones la curva de frecuencia de eventos anuales se describen mejor al calcular curvas separadas para cada tipo de evento, para posteriormente combinarlas en una sola.

Un ejemplo de combinación de diferentes tipos de eventos de diluvio son las tormentas tropicales con tormentas ciclónicas generales.

La separación por periodo del calendario en lugar de la separación por eventos no es considerada hidrología razonable a menos que los eventos en los periodos separados son causados claramente por condiciones hidrometeorológicas diferentes.

Si no pueden identificarse los eventos de diluvio que se cree que comprende dos o más poblaciones y pueden ser separados por un objetivo hidrológico, el criterio significativo, el registro se tratará como venir de una población.

En muchos lugares, especialmente en las zonas costeras de México, los gastos máximos anuales pertenecen a dos poblaciones diferentes: la primera es la de los gastos producidos por precipitaciones relacionadas con los fenómenos meteorológicos dominantes en la región en estudio, y la segunda es la de los gastos producidos por precipitaciones ciclónicas, normalmente mayores que los primeros.

También sucede en algunas zonas, como en el noroeste de la República Mexicana, que una población está compuesta por gastos producidos directamente por precipitaciones, y otras por gastos provenientes de deshielos. En adelante se mencionaran gastos producidos por tormentas ciclónicas o no ciclónicas por ser el más común, pero no debe olvidarse que las mismas ideas se pueden aplicar también a casos como el descrito.

Se ha demostrado que en estos casos, la función de distribución de probabilidad se puede expresar como:

$$F(x) = F_1(x)[p + (1 - p)F_2(x)] \quad (3.17)$$

donde:

$F_1(x)$ y $F_2(x)$ son, respectivamente las funciones de distribución de probabilidad de los gastos máximos anuales no producidos por tormentas ciclónicas y de los producidos por ellas, y p es la probabilidad de que en un año cualquiera el gasto máximo no sea producido por una tormenta ciclónica.

El número de parámetros de la función esta dado por:

$$N = n_1 + n_2 + 1 \quad (3.18)$$

donde:

n_1 = número de parámetros de $F1(x)$,

n_2 = número de parámetros de $F2(x)$

p = parámetro restante

En las zonas del país que son afectadas esporádicamente por los ciclones, los registros hidrométricos de gastos máximos anuales están integrados por eventos que proceden de dos tipos de fenómenos meteorológicos; el análisis estadístico de tales muestras tiene dos enfoques o tendencias, el primero consiste en segregar la muestra y obtener la función de distribución combinada y el segundo, se basa en funciones de distribución Gumbel de dos poblaciones o Gumbel mixta.

Los registros hidrométricos de gastos máximos anuales en las zonas del país afectadas esporádicamente por ciclones tropicales muestran la presencia de dos tipos de crecientes, las precipitaciones comunes o normales de la región y las ciclónicas. El U.S Army Corps of Engineers indica que cuando la cuenca bajo estudio tiene una magnitud mayor de 1300 km², es muy común que las precipitaciones ciclónicas provoquen las más grandes crecientes, debido al comportamiento hidrológico de la cuenca, que por ser grande es sensible a lluvias de baja intensidad, pero de amplia cobertura y duración (Morris, 1982).

El análisis estadístico de un registro de gastos máximos anuales, que está integrado por valores procedentes de dos poblaciones, ha mostrado dos tendencias fundamentales, la primera consiste en segregar la muestra y analizar por separado cada población (Browzin, 1972), o bien, en obtener una función de distribución combinada (Morris, 1982; Campos, 1987); la segunda utiliza una función de distribución mixta o para dos poblaciones mezcladas, enfoque al que se ha recurrido mucho en nuestro país (González, 1970) y (Raynal Santillán, 1986). Otros planteamientos dentro de esta segunda tendencia son el uso de distribuciones normales mixtas (Singh y Sinclair, 1972) y recientemente el de distribuciones basadas en procesos de Poisson compuestos (Rossi et al, 1984).

En términos generales, el primer enfoque de análisis estadístico se considera más adecuado, excepto no esta claro que una determinada población sea la responsable del cambio súbito de tendencia que muestran los datos al llevarlos a un papel de probabilidad Gumbel, o cuando el número de eventos de la muestra es reducido y, por tanto, insuficiente para definir con exactitud la función de distribución de cada población.

Dentro del segundo planteamiento de análisis, cuando se define en el papel Gumbel una curva formada por dos segmentos casi rectos, a partir de los datos de la muestra, (González, 1970) indica que el de menor pendiente se asocia a las crecientes no ciclónicas y el de pendiente mayor, a las avenidas ciclónicas y el de pendiente mayor, a las avenidas ciclónicas. Esto permite ajustar la función de distribución Gumbel de dos poblaciones o Gumbel mixta.

Conviene puntualizar que se acepta que solo las mayores crecientes fueron causadas por los ciclones, esto es, debe existir una correspondencia total entre la presencia de éstos de la cuenca y las fechas de ocurrencia de las avenidas máximas; esto implica, como un mínimo, que las mayores crecientes de registro hayan acontecido de julio a noviembre, que es la temporada normal de incidencia de los ciclones en la República Mexicana. Si no se cumple con tal relación causa efecto, la aplicación de la distribución Gumbel mixta es incorrecta y no trabaja dentro de un contexto hidrológico, sino únicamente estadístico - matemático. Esto debe ser tomado en cuenta al aplicar cualquier función de distribución para dos poblaciones mezcladas.

Gumbel Mixta o de Dos Poblaciones

Esta función de distribución se utiliza en los casos en que se tienen datos tipo ciclónico y no ciclónico: los primeros producidos por tormentas ciclónicas y los segundos resultado de los fenómenos meteorológicos dominantes de la región en estudio.

De la función de distribución de probabilidad Gumbel se deduce la ecuación para el mismo método para el caso de dos poblaciones:

$$f(x) = e^{-e^{-\alpha_1(x-A)}} \left[p + (1-p)e^{-e^{-\alpha_2(x-A)}} \right] \quad (3.19)$$

donde:

α_1 y β_1 son los parámetros correspondientes a la población no ciclónica y α_2 y β_2 corresponden a la población ciclónica.

El valor de P se calcula con:

$$p = \frac{N_n}{N_T} \quad (3.20)$$

donde:

N_n : es el número de años de registro no ciclónico.

N_T : es el número total de años de registro.

Ahora bien, para saber si un dato es ciclónico o no lo es, se pueden inspeccionar los datos. En general se grafican los datos contra su periodo de retorno, en escala Gumbel y en donde se observa un salto brusco. Se dice que la población es ciclónica, lo cual se muestra en la gráfica, para la estación climatológica Poza Rica, y haciendo uso de la tabla 2.1.2. se obtienen σ_y y μ_y para la población no ciclónica y total. Posteriormente se determinan los parámetros de esta función, para cada tipo de población y mediante tanteos se determina el valor ajustado X , hasta dar con la probabilidad de no excedencia P , que es el valor de la función de densidad de probabilidad.

3.2.9. Outliers

Outliers son datos que están alejados significativamente de los datos restantes de la muestra. La omisión o eliminación de estos datos puede afectar a los parámetros de la muestra, sobre todo para muestras pequeñas. Todos los procedimientos por tratar outliers, requieren un juicio que involucra consideraciones matemáticas como hidrológicas. El descubrimiento y tratamiento de outliers tanto altos como bajos se describen a continuación y en el diagrama de flujo del apéndice 5.

Si la inclinación de la estación es mayor que +0.4, la prueba para outliers altos se debe de considerar primero. Si la inclinación de las estaciones es menor que -0.4 la prueba para outliers bajos se realiza primero y si la estación está entre ± 0.4 , deben aplicarse pruebas para el outliers altos y bajos antes de eliminar cualquier outliers de la muestra de datos.

La ecuación siguiente se usa para conocer los outliers altos:

$$X_H = \bar{X} + K_N S \quad (3.21)$$

donde:

X_H = Umbral del outliers alto.

\bar{X} = Logaritmo medio de crestas sistemáticas (X 's) excluyendo ceros eventos de diluvio, crestas debajo de la base de la prenda y los outliers descubiertos previamente.

S = Desviación normal de X 's.

K_N = Valor de K , para el tamaño de la muestra N , que se obtiene del apéndice 3.

Si los logaritmos de crestas en una muestra son mayores que X_H en la ecuación (3.21) entonces son considerados outliers altos.

La ecuación siguiente se usa para descubrir outliers bajos:

$$X_L = \bar{X} - K_N S \quad (3.22)$$

donde:

X_L = el umbral del outliers bajo en unidades del leño y los otros términos es como definió para ecuación (3.21).

Si un ajuste para los datos de diluvio históricos se ha hecho previamente, entonces la ecuación siguiente se usa para descubrir outliers bajos:

$$X_L = \tilde{M} - K_H \tilde{S} \quad (3.23)$$

donde:

X_L = Umbral de outliers bajos.

K_H = Valores de K para el periodo calculado \tilde{M} y \tilde{S} obtenida del apéndice 4.

\tilde{M} = Ajuste del logaritmo medio.

\tilde{S} = Ajuste de la desviación normal.

Si los logaritmos de cualquier cresta anual en una muestra están debajo de X_L en las ecuaciones (3.22) o (3.23), entonces son considerados outliers bajos.

Si valores múltiples que no se han identificado como outliers por el procedimiento recomendado, son muy cercanos del valor del umbral, puede ser deseable probar la sensibilidad de los resultados a tratar estos valores como outliers.

3.2.10.- Límites de confianza

En el registro de diluvio anual su puede considerar que fluye a un sitio tomado al azar, una muestra de la población subyacente de datos anuales, puede usarse para estimar la curva de frecuencia de esa población. Si la misma muestra tomada al azar pudiera seleccionarse en un periodo diferente de tiempo, probablemente resultaría una estimación diferente de la curva de frecuencia de población subyacente. Así, una curva de frecuencia de diluvio estimada puede ser sólo una aproximación a la verdadera curva de frecuencia de la población subyacente de datos de diluvio anuales. Para calibrar la exactitud de esta aproximación, se puede construir un intervalo o rango de curvas de frecuencia hipotéticas que con un grado alto de confianza, contiene la curva de frecuencia de población. Tales intervalos se conocen como intervalos de confianza y sus puntos del extremo como límites de confianza.

En este punto se explica cómo construir intervalos de confianza para descargas de diluvio que han especificado probabilidades de excedencia. Con este fin se permite a X_P denotar la población de descarga logarítmica que tiene probabilidad del excedencia P .

Los límites de confianza tienen las siguientes propiedades:

$$\text{Probabilidad } \{U_{p,c}(X) \geq X_p^*\} = c \quad (3.24)$$

$$\text{Probabilidad } \{L_{p,c}(X) \leq X_p^*\} = c \quad (3.25)$$

Las ecuaciones anteriores explican el significado estadístico de los límites de confianza.

Se llaman límites de confianza definidos sobre que la confianza uno con respecto a la del otro limita, porque cada uno de ellos describe un límite o limita en sólo un lado de la descarga de probabilidad de población. Un intervalo de confianza puede formarse de la unión de los dos intervalos, de la siguiente forma:

$$\text{Probabilidad } \{L_{p,c}(X) \leq X_p \leq U_{p,c}(X)\} = 2c - 1 \quad (3.26)$$

Las fórmulas explícitas para los límites de confianza son derivadas especificando la forma general de los límites y simplificando adicionales para analizar las relaciones entre las estadísticas de la muestra y estadísticas de la población. La forma general de los límites de confianza se especifica con las siguientes ecuaciones:

$$U_{p,c}(X) = \bar{X} + S(K_{p,c}^U) \quad (3.27)$$

$$L_{p,c}(X) = \bar{X} + S(K_{p,c}^L) \quad (3.28)$$

donde:

\bar{X} y S son la desviación media y normal logarítmica

$K_{p,c}^U$ y $K_{p,c}^L$ son coeficientes de confianza superior y bajo respectivamente

Los coeficientes de confianza aproximan la t-distribución no-central. La t-variante no-central puede obtenerse en tablas (41, 32), aunque el proceso es embarazoso cuando G_W tiene no-ceros años. Más conveniente es el uso de las fórmulas aproximadas siguientes (32, pp. 2-15), basó en una aproximación de la muestra grande a la t-distribución no-central (42):

$$K_{p,c}^U = \frac{K_{G_w,p} + \sqrt{K_{G_w,p} - ab}}{a} \quad (3.29)$$

$$K_{p,c}^L = \frac{K_{G_w,p} - \sqrt{K_{G_w,p} - ab}}{a} \quad (3.30)$$

donde:

$$a = 1 - \frac{Z_c^2}{2(N-1)} \quad (3.31)$$

$$b = K_{G_w, r}^2 - \frac{Z_c^2}{N} \quad (3.32)$$

y Z_c es la desviación estándar normal con probabilidad acumulativa c (probabilidad de excedencia $1-c$). La longitud del registro sistemático N se juzga para controlar la fiabilidad estadística de la curva de frecuencia estimada y será usado para los límites de confianza interesados, igual a cuando se ha usado información histórica para estimar la curva de frecuencia.

El uso de la ecuación (3.27) a través de (3.28) es ilustrado calculando el 95 % del límite de confianza para $X_{0.01}$, con un diluvio con probabilidad de excedencia de 0.01, cuando la curva de frecuencia estimada tiene desviación media normal logarítmica y asimetría de 3.00, 0.25, y 0.20, respectivamente basado en 50 años de registro sistemático.

Los límites correspondientes en unidades naturales (metros cúbicos por segundo) son 3270 y 5700, estimados con una probabilidad de excedencia de 0.01 es 4150 pies³/seg.

La tabla del apéndice 4 tiene una inclinación de cero y puede usarse para calcular $K_{P,c}^U$ y $K_{P,c}^L$ para los valores seleccionados de P y c cuando la distribución de logaritmos de datos anuales es normal (*i.e.*, $G_w = 0$).

Un ejemplo de cómo utilizar el apéndice 4 para calcular los límites de confianza es de la siguiente manera: se asume el 95 % de los límites de confianza, que en este caso se desea para $X_{0.01}$, el 0.01 diluvio de probabilidad de excedencia para una curva de frecuencia con la desviación media normal logarítmica y asimetría de 3.00, 0.25 y 0.00, respectivamente, basados en 50 años de registro sistemático.

$K_{0.01,0.95}^U = 2.862$. Se encuentra en el apéndice 4 con un nivel de confianza de 0.05, registro sistemático, longitud de 50 y probabilidad de excedencia de 0.01.

$K_{0.01,0.95}^L = 1.936$. Se encuentra en el apéndice 4 con nivel de confianza de 0.95, registro sistemático, longitud 50 y probabilidad de excedencia de 0.01.

Los límites correspondientes en unidades naturales (metros cúbicos por segundo) son 3050 y 5190, estimados para un diluvio de probabilidad de excedencia de 0.01 es 3820 m³/s.

$$U_{0.01,0.95}(X) = 3.00 + 0.25(2.862) = 3.715$$

$$L_{0.01,0.95}(X) = 3.00 + 0.25(1.936) = 3.484$$

$U_{p,c}(X)$ = Límite de confianza superior.

$L_{p,c}(X)$ = Límite de confianza bajo.

P = Probabilidad de excedencia

C = Nivel de confianza

X_p = Población la descarga logarítmica para probabilidad de excedencia P

\bar{X} = el logaritmo medio de flujos de la cresta

S = la desviación normal de logaritmos de descargas de la cresta anuales

$K_{G_w,P}$ = Coordenada expresada en número de desviaciones medias normales para el sesgo (G_w) y probabilidad de excedencia (P)

G_w = Coeficiente oblicuo

$K_{p,c}^U$ = Coeficiente de confianza superior

$K_{p,c}^L$ = Coeficiente de confianza más bajo

N = Longitud del registro sistemático

Z_c = Desviación normal

3.2.11. Datos Históricos

Puede usarse a menudo información de diluvio fuera del registro sistemático para extender el registro de los eventos más grandes a un tiempo del periodo histórico que el del registro sistemático. En semejante situación, las técnicas analíticas siguientes se usan para calcular un dato históricamente ajustado por la distribución Pearson tipo III para la curva de frecuencia.

1. - Se define el periodo históricamente más largo de "H" años. El número "Z" de eventos que se conocen para ser el más grande en el periodo históricamente más largo "H" un peso de $(H-Z) / (N+L)$ en la asunción que su distribución es representativa del $(H-Z)$ los años restantes del periodo históricamente más largo.

2. - Los cálculos pueden ser hechos directamente aplicando los pesos a cada datos de los años individuales con las ecuaciones (3.33), (3.34), (3.35) y (3.36). Si se han calculado estadísticos previamente para el registro continuo actual, ellos pueden ajustarse para dar los valores históricamente ajustados equivalentes por medio de las ecuaciones (3.33), (3.37), (3.38) y (3.39).

3. - La curva de frecuencia históricamente ajustada se dibuja en papel logarítmico de probabilidad a través de puntos establecidos por medio de la ecuación (3.40). Los eventos de diluvio individuales también deben de ser trazados. El trazado históricamente ajustado de las posiciones para los eventos de diluvio individuales son calculadas con la ecuación (3.43) en que el número de orden históricamente ajustado de cada evento "m" se calcula con las ecuaciones (3.41) y (3.42).

$$W = \frac{H - Z}{N + L} \quad (3.33)$$

$$\bar{M} = \frac{W \sum X_z}{H - WL} \quad (3.34)$$

$$\bar{S}^2 = \frac{W \sum (X - \bar{M})^2 + \sum (X_z - \bar{M})^2}{(H - WL - 1)} \quad (3.35)$$

$$\bar{G} = \frac{H - WL}{(H - WL - 1)(H - WL - 2)} \left[\frac{W \sum (X - \bar{M})^3 + \sum (X_z - \bar{M})^3}{\bar{S}^3} \right] \quad (3.36)$$

$$\bar{M} = \frac{WNM + \sum X_z}{H - WL} \quad (3.37)$$

$$\bar{S}^2 = \frac{W(N - 1)S^2 + WN(M - \bar{M})^2 + \sum (X_z - \bar{M})^2}{(H - WL - 1)} \quad (3.38)$$

$$\bar{G} = \frac{H - WL}{(H - WL - 1)(H - WL - 2)\bar{S}^3} \left[\frac{W(N - 1)(N - 2S^3G)}{N} + 3W(N - 1)(M - \bar{M})S^2 \right. \\ \left. + WN(M - \bar{M})^3 + \sum (X_z - \bar{M})^3 \right] \quad (3.39)$$

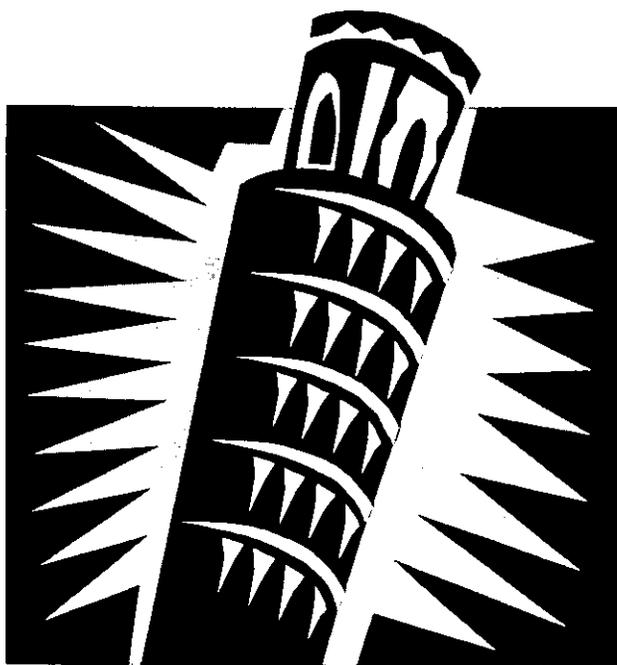
$$\text{Log } Q = \bar{M} + K\bar{S} \quad (3.40)$$

$$\bar{m} = E; \text{ Si: } 1 \leq E \leq Z \quad (3.41)$$

$$\bar{m} = WE - (W-1)(Z+0.5); \text{ Si: } (Z+1) \leq E \leq (Z+N+L) \quad (3.42)$$

$$P\bar{P} = \frac{\bar{m} - a}{H + 1 - 2a} 100 \quad (3.43)$$

CAPITULO IV



CAPITULO CUATRO

Datos disponibles

Años	Ejemplo 1	Ejemplo 2	Ejemplo 3	Ejemplo 4
1929			8750	
1930			15500	
1931			4060	
1932			----	4260
1933			----	345
1934			----	516
1935		1460	----	1320
1936		4050	22000	1200
1937		3570	----	2180
1938		2060	----	3230
1939		1300	6300	115
1940		1390	3130	3440
1941		1720	4160	3070
1942		6280	6700	1880
1943		1360	22400	6450
1944		7440	3880	1290
1945	2290	5320	8050	5970
1946	1470	1400	4020	782
1947	2220	3240	1600	0
1948	2970	2710	4460	0
1949	3020	4520	4230	335
1950	1210	4840	3010	175
1951	2490	8320	9150	2920
1952	3170	13900	5100	3660
1953	3220	71500	9820	147
1954	1760	6250	6200	0
1955	8800	2260	10700	16
1956	8280	318	3880	5620
1957	1310	1330	3420	1440
1958	2500	970	3240	10200
1959	1960	1920	6800	5380
1960	2140	15100	3740	448
1961	4340	2870	4700	0
1962	3060	20600	4380	1740
1963	1780	3810	5190	8300
1964	1380	726	3960	156
1965	980	7500	5600	560
1966	1040	7170	4670	128
1967	1580	2000	7080	4200
1968	3630	829	4640	0
1969		17300	536	5080
1970		4740	6680	1010
1971		13400	8360	584
1972		2940	18700	0
1973		5660	5210	1510

4.1. APLICACION DE LA DISTRIBUCION LOG - PEARSON TIPO III

a.- Descripción de la Estación

Fishkill Creek, New York

Lat: 41° 30 ' 42", Long:73° 56' 58 "

Datos anuales disponibles: 1945 - 1968

b.- Procedimientos de Calculo

Se calculan los logaritmos, cuadrados y cubos de los datos.

Años	Gastos Máximos Anuales	Logaritmos	X²	X³
1945	2290	3.3598	11.2885	37.9275
1946	1470	3.1673	10.0319	31.7742
1947	2220	3.3464	11.1981	37.4727
1948	2970	3.4728	12.0600	41.8816
1949	3020	3.4800	12.1104	42.1444
1950	1210	3.0828	9.5036	29.2975
1951	2490	3.3962	11.5342	39.1723
1952	3170	3.5011	12.2574	42.9139
1953	3220	3.5079	12.3051	43.1644
1954	1760	3.2455	10.5334	34.1861
1955	8800	3.9445	15.5589	61.3720
1956	8280	3.9180	15.3510	60.1455
1957	1310	3.1173	9.7174	30.2917
1958	2500	3.3979	11.5460	39.2326
1959	1960	3.2923	10.8390	35.6846
1960	2140	3.3304	11.0917	36.9398
1961	4340	3.6375	13.2313	48.1288
1962	3060	3.4857	12.1503	42.3524
1963	1780	3.2504	10.5652	34.3414
1964	1380	3.1399	9.8588	30.9556
1965	980	2.9912	8.9474	26.7638
1966	1040	3.0170	9.1025	27.4625
1967	1580	3.1987	10.2314	32.7268
1968	3630	3.5599	12.6729	45.1145
N = 24	Σ	80.8404	273.6863	931.4467

Calculo de estadísticos a través de la ecuación (3.2)

$$\bar{X} = \frac{\sum X}{N} = \frac{80.84043}{24} = 3.3684$$

Calculo de la desviación estándar a través de la ecuación (3.4)

$$S = \left[\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N-1} \right]^{0.5} = \left[\frac{273.68646 - \frac{(80.84043)^2}{24}}{23} \right]^{0.5} = 0.2456$$

Calculo del coeficiente oblicuo por medio de la ecuación (3.6)

$$G = \frac{N^2(\sum X^3) - 3N(\sum X)(\sum X^2) + 2(\sum X)^3}{N(N-1)(N-2)S^3}$$

$$G = \frac{(24)^2(931.44732) - 3(24)(80.84043)(273.68646) + 2(80.84043)^3}{24(23-1)(24-2)(0.24561)^3} = 0.7300$$

Paso 3. – Comprobación para Outliers:

$$X_H = \bar{X} + K_N S = 3.3684 + 2.467(0.2456) = 3.9743$$

$$Q_H = \text{anti log}(3.9743) = 9425$$

Ningún dato de la estación es mayor al dato obtenido, por lo tanto no se detecto ningún outlier alto. Después se aplica la prueba para determinar si existen outliers bajos. Se utiliza el mismo valor de K_N en la formula (3.22) para calcular outliers bajos (Q_L):

$$X_L = \bar{X} - K_N S = 3.3684 - 2.467(0.2456) = 2.7625$$

$$Q_L = \text{anti log}(2.7625) = 579$$

No existen valores por debajo del valor obtenido, por lo cual tampoco existen outliers bajos para esta estación. Para este ejemplo se tomara una inclinación generalizada de 0.6. En práctica real una inclinación generalizada puede obtenerse de otras fuentes o de un estudio especial para la región. Una inclinación es calculada por medio de la ecuación (3.7). El error medio cuadrado de la inclinación de la estación puede encontrarse en el apéndice 2 o puede ser calculado a través de la ecuación (3.8)

$$MSE_G \approx 10^{\left[A - B \left[\log_{10} \left(\frac{N}{10} \right) \right] \right]}$$

donde:

$$A = -0.33 + 0.08 |G| = -0.33 + 0.08 (0.730) = -0.2716$$

$$B = 0.94 - 0.26 |G| = 0.94 - 0.26 (0.730) = 0.7502$$

$$MSE_G \approx 10^{\left[-0.2716 - 0.7502 \left[\log_{10} (2.4) \right] \right]} \approx 0.277$$

El error medio cuadrado de la inclinación generalizada es 0.302.

Calculo la de inclinación a través de la ecuación (3.7)

$$G_w = \frac{MSE_{\bar{G}}(G) + MSE_G(\bar{G})}{MSE_{\bar{G}} + MSE_G} = \frac{0.302(0.7300) + 0.277(0.6)}{0.579} = 0.6678$$

$G_w = 0.7$ (redondeó a más cercano décimo)

Paso 4.- Calculo de las coordenadas de la curva de frecuencia.

El valor de K valora para un coeficiente oblicuo de 0.7 se obtiene del apéndice 3. El calculo para una probabilidad de excedencia de 0.01 se calcula con la ecuación (3.1)

$$\log Q = \bar{X} + KS = 3.3684 + 2.82359(0.2456) = 4.0619$$

$$Q = 11500 \text{ pies}^3/\text{seg}$$

La descarga valorada en este calculo y aquellos obtenidos en la tabla 4.1 se redondean a tres cifras significativas. Las curvas de frecuencia se muestran en la figura 4.1.

Tabla 4.1 Calculo de coordenadas de curva de frecuencia

P	$K_{G_w, P}$ para $G_w = 0.7$	$\text{Log } Q$	Q (pies ³ /seg)
0.99	-1.80621	2.9247	841
0.90	-1.18347	3.0777	1200
0.50	-0.11578	3.3399	2190
0.10	1.33294	3.6957	4960
0.05	1.81864	3.8150	6530
0.02	2.40670	3.9595	9110
0.01	2.82359	4.0619	11500
0.005	3.22281	4.1599	14500
0.002	3.72957	4.2844	19200

Paso 5 .- Calculo de los limites de confianza.

Los límites de confianza superiores y más bajos para los niveles de importancia de 0.05 y 0.95 son calculados por los procedimientos mencionados en el punto 10 del capítulo 3. Se han seleccionado nueve probabilidades de excedencia (P) para definir las curvas del límite de confianza. Los cálculos para dos puntos en la curva a una probabilidad de excedencia de 0.99 se mencionan a continuación.

Se utilizan las ecuaciones del punto 10 calculando un valor aproximado de $K_{P,C}$. La desviación normal Z_C se encuentra en el apéndice 4 con un coeficiente oblicuo de cero. Para un nivel de confianza de 0.05, $Z_C = 1.64485$. La desviación Log Pearson III, $K_{G_w,P}$ se encuentra en el apéndice 4 basado en el coeficiente oblicuo apropiado. Para una probabilidad de excedencia de 0.99 y el coeficiente oblicuo de 0.7, $K_{G_w,P} = -1.80621$.

$$a = 1 - \frac{Z_c^2}{2(N-1)} = 1 - \frac{(1.64485)^2}{2(24-1)} = 0.9412$$

$$b = K_{G_w,P}^2 - \frac{Z_c^2}{N} = (-1.80621)^2 - \frac{(1.64485)^2}{24} = 3.1497$$

$$K_{P,C}^U = \frac{K_{G_w,P} + \sqrt{K_{G_w,P}^2 - ab}}{a} = \frac{-1.80621 + \sqrt{(-1.80621)^2 - (0.9412)(3.1497)}}{0.9412}$$

$$= \frac{-1.80621 + 0.5459}{0.9412} = -1.3392$$

El valor de la descarga es:

$$\text{Log } Q = 3.3684 + (-1.3392)(0.2456) = 3.0395$$

$$Q = 1100 \text{ pies}^3/\text{seg}$$

Para el coeficiente de confianza más bajo:

$$K_{P,C}^L = \frac{K_{G_w,P} - \sqrt{K_{G_w,P}^2 - ab}}{a} = \frac{-1.80621 - 0.5458}{0.9412} = -2.4989$$

El valor de la descarga es:

$$\text{Log } Q = 3.3684 + (-2.4989)(0.2456) = 2.7546$$

$$Q = 568 \text{ pies}^3/\text{seg}$$

Se dan los cálculos que muestran la derivación de los límites de confianza superiores e inferiores en la tabla 4.1.

Tabla 4.2. - Límites de confianza

P	$K_{G_w, P}$ para $G_w = 0.7$	Limite de curva superior 0.05			Limite de curva inferior 0.05		
		$K_{P,c}^U$	Log Q	Q (pies ³ /s)	$K_{P,c}^L$	log Ω	Ω (pies ³ /s)
0.99	-1.80621	-1.3392	3.0395	1100	-2.4989	2.7546	568
0.90	-1.18347	-0.7962	3.1728	1490	-1.7187	2.9462	884
0.50	-0.11578	0.2244	3.4235	2650	-0.4704	3.2528	1790
0.10	1.33294	1.9038	3.8359	6850	0.9286	3.5964	3950
0.05	1.81864	2.5149	3.9860	9680	1.3497	3.6998	5010
0.02	2.40670	3.2673	4.1708	14800	1.8469	3.8220	6640
0.01	2.82359	3.8058	4.3031	20100	2.1943	3.9073	8080
0.005	3.22281	4.3239	4.4303	26900	2.5245	3.9884	9740
0.002	3.72957	4.9841	4.5925	39100	2.9412	4.0907	12300

Paso 6. – Calculo del ajuste de probabilidad esperada.

La probabilidad esperada es determinada de la siguiente forma:

Tabla 4.3. basada en N - 1 de 23.

Tabla de P_N contra P_∞ . Para uso con muestras deducidas de una Distribución Normal

$N-1 \backslash P_\infty$	0.5	0.3	0.1	0.05	0.01	0.001	0.0001
1	0.5	0.372	0.243	0.204	0.154	0.121	0.102
2	0.5	0.347	0.193	0.146	0.09	0.057	0.043
3	0.5	0.336	0.169	0.119	0.064	0.035	0.023
4	0.5	0.33	0.154	0.104	0.05	0.024	0.0137
5	0.5	0.325	0.146	0.094	0.042	0.0179	0.0092
6	0.5	0.322	0.138	0.088	0.036	0.0138	0.006
7	0.5	0.319	0.135	0.083	0.032	0.0113	0.005
8	0.5	0.317	0.131	0.079	0.029	0.0094	0.0039
9	0.5	0.316	0.127	0.076	0.027	0.0082	0.0031
10	0.5	0.315	0.125	0.073	0.025	0.0072	0.0025
11	0.5	0.314	0.123	0.071	0.023	0.0064	0.0021
12	0.5	0.313	0.121	0.069	0.022	0.0058	0.0018
13	0.5	0.312	0.119	0.068	0.021	0.0052	0.0016
14	0.5	0.311	0.118	0.067	0.02	0.0048	0.0014
15	0.5	0.311	0.117	0.066	0.0196	0.0045	0.0013
16	0.5	0.31	0.116	0.065	0.019	0.0042	0.0012
17	0.5	0.31	0.115	0.064	0.0184	0.004	0.0011
18	0.5	0.309	0.114	0.063	0.0179	0.0038	0.001
19	0.5	0.309	0.113	0.062	0.0174	0.0036	0.00091
20	0.5	0.308	0.113	0.062	0.017	0.0034	0.00084
21	0.5	0.308	0.112	0.061	0.0167	0.0033	0.00078
22	0.5	0.308	0.111	0.061	0.0163	0.0031	0.00073
23	0.5	0.307	0.111	0.06	0.0161	0.003	0.00068
24	0.5	0.307	0.11	0.06	0.0158	0.0029	0.00064
25	0.5	0.307	0.11	0.059	0.0155	0.0028	0.0006
26	0.5	0.306	0.109	0.059	0.0153	0.0027	0.00057
27	0.5	0.306	0.109	0.059	0.0151	0.0026	0.00054
28	0.5	0.306	0.109	0.058	0.0149	0.0026	0.00051
29	0.5	0.306	0.108	0.058	0.0147	0.0025	0.00049
30	0.5	0.306	0.108	0.058	0.0145	0.0024	0.00046
40	0.5	0.304	0.106	0.056	0.0133	0.002	0.00034
60	0.5	0.303	0.104	0.054	0.0122	0.0016	0.00025
120	0.5	0.302	0.102	0.052	0.0111	0.0013	0.00017
∞	0.5	0.3	0.1	0.05	0.01	0.001	0.0001

Nota: Valores aproximados de P_N , utilizables con la Distribución Log Pearson tipo III, que contengan coeficientes oblicuos pequeños

Tabla 4.4. Ajuste de probabilidad esperada.

P	Q (pies³/seg)	Probabilidad esperada
0.99	841	0.9839
0.90	1200	0.8890
0.50	2190	0.5000
0.10	4960	0.1110
0.05	6530	0.0600
0.02	9110	0.0280
0.01	11500	0.0161
0.005	14500	0.0095
0.002	19200	0.0049

La curva de frecuencia ajustada para probabilidad esperada se muestra en la figura 4-1.

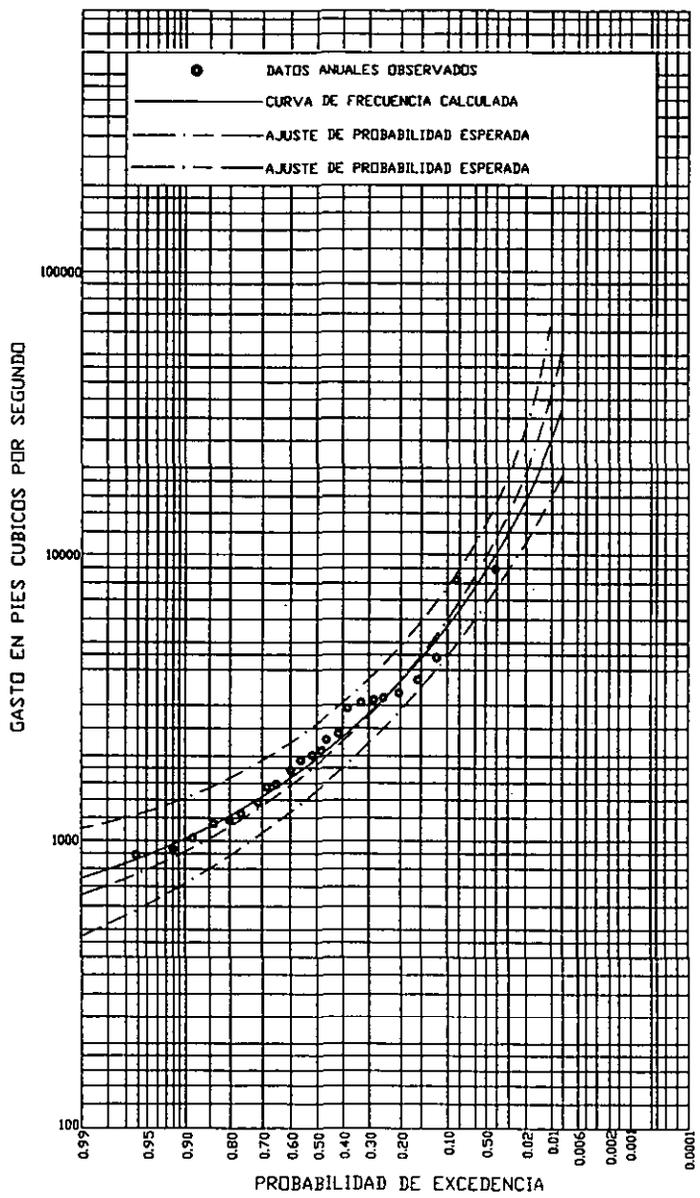


FIGURA 3.1. CURVAS DE FRECUENCIA

4.2. AJUSTE PARA OUTLIERS ALTOS

a.- Descripción de la Estación

Floyd River at James, Iowa.

Lat: 42° 34' 30", Long: 96° 18' 45"

Datos anuales disponibles: 1935 - 1973

b.- Procedimientos de Calculo

Paso 1. – Calculo de estadísticos.

Los cálculos detallados para el registro sistemático se han omitido 1955 - 1999; los resultados de los cálculos son:

Logaritmo medio 3.5553

Desviación Estándar 0.4642

Coficiente oblicuo de leños 0.3566

Años 39

A estas alturas, el analista puede desear ver la curva de frecuencia preliminar basada en las estadísticas del registro sistemático. Figura 12-2 es la curva de frecuencia preliminar basada en la desviación media normal calculada y una inclinación de 0.1 (basado en una inclinación generalizada de -0.3).

Paso 2. Comprobación para Outliers.

La inclinación de la estación está entre ± 0.4 ; por consiguiente, las pruebas para outliers altos y bajos son basadas en las estadísticas del registro sistemático. Del apéndice 3, K_N para un tamaño de la muestra de 39 es 2.671.

El umbral del outliers alto (Q_H) es calculado a través de la ecuación (3.21):

$$X_x = \bar{X} + K_N S = 3.5553 + 2.671(0.4642) = 4.7952$$

$$Q_H = \text{el antilog}(4.7952) = 62400 \text{ pies}^3/\text{seg}$$

En 1953 el valor de 71500 pies³/seg excede este valor. Por consiguiente, este valor se considera como un outlier alto. Si tal información no estuviera disponible, las comparaciones con estaciones cercanas puede adoptarse.

El outlier bajo del umbral (Q_L) es calculado a través de la ecuación (3.22):

$$X_L = \bar{X} - K_N S$$

$$X_L = 3.5553 - 2.671(0.4642) = 2.3154$$

$$Q_L = \text{el antilog}(2.3154) = 207 \text{ pies}^3/\text{seg}$$

No hay valores debajo de este valor del umbral.

Paso 3.- Recalculo de estadísticos

El valor de 1953 se anula y el recalculo de los estadísticos restantes es el siguiente:

Logaritmo medio 3.5212

Desviación normal de leños 0.4177

Coefficiente oblicuo de leños -0.0949

Años 38

Paso 4.- Se utilizan los datos históricos para modificar estadísticos y el trazado de posiciones.

La aplicación de los procedimientos del punto 11 (datos históricos) permite ajustar los estadísticos calculadas por la incorporación de los datos históricos.

1. - El periodo histórico (H) es 1892-1973 o 82 años y el número de valores bajos excluidos (L) es cero.
2. - El periodo sistemático (N) es 1935-1973 (con 1953 anulado) o 38 años.
3. - Hay un evento (Z) conocido para ser el más grande en 82 años.
4. - Calculo de (W) a través de la ecuación (3.3):

$$W = \frac{H - Z}{N + L} = \frac{82 - 1}{38 + 0} = 2.13158$$

Calculo ajustado medio a través de la ecuación (3.37):

$$\tilde{M} = \frac{WNM + \sum X_z}{H - WL}$$

$$\bar{X} = M = 3.5212$$

$$WNM = 285.2173$$

$$\sum X_z = \frac{4.8543}{290.0716}$$

$$\tilde{M} = \frac{290.0716}{(82 - 0)} = 3.5375$$

Calculo de la desviación normal ajustada a través de la ecuación (3.38):

$$\tilde{S}^2 = \frac{W(N-1)S^2 + WN(M - \tilde{M}) + \sum (X_z - \tilde{M})^2}{H - WL - 1}$$

$$S = 0.4177$$

$$W(N-1)S^2 = 13.7604$$

$$WN(M - \tilde{M})^2 = 13.7604$$

$$\frac{\sum (X_z - \tilde{M})^2}{15.5159} = 1.7340$$

$$\tilde{S}^2 = \frac{15.5159}{82 - 0 - 1} = 0.1916$$

$$\tilde{S} = 0.4377$$

$$\tilde{G} = \frac{H - WL}{(H - WL - 1)(H - WL - 2)\tilde{S}^3} \left[\frac{W(N-1)(N-2)S^3G}{N} + 3W(N-1)(M - \tilde{M})S^2 + WN(M - \tilde{M})^3 + \sum (X_z - \tilde{M})^3 \right]$$

$$\tilde{G} = -0.0949$$

$$\frac{W(N-1)(N-2)S^3G}{N} = -0.5168$$

$$3W(N-1)(M - \tilde{M})S^2 = -0.6729$$

$$WN(M - \tilde{M})^3 = -0.0004$$

$$\frac{\sum (X_z - \tilde{M})^3}{1.0932} = 2.2833$$

$$\frac{H}{(H - WL - 1)(H - WL - 2)\tilde{S}^3} = 0.1509$$

$$G = 0.1509 (1.0932) = 0.1650$$

Calculo de la inclinación

Para este ejemplo se tomara una inclinación generalizada de -0.3. El error medio cuadrado será de 0.302. Del apéndice 2, el error medio cuadrado de la inclinación de la estación basado en H de 82 años es 0.073. La inclinación es calculada por medio de la ecuación (3.7):

$$G_w = \frac{0.302(0.1650) + 0.073(-0.3)}{0.302 + 0.073} = 0.0745$$

$$G_w = 0.1 \text{ (redondeó a más cercano décimo)}$$

Paso 5. - Calculo ajustado trazando posiciones para los datos históricos.

Para el evento más grande se utiliza la ecuación (3.41)

$$\tilde{m}_1 = 1$$

Para los eventos subsiguientes se utiliza la ecuación (3.42)

$$\tilde{m} = WE - (W - 1)(Z + 0.5) = 2.1316(2) - (2.1316 - 1)(1 + 0.5) = 2.5658$$

Para esta Distribución $a = 0$; por consiguiente, a través de la ecuación (3.43).

$$P\tilde{P} = \frac{\tilde{m}}{H + 1} (100)$$

$$P\tilde{P}_1 = \frac{1}{82 + 1} (100) = 1.20$$

$$P\tilde{P}_2 = \frac{2.5658}{83} (100) = 3.09 \tag{4.1}$$

Las probabilidades de excedencia son calculadas dividiendo los valores obtenidos de la ecuación (4.1) entre 100.

$$\frac{3.09}{100} = 0.0309$$

Tabla 4.5. Trazo de posiciones

Años	Q	W	Trazo de posición			
			Número de evento	Orden	Oportunidad de porcentaje	Probabilidad de excedencia
			E	m	$\frac{E}{m}$	$\frac{E^2}{m}$
1953	71500	2.1316	1	1.0000	1.20	0.0120
1962	20600	2.1316	2	2.5658	3.09	0.0309
1969	17300	2.1316	3	4.6974	5.66	0.0566
1960	15100	2.1316	4	6.8290	8.23	0.0823
1952	13900	2.1316	5	8.9606	10.80	0.1080
1971	13400	2.1316	6	11.0922	13.36	0.1336
1951	8320	2.1316	7	13.2238	15.93	0.1593
1965	7500	2.1316	8	15.3554	18.50	0.1850
1944	7440	2.1316	9	17.4870	21.07	0.2107
1966	7170	2.1316	10	19.6186	23.64	0.2364

Paso 6. – Calculo de la curva de frecuencia.

Tabla 4.6. Calculo de coordenadas de curva de frecuencia.

P	$K_{G_w, P}$ para $G_w = 0.1$	Log Q	Q (pies ³ /seg)
0.99	-2.25258	2.5515	356
0.90	-1.27037	2.9815	958
0.50	-0.01662	3.5302	3390
0.10	1.29178	4.1029	12700
0.05	1.67279	4.2697	18600
0.02	2.10697	4.4597	28800
0.01	2.39961	4.5878	38700
0.005	2.66965	4.7060	50800
0.002	2.99978	4.8504	70900

La curva de frecuencia final se muestra en la figura 4.3.

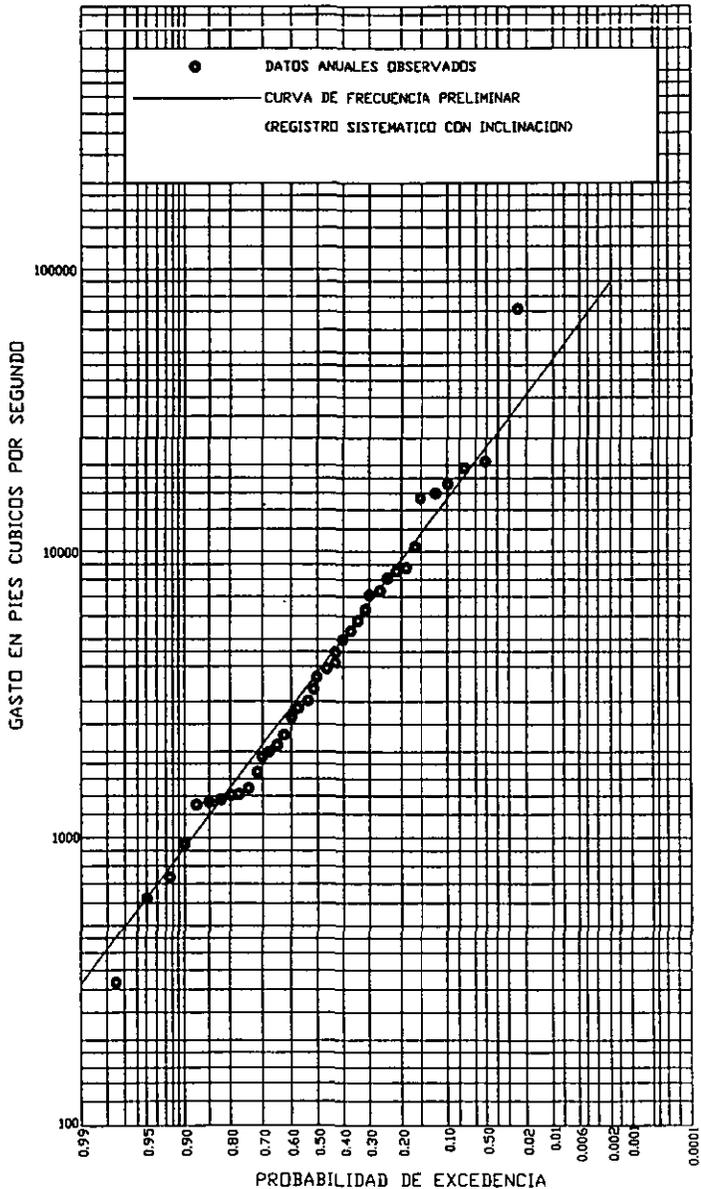


FIGURA 3.2. CURVAS DE FRECUENCIA PRELIMINAR

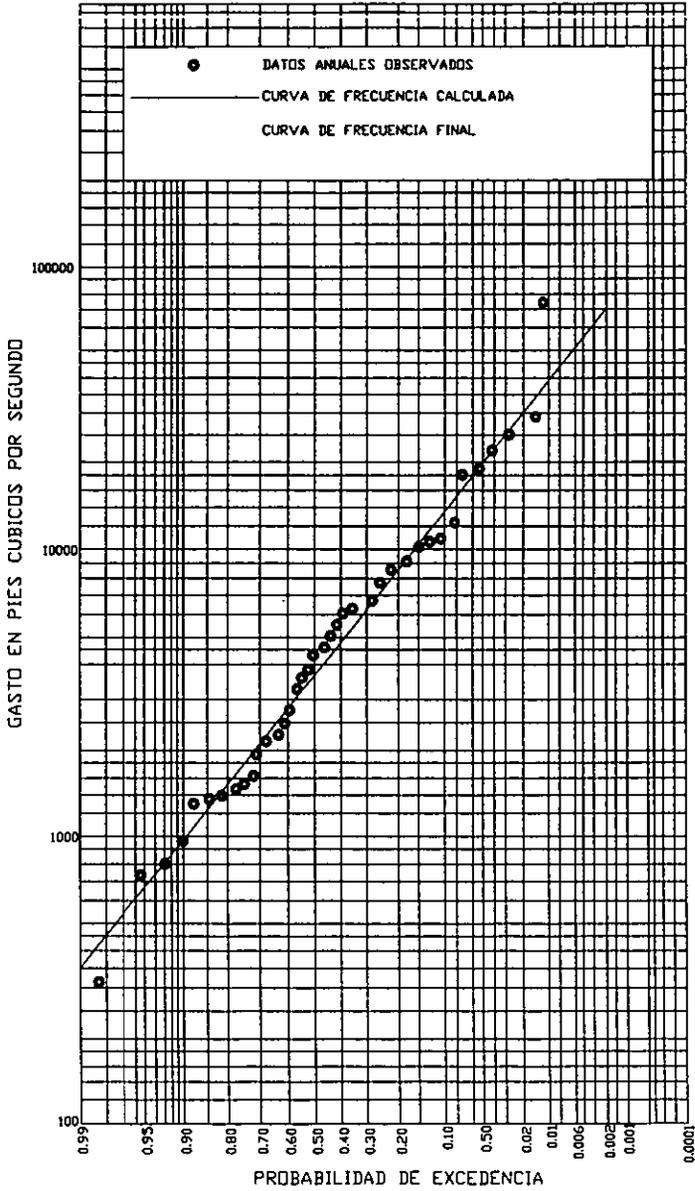


FIGURA 3.3. CURVA DE FRECUENCIA FINAL

4.3. Prueba y ajuste para outliers bajos

a.- Descripción de la Estación

Jones Springs, West Virginia

Lat: 39° 30' 43", Long: 78° 02' 15"

Datos anuales Disponible: 1929-1931, 1939-1973,

b. Procedimientos de Calculo

Paso 1.- Calculo de los estadísticos del registro sistemático.

Los cálculos detallados se han omitido; los resultados de los cálculos son:

Logaritmo medio 3.7220

Desviación normal de leños 0.2804

Coefficiente oblicuo de leños -0.7311

Años 38

La figura 4.4 es la curva de frecuencia preliminar basada en la desviación media normal calculada y una inclinación de -0.2.

Paso 2.- Comprobación para outliers.

Como los coeficientes oblicuos calculados son menores de -0.4, la prueba por descubrir posible outliers bajo se realiza primero. Del Apéndice 3 se obtiene el valor de K_N para un tamaño de la muestra de 38 y su valor es de 2.661.

El umbral de outliers bajos es calculado a través de la ecuación (3.22):

$$X_L = \bar{X} - K_N S = 3.7220 - 2.661(0.2804) = 2.9759$$

$$Q_L = \text{anti log}(2.9759) = 946 \text{ pies}^3/\text{seg}$$

El valor de 1969 es de 536 está debajo del valor del umbral de 946 y se tratará como outlier bajo.

Se anula el outlier bajo y se recalculan los estadísticos.

Logaritmo medio 3.7488

Desviación normal de leños 0.2296

Coefficiente oblicuo de leños 0.6311

Años 37

Paso 4. – Comprobación para outliers altos.

El outlier alto se calcula para 22760 basado en los procedimientos del paso 3 y la muestra es de 37 eventos. Ningún evento grabado excede el valor del umbral.

Paso 5. - Calculo y ajuste de la curva de frecuencia condicional.

Una curva de frecuencia condicional se calcula basada en las estadísticas del paso 3 y entonces es modificada por el ajuste de probabilidad condicional.

El coeficiente oblicuo ha sido redondeado a 0.6 para facilitar el calculo. La proporción de ajuste calculada de la ecuación (3.9) es:

$$\tilde{P} = \frac{N}{n} = \frac{37}{38} = 0.9737$$

Tabla 4.7. Calculo de coordenadas de curva de frecuencia condicionales.

P_d	$K_{G.P.}$ para $G = 0.6$	$\log Q$	Q (pies ³ /seg)	Ajuste de la probabilidad de excelencia
0.99	-1.88029	3.3171	2080	0.9639
0.90	-1.20028	3.4732	2970	0.8760
0.50	-0.09945	3.7260	5320	0.4870
0.10	1.32850	4.0538	11300	0.0970
0.05	1.79701	4.1614	14500	0.0490
0.02	2.35931	4.2905	19500	0.0195
0.01	2.75714	4.3814	24100	0.0097
0.005	3.13232	4.4680	29400	0.0049
0.002	3.60872	4.5774	37800	0.0019

La curva de frecuencia condicional se muestra en la figura 4.5.

Paso 6.- Calculo de los estadísticos sintéticos.

Los estadísticos de la curva del ajuste de frecuencia son desconocidos. El uso de sintéticos proporciona una curva de frecuencia con una distribución Log - Pearson tipo III. Primero se determinan las $Q_{0.01}$, $Q_{0.10}$, y $Q_{0.50}$ descargas de la curva ajustada en la figura 4.5.

$$Q_{0.01} = 23880 \text{ pies}^3/\text{seg}$$

$$Q_{0.10} = 11210 \text{ pies}^3/\text{seg}$$

$$Q_{0.50} = 5230 \text{ pies}^3/\text{seg}$$

Calculo del coeficiente oblicuo sintético a través de la ecuación (3.14).

$$G_s = -2.50 + 3.12 \frac{\log\left(\frac{Q_{0.01}}{Q_{0.10}}\right)}{\log\left(\frac{Q_{0.10}}{Q_{0.50}}\right)} = -2.50 + 3.12 \frac{\log\left(\frac{23880}{11210}\right)}{\log\left(\frac{11210}{5230}\right)} = 0.5948$$

Se calcula la desviación normal sintética a través de la ecuación (3.15).

$$S_s = \frac{\log\left(\frac{Q_{0.01}}{Q_{0.50}}\right)}{(K_{0.01} - K_{0.50})} = \frac{\log\left(\frac{23880}{5230}\right)}{[2.75514 - (-0.09945)]} = 0.2310$$

Calculo del sintético medio a través de la ecuación (3.16).

$$\bar{X} = \log(Q_{0.50}) - K_{0.50}(S_s) = \log(5230) - (-0.09945)(0.2310) = 3.7415$$

Paso 7. - Calculo del coeficiente oblicuo.

El error medio cuadrado de la inclinación de la estación, se ingresa al apéndice 2 y se obtiene el valor de 0.183 basado en $n = 38$ y usando G_s para G

$$G_w = \frac{0.302(0.5948) + 0.183(0.5)}{0.302 + 0.183} = 0.5590$$

$G_w = 0.6$ (redondeó al termino más cercano)

Paso 8.- Calculo de la curva de frecuencia final.

Tabla 4.8. Calculo de coordenadas de la curva de frecuencia

P	$K_{G_w, P}$ para $G_w = 0.6$	log Q	Q (pies³/seg)
0.99	-1.88029	3.3072	2030
0.90	-1.20028	3.4642	2910
0.50	-0.09945	3.7185	5230
0.10	1.32850	4.0484	11200
0.05	1.79701	4.1566	14300
0.02	2.35931	4.2865	19300
0.01	2.75514	4.3780	23900
0.005	3.13232	4.4651	29200
0.002	3.60872	4.5751	37600

La curva de frecuencia final se indica en la figura 4.6.

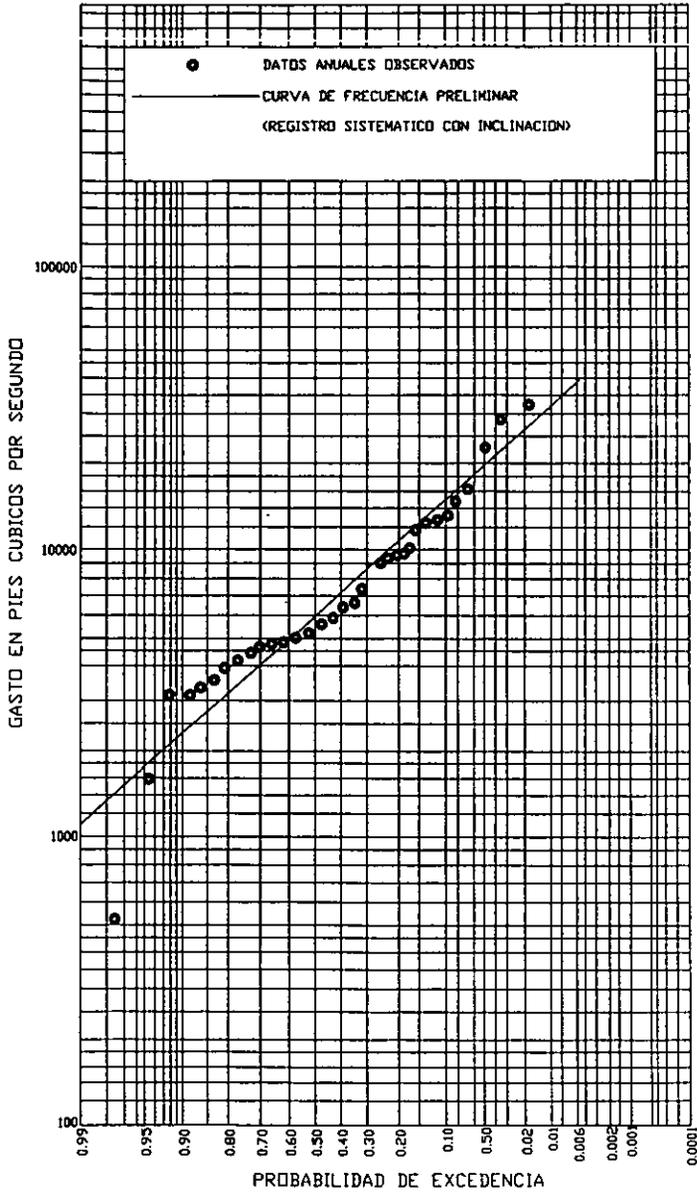


FIGURA 3.4. CURVA DE FRECUENCIA PRELIMINAR

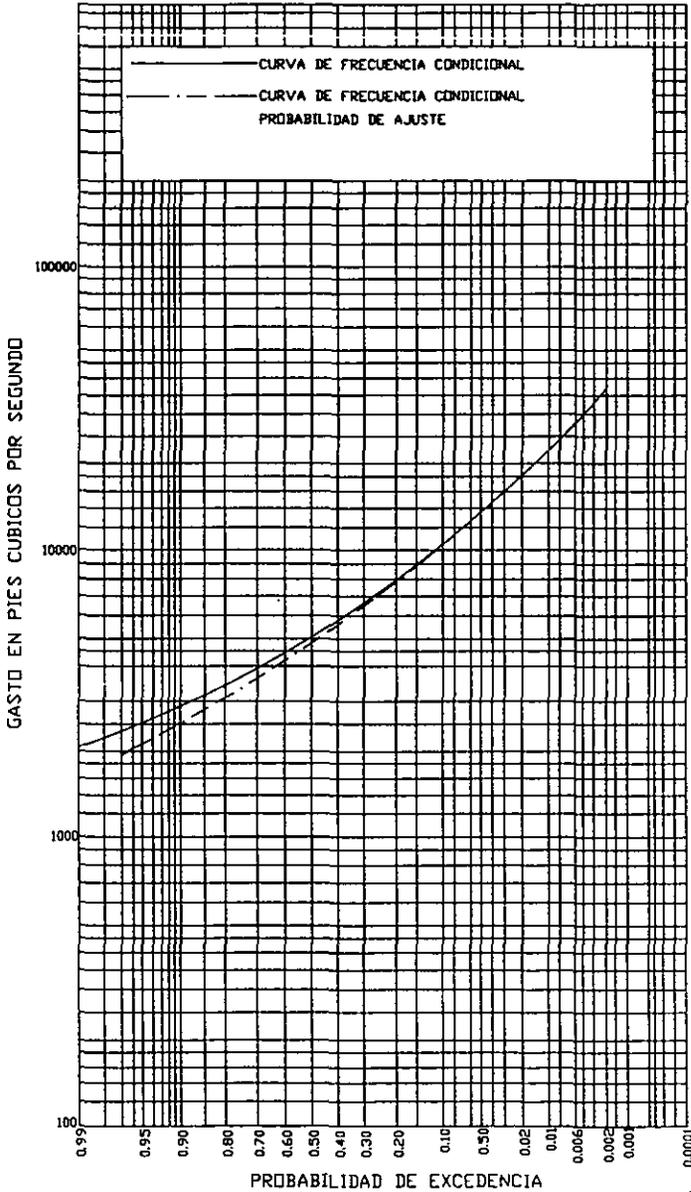


FIGURA 3.5. CURVA DE FRECUENCIA AJUSTADA

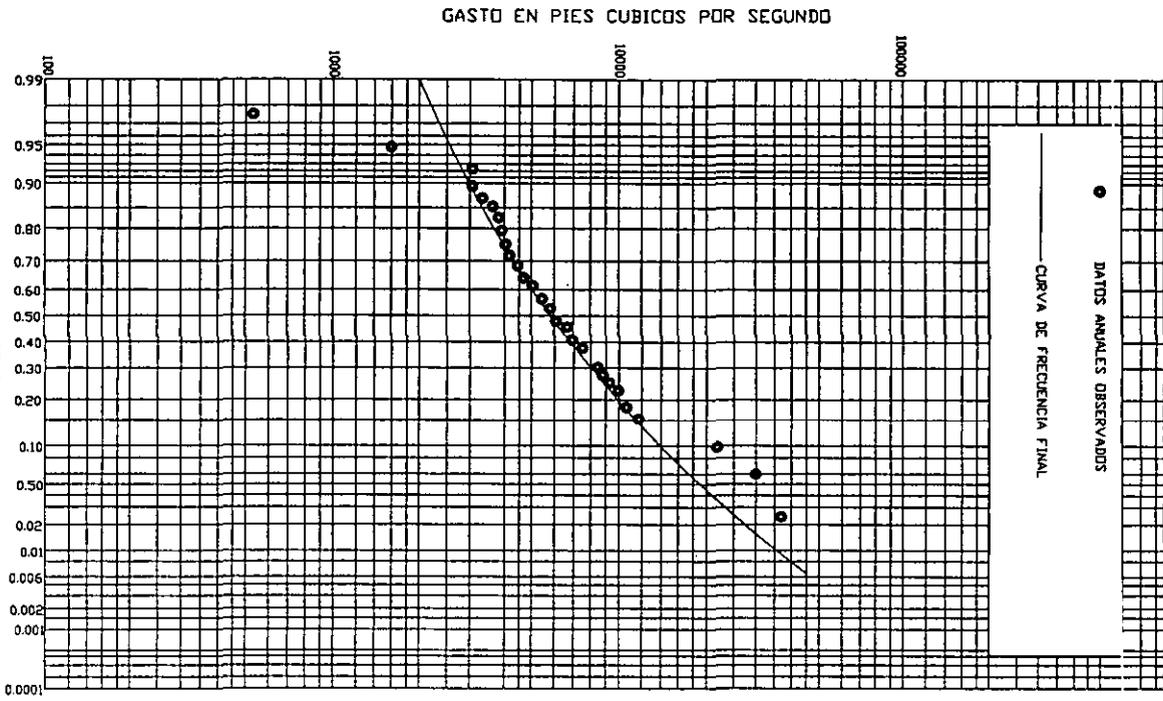


FIGURA 3.6. CURVA DE FRECUENCIA FINAL

4.4. AJUSTANDO DURANTE AÑOS CON DILUVIO DE MAGNITUD CERO

a.- Descripción de la Estación

Orestimba Creek near Newman, California

Lat: 37° 19' 01", Long: 121° 07' 39"

Datos anuales disponibles: 1932 - 1973

b.- Procedimiento de calculo

Paso 1.- Se eliminan los años que contengan eventos con cero.

Hay 6 años con eventos de cero y quedan 36 eventos.

Paso 2.- Calculo de los estadísticos eliminando los datos que tengan magnitud cero.

Logaritmo medio 3.0786

Desviación normal de leños 0.6443

Coefficiente oblicuo de leños -0.8360

Años 36 (número de eventos que no tienen magnitud cero)

Paso 3.- Curva de frecuencia condicional para outliers.

Como los coeficientes oblicuos calculados son menores de -0.4 , la prueba para descubrir outliers bajos se realiza primero, con los registros de los 36 años, se realizan los cálculos correspondientes para detectar outliers bajos, el umbral de estos outliers es 23.9 pies³/seg, el evento de 1955 esta debajo de este valor; por lo cual será tomado como un outlier bajo y el recalcu de los estadísticos es el siguiente:

Logaritmo medio 3.1321

Desviación normal de leños 0.5665

Coefficiente oblicuo de leños -0.4396

Años 35 (Se anularon los eventos con cero, así como los outliers bajos)

Paso 4.- Comprobación para outliers altos

El umbral del outlier alto es calculado para 41,770 pies³/seg basado en el procedimiento del paso 3. Ningún evento excede el valor del umbral.

Paso 5.- Calculo y ajuste la curva de frecuencia condicional.

Una curva de frecuencia condicional se calcula basándose en los estadísticos del paso 3 y entonces es ajustada por la probabilidad condicional.

El coeficiente oblicuo se ha redondeado a -0.4 para facilidad en el calculo. La proporción de ajuste es $35/42 = 0.83333$.

Tabla 4.9. Calculo de coordenadas de curva de frecuencia condicional.

P_d	$K_{G,P}$ para $G = -0.4$	$\log Q$	Q (pies ³ /seg)	Ajuste de la probabilidad de excedencia (P.P _d)
0.99	-2.61539	1.6505	44.7	0.825
0.90	-1.31671	2.3862	243	0.750
0.50	0.06651	3.1696	1460	0.417
0.10	1.23114	3.8295	6750	0.083
0.05	1.52357	3.9952	98900	0.042
0.02	1.83361	4.1708	14800	0.017
0.01	2.02933	4.2817	19100	0.0083
0.005	2.20092	4.3789	23900	0.0042
0.002	2.39942	4.4914	31000	0.0017

Las curvas de frecuencia condicional se muestran en la figura 4.7

Paso 6.- Calculo de las estadísticos sintéticos.

Primero se determinan los $Q_{0.01}$, $Q_{0.10}$, y $Q_{0.50}$ descargas de la curva ajustada en la figura (4.7).

$$Q_{0.01} = 17940$$

$$Q_{0.10} = 6000$$

$$Q_{0.50} = 1060$$

Calculo del coeficiente oblicuo sintético a través de la ecuación (3.14).

$$G_s = -2.50 + 3.12 \frac{\log\left(\frac{17940}{6000}\right)}{\log\left(\frac{6000}{1060}\right)} = -0.5287$$

$$G_s = -0.5 \text{ (redondeó a más cercano décimo)}$$

Calculo del sintético de la desviación normal a través de la ecuación (3.15).

$$S_s = \frac{\log\left(\frac{17940}{1060}\right)}{(1.95472 - 0.08302)} = 0.6564$$

Calculo del sintético medio a través de la ecuación (3.16).

$$\bar{X}_s = \log(1060) - (0.08302)(0.6564) = 2.9708$$

Paso 7. - Calculo del coeficiente oblicuo a través de la ecuación (3.7).

Se determina una inclinación generalizada de -0.3. Del apéndice 2 el error medio cuadrado de la inclinación de la estación es 0.163.

$$G_w = \frac{0.302(-0.529) + 0.163(-0.3)}{0.302 + 0.163} = -0.4487$$

$G_w = -0.4$ (redondeó a más cercano décimo)

Paso 8. - Calculo de la curva de frecuencia final.

Tabla 4.10. Calculo de ordenadas de curva de frecuencia

P	$K_{G_w, P}$ para $G_w = -0.4$	log Q	Q (pies³/seg)
0.99	-2.61539	1.2541	17.9
0.90	-1.31671	2.1065	128
0.50	0.06651	3.0145	1030
0.10	1.23114	3.7789	6010
0.05	1.52357	3.9709	9350
0.02	1.83361	4.1744	14900
0.01	2.02933	4.3029	20100
0.005	2.20092	4.4155	26000
0.002	2.39942	4.5458	351000

Las curvas de frecuencia se muestran en la figura 4.8.

GASTO EN PIES CUBICOS POR SEGUNDO

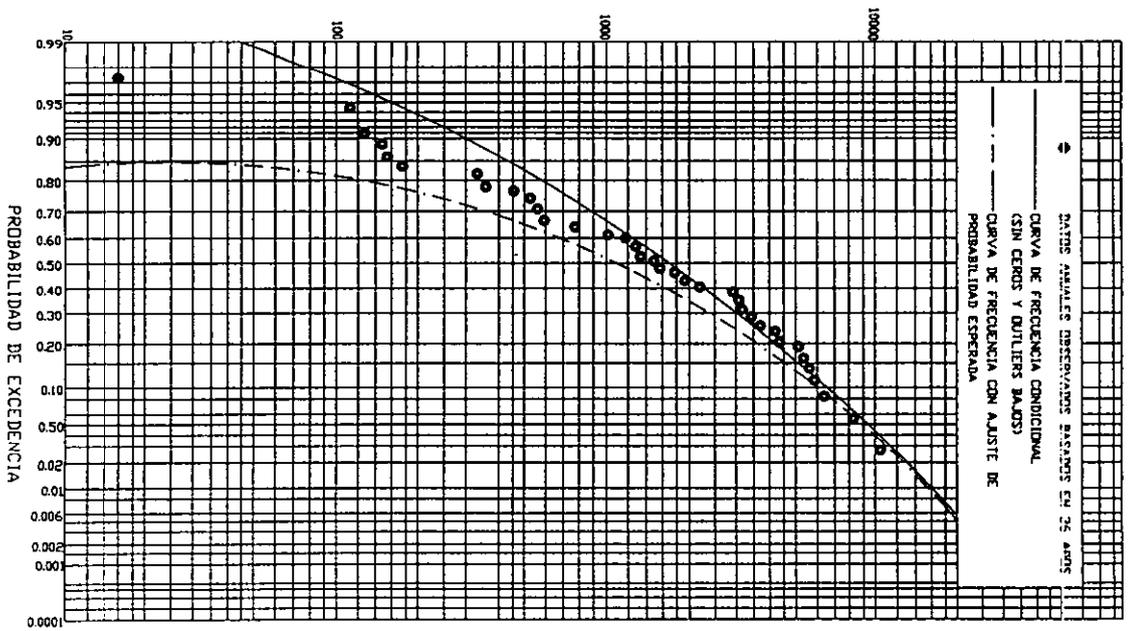
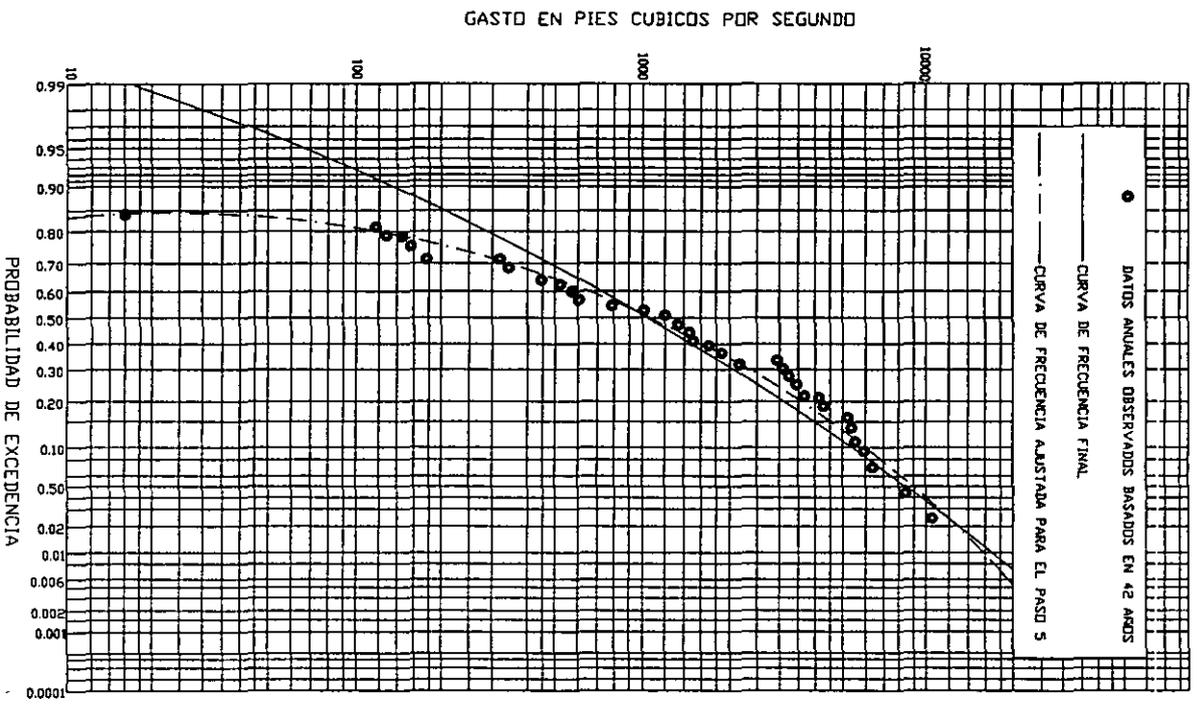


FIGURA 3.7. CURVAS DE FRECUENCIA AJUSTADAS



4.5. Aplicación de la distribución Doble Gumbel, para la estación Poza Rica. Ver

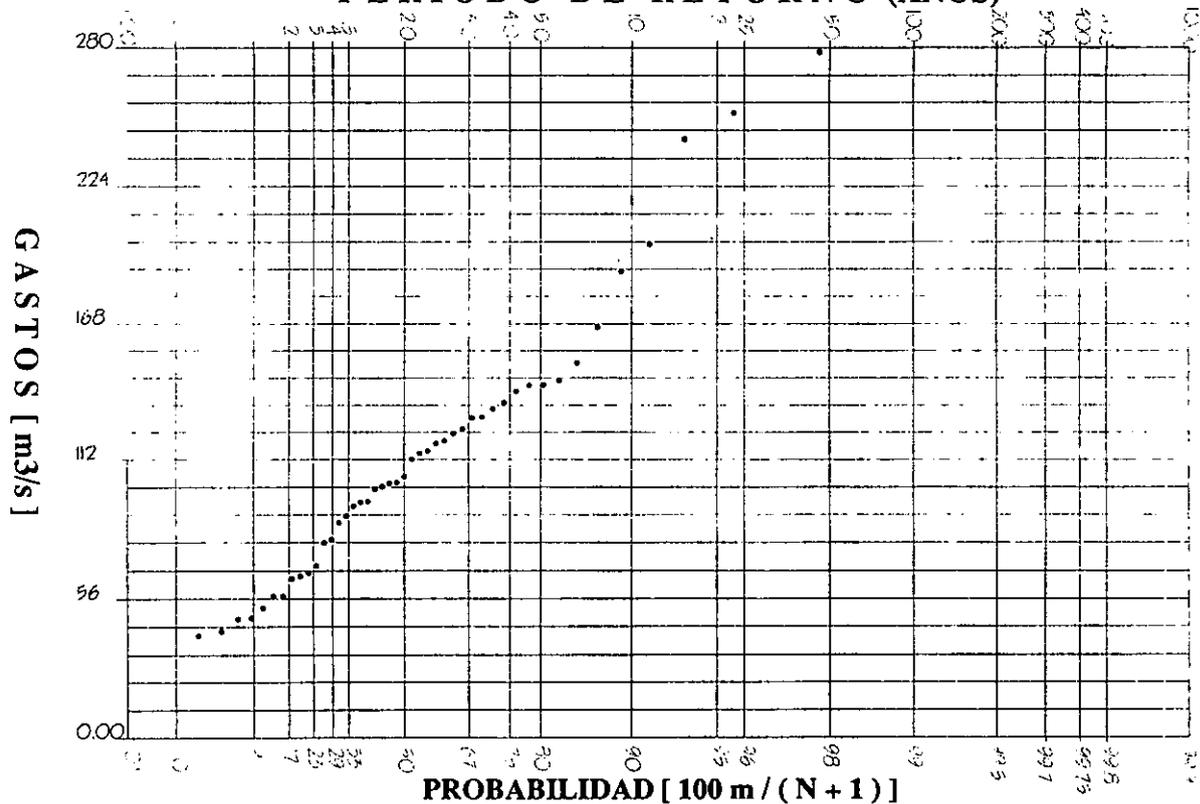
m	Tr (años)	\bar{h}_p 24 horas	Población Ciclónica	Población No ciclónica	Probabilidad de no excedencia	\bar{h}_p ajustada	$(X_i - \bar{h}_p \text{ ajust})^2$
1	46.00	278.20	278.20		0.978	260.15	325.80
2	23.00	253.30	253.30		0.957	230.88	502.66
3	15.33	242.80	242.80		0.935	212.81	899.40
4	11.50	200.50	200.50		0.913	199.25	1.56
5	9.20	189.30	189.30		0.891	188.09	1.46
6	7.67	166.50		166.50	0.870	178.38	141.13
7	6.57	151.60		151.60	0.848	169.65	325.80
8	5.75	144.40		144.40	0.826	161.62	296.53
9	5.11	142.70		142.70	0.804	154.17	131.56
10	4.60	142.50		142.50	0.783	147.26	22.66
11	4.18	140.00		140.00	0.761	140.89	0.79
12	3.83	135.50		135.50	0.739	135.10	0.16
13	3.54	132.80		132.80	0.717	129.87	8.58
14	3.29	129.40		129.40	0.696	125.18	17.81
15	3.07	129.00		129.00	0.674	120.98	64.32
16	2.88	124.50		124.50	0.652	117.19	53.44
17	2.71	122.50		122.50	0.630	113.76	76.39
18	2.56	119.50		119.50	0.609	110.62	78.85
19	2.42	118.50		118.50	0.587	107.73	115.99
20	2.30	115.50		115.50	0.565	105.04	109.41
21	2.19	114.50		114.50	0.543	102.53	143.28
22	2.09	112.00		112.00	0.522	100.17	139.95
23	2.00	105.00		105.00	0.500	97.93	49.98
24	1.92	102.70		102.70	0.478	95.79	47.75
25	1.84	102.40		102.40	0.457	93.74	75.00
26	1.77	101.00		101.00	0.435	91.76	85.38
27	1.70	100.00		100.00	0.413	89.84	103.23
28	1.64	95.00		95.00	0.391	87.97	49.42
29	1.59	94.70		94.70	0.370	86.15	73.10
30	1.53	93.00		93.00	0.348	84.35	74.82
31	1.48	89.00		89.00	0.326	82.58	41.22
32	1.44	86.30		86.30	0.304	80.82	30.03
33	1.39	79.50		79.50	0.283	79.07	0.18
34	1.35	78.20		78.20	0.261	77.31	0.79
35	1.31	69.10		69.10	0.239	75.53	41.34
36	1.28	66.30		66.30	0.217	73.73	55.20
37	1.24	65.00		65.00	0.196	71.88	47.33
38	1.21	64.00		64.00	0.174	69.97	35.64
39	1.18	57.00		57.00	0.152	67.98	120.56
40	1.15	57.00		57.00	0.130	65.87	78.68
41	1.12	52.30		52.30	0.109	63.59	127.46
42	1.10	48.20		48.20	0.087	61.08	165.89
43	1.07	47.80		47.80	0.065	58.18	107.74
44	1.05	43.00		43.00	0.043	54.62	135.02
45	1.02	41.30		41.30	0.022	49.55	68.06
Media =			232.82	99.48		Suma =	5071.40
Desviación Estándar =			37.14	33.88		Error Cuadrático	71.21

Valores extrapolados

Periodo de retorno (años)	Valor estimado \bar{h}_p mm
5	152.7
10	192.3
20	224.8
50	263.6
100	291.8
500	356.1
1000	383.6
10000	473.9

PAPEL DE PROBABILIDADES DE GUMBEL

PERIODO DE RETORNO (ANOS)



CAPITULO V



CAPITULO CINCO

CONCLUSIONES

La hidrología es una de las ramas más importantes de la ingeniería civil, ya que de ella depende la construcción y buen funcionamiento de un gran número de obras civiles como todo tipo de presas, bordos, encauzamientos, sistemas de drenaje, sistemas de riego, puentes, etc.

La infraestructura hidráulica con que cuenta nuestro país es muy deficiente, ya que si nos remontamos a épocas muy antiguas siempre han existido problemas principalmente de inundaciones, esto debido a la excesiva precipitación que por ella misma provoca inundaciones o también que se desborden ríos, esto debido a el mal funcionamiento de la infraestructura existente. Esto se podría evitar realizando un buen estudio hidrológico con un periodo de retorno adecuado, en las zonas que estén más propensas a este tipo de problemas, esto resultaría muy costoso pero a la larga sería más económico, ya que se evitarían este tipo de problemas y después no se tendría que gastar en resolver los problemas de inundaciones.

La mayoría de las estaciones tanto climatológicas como hidrométricas con que cuenta nuestro país tienen muy pocos datos o tienen registros incompletos, y otro problema sería que no se cuenta con el número necesario de estaciones, por que en algunos casos las estaciones están muy distantes del sitio donde se va a realizar la obra, lo cual repercute en los resultados de las distribuciones que se aplican, ya que con esto crece el error cuadrático.

Cuando existe una gran dispersión de los datos, el error cuadrático también es muy alto, esto se puede tratar con la distribución doble Gumbel, que trata los datos como dos tipos de poblaciones, lo que reduce el error cuadrático. Otra forma de tratar las estaciones que cuentan con muy pocos datos o que tienen registros con cero, es con los casos especiales; ya que trata este tipo de datos de manera distinta que los demás, lo que reduce el error.

El objetivo primordial de este trabajo es dar a conocer un método para estimar eventos hidrológicos con mayor precisión, ya que indica que hacer cuando existen datos muy elevados o muy bajos, sin tomarlos como dos poblaciones, y también que hacer cuando la estación tiene registros incompletos, o cuando algunos de sus registros tienen cero. A pesar de que este estudio solo se ha aplicado en los Estados Unidos, este método sería muy aplicable a nuestro país ya que se ajusta perfectamente a necesidades que existen en México, pero desafortunadamente no se pueden obtener todos los coeficientes que se necesitan para su aplicación, en Estados Unidos existen incluso tablas o mapas con los valores de estos coeficientes, pero antes de esto ya realizaron los estudios correspondientes a todo el país, cosa que no se ha hecho en México, ya que aquí solo se realizan los estudios hasta que se va a realizar algún tipo de obra y se trata de economizar a el máximo, aunque después existan problemas.

Espero que se lleguen a realizar algunos estudios para que este método se pueda aplicar en nuestro país, o en su defecto por lo menos que se realicen mejores estudios hidrológicos para evitar muchos problemas que año con año aquejan a nuestro país y que se podrían evitar, principalmente las inundaciones.

También esperando que este trabajo sirva a el estudiante de ingeniería civil para conocer un poco mas sobre la hidrología y sus aplicaciones, el conocer métodos nuevos y no solo los que conocemos en México, ya que este método no es tan reciente, pero se conoce poco de el en Latinoamérica, debido a que existe poca bibliografía sobre el.

APENDICES



La tabla siguiente contiene los valores de K para su uso en la ecuación (3.1), para los coeficientes oblicuos, G de 0 a 9.0 y 0 a -9.0 y probabilidades de excedencia, P , de 0.9999 a 0.0001.

Pueden obtenerse valores aproximados de K cuando los coeficientes oblicuos están entre 1.0 y -1.0:

$$K = \frac{2}{G} \left[\left[\left(K_n - \frac{G}{6} \right) \frac{G}{6} + 1 \right]^3 - 1 \right]$$

donde: K_n es la desviación normal y G es el coeficiente oblicuo. Debido a las limitaciones involucradas en el uso de esta ecuación es preferible el uso de las tablas.

P	$G = 0.0$	$G = 0.1$	$G = 0.2$	$G = 0.3$	$G = 0.4$	$G = 0.5$	$G = 0.6$
0.9999	-3.71902	-3.50703	-3.29921	-3.09631	-2.89907	-2.70836	-2.52507
0.9995	-3.29053	-3.12767	-2.96670	-2.80889	-2.65390	-2.50257	-2.35549
0.9990	-3.09023	-2.94834	-2.80786	-2.66915	-2.53261	-2.39867	-2.26780
0.9980	-2.87816	-2.75706	-2.63672	-2.51741	-2.39942	-2.28311	-2.16884
0.9950	-2.57583	-2.48187	-2.38795	-2.29423	-2.20092	-2.10825	-2.01644
0.9900	-2.32635	-2.25258	-2.17840	-2.10394	-2.02933	-1.95472	-1.88029
0.9800	-2.05375	-1.99973	-1.94499	-1.88959	-1.83361	-1.77716	-1.72033
0.9750	-1.95996	-1.91219	-1.86360	-1.81427	-1.76427	-1.71366	-1.66253
0.9600	-1.75069	-1.71580	-1.67999	-1.64329	-1.60574	-1.56740	-1.52830
0.9500	-1.64485	-1.61594	-1.58607	-1.55527	-1.52357	-1.49101	-1.45762
0.9000	-1.28155	-1.27037	-1.25824	-1.24516	-1.23114	-1.21618	-1.20028
0.8000	-0.84162	-0.84611	-0.84986	-0.85285	-0.85508	-0.85653	-0.85718
0.7000	-0.52440	-0.53624	-0.54757	-0.55839	-0.56867	-0.57840	-0.58757
0.6000	-0.25335	-0.26882	-0.28403	-0.29897	-0.31362	-0.32796	-0.34198
0.5704	-0.17733	-0.19339	0.20925	-0.22492	-0.24037	-0.25558	-0.27047
0.5000	0.00000	-0.01662	-0.03325	-0.04993	-0.06651	-0.08302	-0.09945
0.4296	0.17733	0.16111	0.14472	0.12820	0.11154	0.09478	0.07791
0.4000	0.25335	0.23763	0.22168	0.20552	0.18916	0.17261	0.15589
0.3000	0.52440	0.51207	0.49927	0.48600	0.47228	0.45812	0.44352
0.2000	0.84162	0.83639	0.83044	0.82377	0.81638	0.80829	0.79950
0.1000	1.28155	1.29178	1.30105	1.30936	1.31671	1.32309	1.32850
0.0500	1.65485	1.67279	1.69971	1.72562	1.75048	1.77428	1.79701
0.0400	1.75069	1.78462	1.81756	1.84949	1.88039	1.91022	1.93896
0.0250	1.95996	2.00688	2.05290	2.09795	2.14202	2.18505	2.22702
0.0200	2.05375	2.10697	2.15935	2.21081	2.26133	2.31084	2.35931
0.0100	2.32635	2.39961	2.47226	2.54442	2.61539	2.68572	2.75514
0.0050	2.57583	2.66965	2.76321	2.85636	2.94900	3.04102	3.13232
0.0020	2.87816	2.99978	3.12169	3.24371	3.36566	3.48737	3.60872
0.0010	3.09023	3.23322	3.37703	3.52139	3.66608	3.81090	3.95567
0.0005	3.29053	3.45513	3.62113	3.78820	3.95605	4.12443	4.29311
0.0001	3.71902	3.93453	4.15301	4.37394	4.59687	4.82141	5.04718

P	G = 0.7	G = 0.8	G = 0.9	G = 1.0	G = 1.1	G = 1.2	G = 1.3
0.9999	-2.35015	-2.18448	-2.02891	-1.88410	-1.75053	-1.62838	-1.51752
0.9995	-2.21328	-2.07661	-1.94611	-1.82241	-1.70660	-1.59738	-1.49673
0.9990	-2.14053	-2.01739	-1.89894	-1.78572	-1.67825	-1.57695	-1.48216
0.9980	2.05701	-1.04806	-1.84244	-1.74062	-1.64305	-1.55016	1.46232
0.9950	-1.92580	-1.83660	-1.74919	-1.66390	-1.58110	-1.50114	-1.42439
0.9900	-1.80621	-1.73271	-1.66001	-1.58838	-1.51808	-1.44942	-1.38267
0.9800	-1.66325	-1.60604	-1.54886	-1.49188	-1.43529	-1.37929	-1.32412
0.9750	-1.61099	-1.55914	-1.50712	-1.45507	-1.40314	-1.35153	-1.30042
0.9600	-1.48852	-1.44813	-1.40720	-1.36584	-1.32414	-1.28225	-1.24028
0.9500	-1.42325	-1.38855	-1.35299	-1.31684	-1.28019	-1.24313	-1.20578
0.9000	-1.18347	-1.16574	-1.14712	-1.12762	-1.10726	-1.08608	-1.06413
0.8000	-0.85703	-0.85607	-0.85426	-0.85161	-0.84809	-0.84369	-0.83841
0.7000	-0.59615	-0.60412	-0.61146	-0.61815	-0.62415	-0.62944	-0.63400
0.6000	-0.35565	-0.36889	-0.38186	-0.39434	-0.40638	-0.41794	-0.42899
0.5704	-0.28516	-0.29961	-0.31368	-0.32740	-0.34075	-0.35370	-0.36620
0.5000	-0.11578	-0.13199	-0.14807	-0.16397	-0.17968	-0.19517	-0.21040
0.4296	0.06097	0.04397	0.02693	0.00987	-0.00719	-0.02421	-0.04116
0.4000	0.13901	0.12199	0.10486	0.08763	0.07032	0.05297	0.03560
0.3000	0.42851	0.41309	0.39729	0.38111	0.36458	0.34772	0.33054
0.2000	0.79002	0.77986	0.76902	0.75752	0.74537	0.73257	0.71915
0.1000	1.33294	1.33640	1.33889	1.34039	1.34092	1.34047	1.33904
0.0500	1.81864	1.83916	1.85856	1.87683	1.89395	1.90992	1.92472
0.0400	1.96660	1.99311	2.01848	2.04269	2.06573	2.08758	2.10823
0.0250	2.26790	2.30764	2.34623	2.38364	2.41984	2.45482	2.48855
0.0200	2.40670	2.45298	2.49811	2.54206	2.58480	2.62631	2.66657
0.0100	2.82359	2.89101	2.95735	3.02256	3.08660	3.14944	3.21103
0.0050	3.22281	3.31243	3.40109	3.48874	3.57530	3.66073	3.74497
0.0020	3.72957	3.84981	3.96932	4.08802	4.20582	4.32226	4.43839
0.0010	4.10022	4.24439	4.38807	4.53112	4.67344	4.81492	4.95549
0.0005	4.46489	4.63057	4.79899	4.96701	5.13449	5.30130	5.46735
0.0001	5.27389	5.50124	5.72899	5.95691	6.18480	6.41249	6.63980

Error medio cuadrado de la inclinación de la estación

Inclinación de la estación (G o G')	10	20	30	40	50	60	70	80	90	100
0.0	0.468	0.244	0.167	0.127	0.103	0.087	0.075	0.066	0.059	0.054
0.1	0.476	0.253	0.175	0.134	0.109	0.093	0.080	0.071	0.064	0.058
0.2	0.485	0.262	0.183	0.142	0.116	0.099	0.086	0.077	0.069	0.063
0.3	0.494	0.272	0.192	0.150	0.123	0.105	0.092	0.082	0.074	0.068
0.4	0.504	0.282	0.201	0.158	0.131	0.113	0.099	0.089	0.080	0.073
0.5	0.513	0.293	0.211	0.167	0.139	0.120	0.106	0.095	0.087	0.079
0.6	0.522	0.303	0.221	0.176	0.148	0.128	0.114	0.102	0.093	0.086
0.7	0.532	0.315	0.231	0.186	0.157	0.137	0.122	0.110	0.101	0.093
0.8	0.542	0.326	0.243	0.196	0.167	0.146	0.130	0.118	0.109	0.100
0.9	0.562	0.345	0.259	0.211	0.181	0.159	0.142	0.130	0.119	0.111
1.0	0.603	0.376	0.285	0.235	0.202	0.178	0.160	0.147	0.135	0.126
1.1	0.646	0.410	0.315	0.261	0.225	0.200	0.181	0.166	0.153	0.143
1.2	0.692	0.448	0.347	0.290	0.252	0.225	0.204	0.187	0.174	0.163
1.3	0.741	0.488	0.383	0.322	0.281	0.252	0.230	0.212	0.197	0.185
1.4	0.794	0.533	0.422	0.357	0.314	0.283	0.259	0.240	0.224	0.211
1.5	0.851	0.581	0.465	0.397	0.351	0.318	0.292	0.271	0.254	0.240
1.6	0.912	0.623	0.498	0.425	0.376	0.340	0.313	0.291	0.272	0.257
1.7	0.977	0.667	0.534	0.456	0.403	0.365	0.335	0.311	0.292	0.275
1.8	1.047	0.715	0.572	0.489	0.432	0.391	0.359	0.334	0.313	0.295
1.9	1.122	0.766	0.613	0.523	0.463	0.419	0.385	0.358	0.335	0.316
2.0	1.202	0.821	0.657	0.561	0.496	0.449	0.412	0.383	0.359	0.339
2.1	1.288	0.880	0.704	0.601	0.532	0.481	0.442	0.410	0.385	0.363
2.2	1.380	0.943	0.754	0.644	0.570	0.515	0.473	0.440	0.412	0.389
2.3	1.479	1.010	0.808	0.690	0.610	0.552	0.507	0.471	0.442	0.417
2.4	1.585	1.083	0.866	0.739	0.654	0.592	0.543	0.505	0.473	0.447
2.5	1.698	1.160	0.928	0.792	0.701	0.634	0.582	0.541	0.507	0.479
2.6	1.820	1.243	0.994	0.849	0.751	0.679	0.624	0.580	0.543	0.513
2.7	1.950	1.332	1.066	0.910	0.805	0.728	0.669	0.621	0.582	0.550
2.8	2.089	1.427	1.142	0.975	0.862	0.780	0.716	0.666	0.624	0.589
2.9	2.239	1.529	1.223	1.044	0.924	0.836	0.768	0.713	0.669	0.631
3.0	2.399	1.638	1.311	1.119	0.990	0.895	0.823	0.764	0.716	0.676

Valores de K para Outliers

Valores nivelados de K

La siguiente tabla contiene los valores nivelados de K_N para una distribución normal. Las Pruebas dirigieron a seleccionar que los procedimientos para determinar si existen outliers, usados en este trabajo indican que los valores de K_N son aplicables a la distribución a Log Pearson Tipo III.

Tamaño de la muestra	Valores de K_N	Tamaño de la muestra	Valores de K_N	Tamaño de la muestra	Valores de K_N	Tamaño de la muestra	Valores de K_N
10	2.036	45	2.727	80	2.940	115	3.064
11	2.088	46	2.736	81	2.945	116	3.067
12	2.134	47	2.744	82	2.949	117	3.070
13	2.175	48	2.753	83	2.953	118	3.073
14	2.213	49	2.760	84	2.957	119	3.075
15	2.247	50	2.768	85	2.961	120	3.078
16	2.279	51	2.775	86	2.966	121	3.081
17	2.309	52	2.783	87	2.970	122	3.083
18	2.335	53	2.790	88	2.973	123	3.086
19	2.361	54	2.798	89	2.977	124	3.089
20	2.385	55	2.804	90	2.981	125	3.092
21	2.408	56	2.811	91	2.984	126	3.095
22	2.429	57	2.818	92	2.989	127	3.097
23	2.448	58	2.824	93	2.993	128	3.100
24	2.467	59	2.831	94	2.996	129	3.102
25	2.486	60	2.837	95	3.000	130	3.104
26	2.502	61	2.842	96	3.003	131	3.107
27	2.519	62	2.849	97	3.006	132	3.109
28	2.534	63	2.854	98	3.011	133	3.112
29	2.549	64	2.860	99	3.014	134	3.114
30	2.563	65	2.866	100	3.017	135	3.116
31	2.577	66	2.871	101	3.021	136	3.119
32	2.591	67	2.877	102	3.024	137	3.122
33	2.604	68	2.883	103	3.027	138	3.124
34	2.616	69	2.888	104	3.030	139	3.126
35	2.628	70	2.893	105	3.033	140	3.129
36	2.639	71	2.897	106	3.037	141	3.131
37	2.650	72	2.903	107	3.040	142	3.133
38	2.661	73	2.908	108	3.043	143	3.135
39	2.671	74	2.912	109	3.046	144	3.138
40	2.682	75	2.917	110	3.049	145	3.140
41	2.692	76	2.922	111	3.052	146	3.142
42	2.700	77	2.927	112	3.055	147	3.144
43	2.710	78	2.931	113	3.058	148	3.146
44	2.719	79	2.935	114	3.061	149	3.148

LIMITE DE CONFIANZA VALORES POR LA DISTRIBUCION NORMAL

PROBABILIDAD DE EXCEDENCIA

Nivel de confianza Longitud de registro sistemático

Nivel de confianza	Longitud de registro sistemático	PROBABILIDAD DE EXCEDENCIA											
		0.002	0.005	0.010	0.020	0.040	0.100	0.200	0.500	0.800	0.900	0.950	0.990
0.01	10	6.18	5.572	5.074	4.535	3.942	3.048	2.243	0.892	-0.107	-0.508	-0.804	-1.314
	15	5.15	4.639	4.222	3.770	3.274	2.521	1.841	0.678	-0.236	-0.629	-0.929	-1.458
	20	4.68	4.212	3.832	3.419	2.965	2.276	1.651	0.568	-0.313	-0.705	-1.008	-1.550
	25	4.40	3.960	3.601	3.211	2.782	2.129	1.536	0.498	-0.364	-0.757	-1.064	-1.616
	30	4.21	3.792	3.447	3.071	2.658	2.030	1.457	0.450	-0.403	-0.797	-1.107	-1.667
	40	3.98	3.577	3.249	2.893	2.500	1.902	1.355	0.384	-0.457	-0.854	-1.169	-1.741
	50	3.83	3.442	3.125	2.781	2.401	1.821	1.290	0.340	-0.496	-0.894	-1.212	-1.793
	60	3.72	3.347	3.038	2.702	2.331	1.764	1.244	0.309	-0.524	-0.924	-1.245	-1.833
	70	3.65	3.278	2.974	2.644	2.280	1.722	1.210	0.285	-0.545	-0.948	-1.272	-1.865
	80	3.59	3.223	2.924	2.599	2.239	1.688	1.183	0.265	-0.565	-0.968	-1.293	-1.891
90	3.54	3.179	2.883	2.561	2.206	1.661	1.160	0.250	-0.578	-0.984	-1.311	-1.913	
100	3.5	3.143	2.850	2.531	2.179	1.639	1.142	0.236	-0.591	-0.998	-1.326	-1.932	
0.05	10	4.86	4.379	3.981	3.549	3.075	2.355	1.702	0.580	-0.317	-0.712	-1.017	-1.563
	15	4.3	3.874	3.52	3.136	2.713	2.068	1.482	0.455	-0.406	-0.802	-1.114	-1.677
	20	4.03	3.628	3.295	2.934	2.534	1.926	1.370	0.387	-0.460	-0.858	-1.175	-1.749
	25	3.87	3.478	3.158	2.809	2.425	1.838	1.301	0.342	-0.497	-0.898	-1.217	-1.801
	30	3.76	3.376	3.064	2.724	2.350	1.777	1.252	0.310	-0.525	-0.928	-1.250	-1.840
	40	3.61	3.242	2.941	2.613	2.251	1.697	1.188	0.266	-0.565	-0.970	-1.297	-1.896
	50	3.52	3.157	2.862	2.542	2.188	1.646	1.146	0.237	-0.592	-1.000	-1.329	-1.936
	60	3.45	3.096	2.807	2.492	2.143	1.609	1.116	0.216	-0.612	-1.022	-1.354	-1.966
	70	3.4	3.051	2.765	2.454	2.110	1.581	1.093	0.199	-0.629	-1.040	-1.374	-1.990
	80	3.36	3.016	2.733	2.425	2.083	1.559	1.076	0.186	-0.642	-1.054	-1.390	-2.010
90	3.33	2.987	2.706	2.400	2.062	1.542	1.061	0.175	-0.652	-1.066	-1.403	-2.026	
100	3.3	2.963	2.684	2.380	2.044	1.527	1.049	0.166	-0.662	-1.077	-1.414	-2.040	

LIMITE DE CONFIANZA VALORES POR LA DISTRIBUCION NORMAL

PROBABILIDAD DE EXCEDENCIA

Nivel de confianza Longitud de registro sistemático

	N	0.002	0.005	0.010	0.020	0.040	0.100	0.200	0.500	0.800	0.900	0.950	0.990
0.1	10	4.32	3.889	3.532	3.144	2.716	2.066	1.474	0.437	-0.429	-0.828	-1.144	-1.715
	15	3.94	3.539	3.212	2.857	2.465	1.867	1.320	0.347	-0.499	-0.901	-1.222	-1.808
	20	3.74	3.364	3.052	2.712	2.338	1.765	1.240	0.297	-0.541	-0.946	-1.271	-1.867
	25	3.62	3.255	2.952	2.623	2.258	1.702	1.190	0.264	-0.570	-0.978	-1.306	-1.908
	30	3.54	3.181	2.884	2.561	2.204	1.657	1.154	0.239	-0.593	-1.002	-1.332	-1.940
	40	3.43	3.082	2.793	2.479	2.131	1.598	1.106	0.206	-0.624	-1.036	-1.369	-1.986
	50	3.36	3.019	2.735	2.426	2.084	1.559	1.075	0.184	-0.645	-1.059	-1.396	-2.018
	60	3.31	2.974	2.694	2.389	2.051	1.532	1.052	0.167	-0.662	-1.077	-1.415	-2.042
	70	3.28	2.94	2.662	2.360	2.025	1.511	1.035	0.155	-0.674	-1.091	-1.431	-2.061
	80	3.25	2.913	2.638	2.338	2.006	1.495	1.021	0.144	-0.684	-1.103	-1.444	-2.077
	90	3.22	2.891	2.618	2.319	1.989	1.481	1.010	0.136	-0.693	-1.112	-1.454	-2.090
	100	3.20	2.873	2.601	2.305	1.976	1.470	1.001	0.129	-0.701	-1.120	-1.463	-2.101
0.25	10	3.60	3.231	2.927	2.596	2.231	1.671	1.155	0.222	-0.625	-1.043	-1.382	-2.008
	15	3.42	3.064	2.775	2.460	2.112	1.577	1.083	0.179	-0.661	-1.081	-1.422	-2.055
	20	3.32	2.978	2.697	2.390	2.050	1.528	1.045	0.154	-0.683	-1.104	-1.448	-2.085
	25	3.26	2.925	2.648	2.346	2.011	1.497	1.020	0.137	-0.699	-1.121	-1.466	-2.106
	30	3.22	2.888	2.614	2.315	1.984	1.475	1.002	0.125	-0.710	-1.133	-1.479	-2.123
	40	3.17	2.838	2.568	2.274	1.948	1.445	0.978	0.108	-0.726	-1.151	-1.499	-2.147
	50	3.13	2.805	2.538	2.274	1.924	1.425	0.962	0.096	-0.738	-1.164	-1.513	-2.163
	60	3.11	2.783	2.517	2.227	1.907	1.411	0.950	0.088	-0.747	-1.173	-1.523	-2.176
	70	3.09	2.765	2.501	2.213	1.893	1.401	0.942	0.081	-0.753	-1.181	-1.532	-2.186
	80	3.07	2.752	2.489	2.202	1.883	1.392	0.935	0.076	-0.759	-1.187	-1.538	-2.194
	90	3.06	2.74	2.478	2.192	1.875	1.386	0.929	0.071	-0.763	-1.192	-1.544	-2.201
	100	3.05	2.731	2.47	2.184	1.868	1.380	0.925	0.068	-0.767	-1.196	-1.549	-2.207

LIMITE DE CONFIANZA VALORES POR LA DISTRIBUCION NORMAL

PROBABILIDAD DE EXCEDENCIA

Nivel de
confianza Longitud de
registro
sistemático

	N	0.002	0.005	0.010	0.020	0.040	0.100	0.200	0.500	0.800	0.900	0.950	0.990
0.75	10	2.51	2.235	2.008	1.759	1.480	1.043	0.625	-0.222	-1.155	-1.671	-2.104	-2.927
	15	2.56	2.284	2.055	1.803	1.521	1.081	0.661	-0.179	-1.083	-1.577	-1.991	-2.775
	20	2.60	2.317	2.085	1.831	1.547	1.104	0.683	-0.154	-1.045	-1.528	-1.932	-2.697
	25	2.62	2.339	2.106	1.851	1.566	1.121	0.699	-0.137	-1.020	-1.497	-1.895	-2.648
	30	2.64	2.357	2.123	1.867	1.580	1.133	0.710	-0.125	-1.002	-1.475	-1.869	-2.614
	40	2.67	2.383	2.147	1.888	1.600	1.151	0.726	-0.108	-0.978	-1.445	-1.834	-2.568
	50	2.69	2.4	2.163	1.903	1.614	1.164	0.738	-0.096	-0.962	-1.425	-1.811	-2.538
	60	2.70	2.414	2.176	1.916	1.625	1.173	0.747	-0.088	-0.950	-1.411	-1.795	-2.517
	70	2.71	2.425	2.186	1.925	1.634	1.181	0.753	-0.081	-0.942	-1.401	-1.782	-2.501
	80	2.72	2.434	2.194	1.932	1.640	1.187	0.759	-0.076	-0.935	-1.392	-1.772	-2.489
90	2.73	2.441	2.201	1.938	1.646	1.192	0.763	-0.071	-0.929	-1.386	-1.764	-2.478	
100	2.74	2.447	2.207	1.944	1.652	1.196	0.767	-0.068	-0.925	-1.380	-1.758	-2.470	
0.9	10	2.17	1.919	1.715	1.489	1.234	0.828	0.429	-0.437	-1.474	-2.066	-2.568	-3.532
	15	2.27	2.019	1.808	1.576	1.314	0.901	0.499	-0.347	-1.320	-1.867	-2.329	-3.212
	20	2.34	2.082	1.867	1.630	1.364	0.946	0.541	-0.297	-1.240	-1.765	-2.208	-3.052
	25	2.39	2.126	1.908	1.669	1.400	0.978	0.570	-0.264	-1.190	-1.702	-2.132	-2.952
	30	2.43	2.16	1.94	1.698	1.427	1.002	0.593	-0.239	-1.154	-1.657	-2.080	-2.884
	40	2.48	2.209	1.986	1.740	1.465	1.036	0.624	-0.206	-1.106	-1.598	-2.010	-2.793
	50	2.52	2.244	2.018	1.770	1.493	1.059	0.645	-0.184	-1.075	-1.559	-1.965	-2.735
	60	2.54	2.269	2.042	1.792	1.513	1.077	0.662	-0.167	-1.052	-1.532	-1.933	-2.694
	70	2.57	2.29	2.061	1.810	1.529	1.091	0.674	-0.155	-1.035	-1.511	-1.909	-2.662
	80	2.59	2.307	2.077	1.824	1.543	1.103	0.684	-0.144	-1.021	-1.495	-1.890	-2.638
90	2.60	2.321	2.09	1.836	1.553	1.112	0.693	-0.136	-1.010	-1.481	-1.874	-2.618	
100	2.61	2.333	2.101	1.847	1.563	1.120	0.701	-0.129	-1.001	-1.470	-1.861	-2.601	

LIMITE DE CONFIANZA VALORES POR LA DISTRIBUCION NORMAL

PROBABILIDAD DE EXCEDENCIA

Nivel de confianza Longitud de registro sistemático

		N	0.002	0.005	0.010	0.020	0.040	0.100	0.200	0.500	0.800	0.900	0.950	0.990
0.95	10	1.99	1.757	1.563	1.348	1.104	0.712	0.317	-0.580	-1.702	-2.355	-2.911	-3.981	
	15	2.12	1.878	1.677	1.454	1.203	0.802	0.406	-0.455	-1.482	-2.068	-2.566	-3.520	
	20	2.20	1.955	1.749	1.522	1.266	0.858	0.460	-0.387	-1.370	-1.926	-2.396	-3.295	
	25	2.26	2.011	1.801	1.569	1.309	0.898	0.497	-0.342	-1.301	-1.838	-2.292	-3.158	
	30	2.31	2.053	1.84	1.605	1.342	0.928	0.525	-0.310	-1.252	-1.777	-0.220	-3.064	
	40	2.38	2.113	1.896	1.657	1.910	0.970	0.565	-0.266	-1.188	-1.697	-2.125	-2.941	
	50	2.42	2.156	1.936	1.694	1.424	1.000	0.592	-0.237	-0.146	-1.646	-2.065	-2.862	
	60	2.46	2.188	1.966	1.722	1.450	1.022	0.612	-0.216	-1.116	-1.609	-2.022	-2.807	
	70	2.48	2.214	1.99	1.745	1.470	1.040	0.629	-0.199	-1.093	-1.571	-1.990	-2.765	
	80	2.51	2.235	2.01	1.762	1.487	1.054	0.642	-0.186	-1.076	-1.559	-1.964	-2.733	
	90	2.53	2.252	2.026	1.778	1.500	1.066	0.652	-0.175	-1.061	-1.542	-1.944	-2.706	
	100	2.54	2.267	2.040	1.791	1.512	1.077	0.662	-0.166	-1.049	-1.527	-1.927	-2.684	
0.99	10	1.70	1.492	1.314	1.115	0.886	0.508	0.107	-0.892	-2.243	-3.048	-3.738	-5.074	
	15	1.87	1.645	1.458	1.251	1.014	0.629	0.236	-0.678	-1.841	-2.521	-3.102	-4.222	
	20	1.97	1.743	1.55	1.336	1.094	0.705	0.313	-0.568	-1.651	-2.276	-2.808	-3.382	
	25	2.05	1.813	1.616	1.399	1.152	0.757	0.364	-0.498	-1.536	-2.129	-2.633	-3.601	
	30	2.11	1.867	1.667	1.446	1.196	0.797	0.403	-0.450	-1.457	-2.030	-2.515	-3.447	
	40	2.19	1.946	1.741	1.515	1.259	0.854	0.457	-0.384	-1.355	-1.902	-2.364	-3.249	
	50	2.26	2.002	1.793	1.563	1.304	0.894	0.496	-0.340	-1.290	-1.821	-2.269	-3.125	
	60	2.30	2.045	1.833	1.600	1.337	0.924	0.524	-0.309	-1.244	-1.764	-2.202	-3.038	
	70	2.34	2.079	1.865	1.630	1.365	0.948	0.545	-0.285	-1.210	-1.722	-2.153	-2.974	
	80	2.37	2.107	1.891	1.653	1.387	0.968	0.563	-0.265	-1.183	-1.688	-2.114	-2.924	
	90	2.39	2.131	1.913	1.674	1.405	0.984	0.578	-0.250	-1.160	-1.661	-2.082	-2.883	
	100	2.42	2.151	1.932	1.691	1.421	0.998	0.591	-0.236	-1.142	-1.639	-2.056	-2.850	

DIAGRAMA DE FLUJO PARA DATOS HISTORICOS Y GASTOS EXTRAÑOS

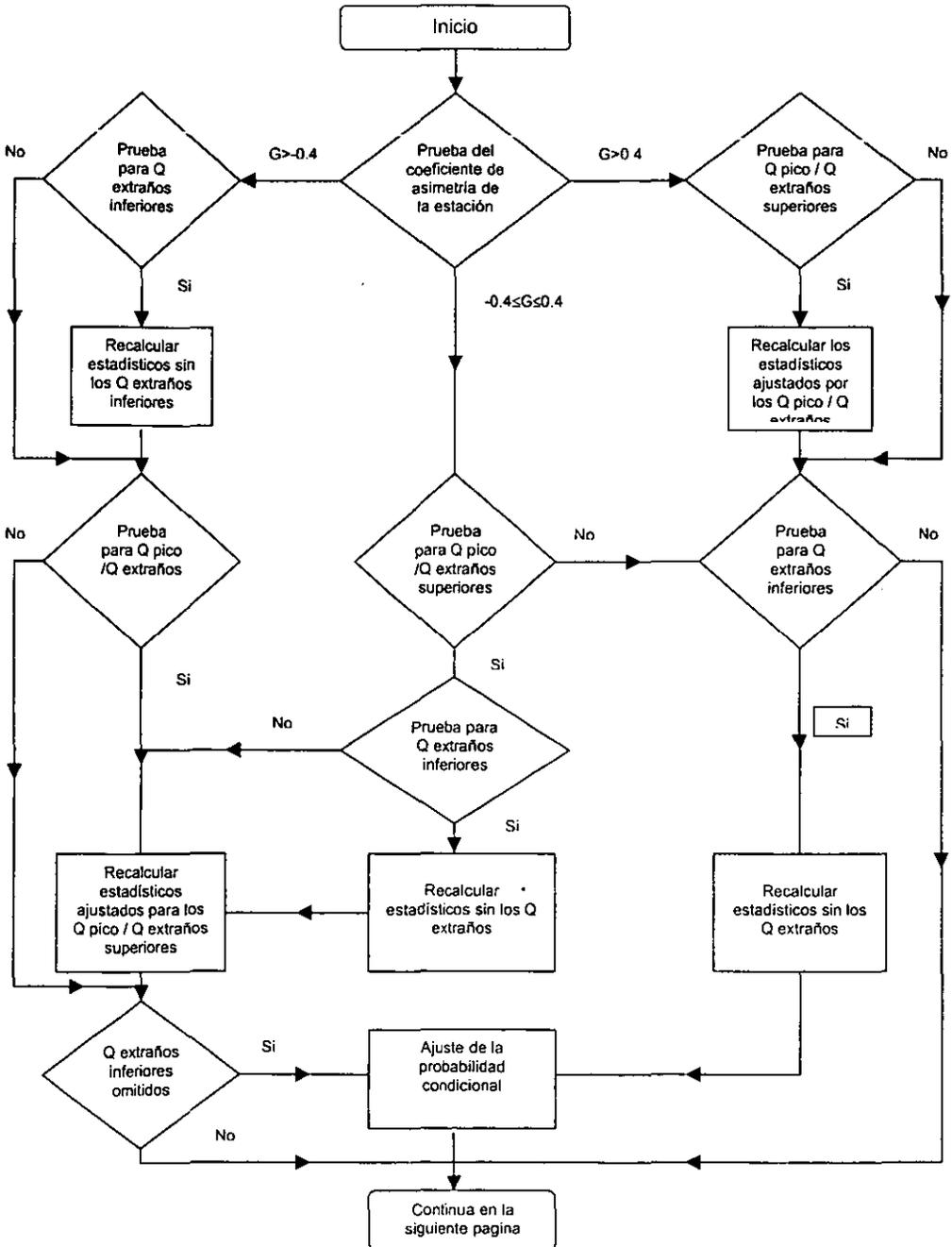
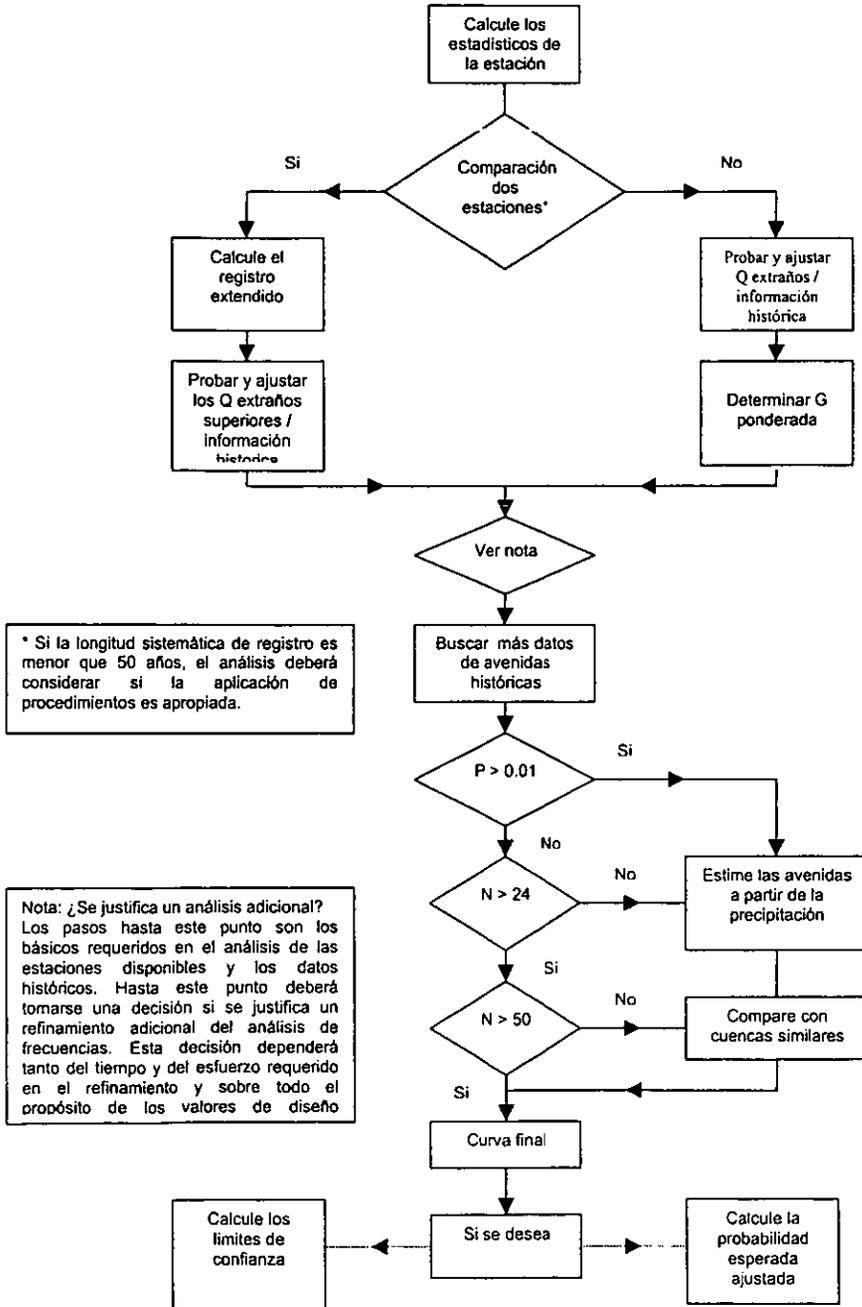


DIAGRAMA DE FLUJO PARA EL ANALISIS DE FRECUENCIA DE AVENIDAS



Descripción del algoritmo de Rosenbrock y del programa GUM2P

El propósito general del algoritmo de Rosenbrock es minimizar una función no lineal de múltiples variables no restringidas, esto es:

Minimizar $F(X_1, X_2, \dots, X_N)$

El procedimiento del algoritmo es como sigue (Rosenbrock, 1960):

- 1.- Se escogen un punto inicial e incrementos también iniciales S_i , $i = 1, 2, 3, \dots, n$. Se evalúa la función objetivo (FX).
- 2.- La primera variable X_1 se incrementa una distancia S_1 y se evalúa la función objetivo (FX); si el valor de FX decreció, el incremento denominado éxito y S_1 se multiplica por un factor α mayor o igual a uno. Si el valor de FX aumentó, el incremento es denominado falla y S_1 se multiplica por el factor β ($0 \leq \beta \leq 1.0$) y se invierte la dirección del movimiento.
- 3.- El procedimiento descrito se sigue para todas las variables, hasta encontrar un éxito (reducción de FX) y una falla (aumento de FX) en todas las N direcciones.
- 4.- Se realiza la rotación de ejes, la cual es designada como una etapa, por medio de las ecuaciones citadas por Kuester y Mize (1973).
- 5.- La búsqueda prosigue en cada una de las X direcciones usando los nuevos ejes de coordenados.
- 6.- El procedimiento termina cuando se satisface el criterio de convergencia, es decir, cuando la diferencia absoluta entre el valor actual de FX y el de la etapa anterior es inferior a la tolerancia establecida.

El código en Basic para el programa *GUM2P* tiene coma característica principal utilizar variables de sólo dos alfanuméricos, lo cual permite su fácil adaptación y uso en computadoras, siempre que se hagan los cambios necesarios para las impresiones o salidas de datos y resultados. La porción correspondiente al algoritmo de Rosenbrock (líneas 1580 a 2500) es una transcripción de la subrutina ROSENB en Fortran presentada por Kuester y Mize.

Las principales variables y parámetros del algoritmo de Rosenbrock y del programa *GUM2P* son:

NQ, Nc = número de gastos máximos anuales y ciclónicos, respectivamente.

$Q(NQ)$ = vector de gastos máximos anuales observados.

$X(I)$ = vector de valores iniciales supuestos para las cinco variables independientes ($I = 1, 2, \dots, 5$) o parámetros de ajuste.

$EP(I)$ = vector de incrementos iniciales para las variables ($I = 1, \dots, 5$).

MF = número máximo de veces que el programa evalúa la función objetivo ($MF = 1000$).

KF, KE = contadores del número de evaluaciones de la función objetivo y de etapas, respectivamente.

ME = número máximo de veces que se rotan los ejes ($ME = 30$).

MC = número de fallas sucesivas encontradas en todas las direcciones antes de terminar ($MC = 50$).

EY = tolerancia de error en la función objetivo o diferencia entre el valor actual y el valor de la etapa previa ($EY = 10^{-5}$).

NP = variables para el control de los incrementos que se utilizarán después de cada rotación de ejes; si es igual a uno, se utilizarán los incrementos iniciales; si es igual a dos, se emplearán los incrementos finales de la etapa previa ($NP = 2$).

AA = factor de escalamiento α para aumentar los incrementos ($AA = 2.0$).

BA = factor de escalamiento β para reducir los incrementos ($BA = 0.50$).

En la línea 1340 se define o calcula el parámetro c_2 , y se observa que la desviación estándar S_2 se obtiene a través de su estimador insesgado, esto es, dividiendo el segundo momento central entre $(nc - 1)$; por lo tanto, el número mínimo de eventos ciclónicos por considerar será dos. Por otra parte, como indica Bunday (1984), a la fecha no existe algoritmo de optimización universal garantizado bajo todas las circunstancias y, por ello en ocasiones, puede fallar; algunas veces con solo cambiar el criterio de terminación sensiblemente, por ejemplo de ≤ 0.00001 a ≤ 0.0001 , se logra alcanzar su conclusión. Para el algoritmo de Rosenbrock tal cambio se realiza en el parámetro EY de la línea 1400.

También conviene considerar que el cálculo del error estándar de ajuste se realiza por medio de una técnica numérica (bisección) y por ello, quizás en algunas ocasiones habrá que hacer cambios en las líneas 2690 ó 2720, para conseguir su inicio, funcionamiento y/o terminación. Si durante la ejecución del programa aparece en pantalla overflow, significa que el resultado de un cálculo es demasiado grande y que se está utilizando el valor infinito de la máquina; cuando se hace la rotación de ejes, es decir, al iniciar la otra etapa el mensaje desaparece, excepto si vuelve a ocurrir en dicha etapa.

A continuación se presenta el programa *GUM2P* desarrollado y aplicado en una computadora personal.

Lista del programa

```

1000 REM ESTE PROGRAMA UTILIZA EL ALGORITMO DE ROSENBROCK
1010 REM PARA OBTENER LOS PARAMETROS OPTIMOS DE AJUSTE DE
1020 REM LA FUNCION DE DISTRIB. GUMBEL DE DOS POBLACIONES
1030 REM PROGRAMA REALIZADO DURANTE ABRIL DE 1988
1040 REM POR DANIEL FCO. CAMPOS ARANDA EN PC: NCR
1050 REM SE INTRODUCEN LOS GASTOS MAXIMOS ANUALES
1060 CLS: PRINT"CALCULO DE LA DISTRIBUCION GUMBEL DE DOS POBLACIONES": PRINT
1070 INPUT"NOMBRE DE LA ESTACION HIDROMETRICA: "; EH$: PRINT
1080 PRINT"SE INTRODUCEN GASTOS MAXIMOS ANUALES (m^3/S)": PRINT
1090 INPUT"NUMERO DE GASTOS MAXIMOS ANUALES ="; NQ
1100 INPUT"NUMERO DE GASTOS MAXIMOS CICLONICOS ="; NC
1110 DIM EP(5),X(5),V(5,5),D(5),BL(5,5),Q(NQ)
1120 DIM BN(5),AJ(5),E(5),AL(5,5),XF(5),QO(NQ)
1130 FOR I=1 TO NQ
1140 PRINT"GASTO MAXIMO No.": I: INPUT "="; Q(I): NEXT I: PRINT
1150 INPUT"DESEA CORREGIR ALGUN GASTO (S/N)": A$: IF A$="N" THEN 1180: PRINT
1160 INPUT"INTRODUCIR No. DE GASTO Y VALOR CORRECTO": I,Q(I): PRINT: GOTO 1150
1170 REM INICIA ORDENAMIENTO DE DATOS
1180 CLS: LOCATE 8,10: PRINT"ORDENAMIENTO DE DATOS"
1190 FOR I=1 TO NQ
1200 QO(I)=Q(I): NEXT I
1210 FOR I=1 TO NQ-1: FOR J=I+1 TO NQ
1220 IF QO(J)<QO(I) THEN 1240
1230 Q1=QO (I): QO(I)=QO(J): QO(J)=Q1
1240 NEXT J: NEXT I
1250 REM SE EVALUAN VALORES INICIALES DE LOS PARAMETROS DE AJUSTE
1260 S1=0: S2=0: S3=0: S4=0
1270 FOR I=1 TO NQ
1280 IF I # (NQ-NC) THEN S1=S1+QO(I): S2=S2+QO(I)^2:GOTO 1300
1290 S3=S3+QO(I): S4=S4+QO(I)^2
1300 NEXT I
1310 X1=S1/(NQ-NC): X2=S3/NC
1320 X(2)=.779*SQR((S2-(NQ-NC)*X1^2)/(NQ-NC-1))
1330 X(1)=-.577*X(2): X(5)=(NQ-NC)/NQ
1340 X(4)=.779*SQR((S4-NC*X2^2)/(NC-1))
1350 X(3)=X2-.577*X(4)
1360 FOR I=1 TO 5
1370 EP(I) = .1*X(I): NEXT I
1380 LOCATE 10,10: IMPUT"DESEA IMPRESIÓN DE RESULTADOS POR ETAPAS (S/N)": B$
1390 REM SE DEFINEN PARAMETROS BASICOS DEL ALGORITMO DE ROSENBROCK
1400 MF=1000: ME=30: MC=50: NP=2: EY=.00001: AA=2: BA=.5: KE=1
1410 LPRINT" FUNCION DE DISTRIBUCION GUMBEL DE 2 POBLACIONES"
1420 LPRINT "AJUSTE DE PARAMETROS A TRAVES DEL ALGORITMO DE ROSENBROCK"
1430 LPRINT: LPRINT"ESTACION HIDROMETRICA: "; EH$
1440 LPRINT: LPRINT"TOLERANCIA DE ERROR EN LA FUNCION OBJETIVO ="; EY
1450 LPRINT" No. GASTOS MAXIMOS: ANUALES Y ORDENADOS EN m^3/S"
1460 FOR I=1 TO NQ
1470 LPRINT USING"#####": I: LPRINT
    USING"#####": Q(I):
1480 LPRINT USING"#####": QO(I): NEXT I

```

```

1490 LPRINT"NUMERO DE GASTOS MAXIMOS CICLONICOS ="; NC: LPRINT
1500 LPRINT"VALORES INICIALES DE LOS PARAMETROS DE AJUSTE EN m3/S:"
1510 LPRINT"PARAMETRO DE UBICACIÓN DE LA POBLACION NO CICLONICA (a1) = "; X(1)
1520 LPRINT"PARAMETRO DE ESCALA DE LA POBLACION NO CICLONICA (c1) ="; X(2)
1530 LPRINT"PARAMETRO DE UBICACIÓN DE LA POBLACION CICLONICA (a2) ="; X(3)
1540 LPRINT"PARAMETRO DE ESCALA DE LA POBLACION CICLONICA (c2) ="; X(4)
1550 LPRINT"PROBABILIDAD DE EVENTOS NO CICLONICOS (p) ="; X(5)
1560 LPRINT: GOSUB 2960: LPRINT"PRIMER VALOR DE LA FUNCION OBJETIVO ="; FX
1570 REM INICIAN CALCULOS DEL ALGORITMO DE ROSENBROCK
1580 LOCATE 12,10: PRINT"CALCULOS EN PROCESO": KF=1
1590 FOR II=1 TO 5
1600 FOR JJ=1 TO 5
1610 V(II,JJ)=0!
1620 IF(II-JJ) = 0 THEN V(II,JJ)=1!
1630 NEXT JJ: NEXT II
1640 GOSUB 2960: FO=FX
1650 FOR K=1 TO 5
1660 XF(K)=X(K): NEXT K
1670 IF(NP-1) = 0 THEN 1700
1680 FOR I=1 TO 5
1690 E(I)=EP(I): NEXT I
1700 FOR I=1 TO 5
1710 FB=FX: AJ(I)=2!
1720 IF (NP-1) = 0 THEN 1740
1730 GOTO 1750
1740 E(I)=EP(I)
1750 D(I)=0!: NEXT I
1760 III=0
1770 III=III+1
1780 I=1
1790 FOR J=1 TO 5
1800 X(J)=X(J)+E(I)*V(I,J)
1810 NEXT J=: GOSUB 2960
1820 KF=KF+1: DF=FB-FX
1830 IF(ABS(DF)-EY) i= 0 THEN 2520
1840 IF(KF-MF) z= 0 THEN 2520
1850 IF(FX-FO) i= 0 THEN 1920
1860 FOR J=1 TO 5
1870 X(J)=X(J)-E(I)*V(I,J)
1880 NEXT J: E(I)=-BA*E(I)
1890 IF(AJ(I)-1.5) z= 0 THEN 1910
1900 AJ(I) =0!
1910 GOTO 1980
1920 D(I)=D(I)+E(I)
1930 E(I)=AA*E(I): FO=FX
1940 FOR K = 1 TO 5
1950 XF(K)=X(K): NEXT K
1960 IF(AJ(I)-1.5) i= 0 THEN 1980
1970 AJ(I)=0!
1980 FOR J=1 TO 5
1990 IF(AJ(J)-5) z= 0 THEN 2010
2000 NEXT J: GOTO 2070
2010 IF(I-5) = 0 THEN 2030
2020 I=I+1: GOTO 1790
2030 FOR J=1 TO 5
2040 IF(AJ(J)-2!) I 0 THEN 1780

```

```

2050 NEXT J: IF(III-MC) i 0 THEN 1770
2060 GOTO 2520
2070 FOR I=1 TO 5
2080 FOR J=1 TO 5
2090 AL(I,J)=0!
2100 NEXT J: NEXT I
2110 REM IMPRESIÓN DE RESULTADOS PARCIALES
2120 IF B$="N" THEN 2230
2130 LPRINT: LPRINT"ETAPA NUMERO: "; KE
2140 LPRINT"VALOR DE LA FUNCION OBJETIVO ="; FO
2150 LPRINT"PARAMETROS DE LA POBLACION NO CICLONICA (m3/s): "
2160 LPRINT"PARAMETRO DE UBICACIÓN a1 ="; X(1)
2170 LPRINT"PARAMETRO DE ESCALA c1 ="; X(2)
2180 LPRINT"PARAMETROS DE LA POBLACION CICLONICA (m3/s): "
2190 LPRINT"PARAMETRO DE UBICACIÓN a2 ="; X(3)
2200 LPRINT"PARAMETRO DE ESCALA c2 ="; X(4)
2210 LPRINT"PROBABILIDAD (p) DE EVENTOS NO CICLONICOS ="; X(5)
2220 REM SE EFECTUA LA ROTACION DE EJES
2230 CLS: LOCATE 12,10: PRINT"CALCULOS EN PROCESO"
2240 FOR I=1 TO 5: KL=I
2250 FOR J=1 TO 5
2260 FOR K= KL TO 5
2270 AL(I,J)=AL(I,J)+D(K) *V(K,J)
2280 NEXT K: BL(I,J)=AL(I,J)
2290 NEXT J: NEXT I: BN(1)=0!
2300 FOR K=1 TO 5
2310 BN(1)=BN(1)+BL(1,K)^2: NEXT K
2320 BN(1)=SQR(BN(1))
2330 FOR J=1 TO 5
2340 V(1,J)=BL(1,J)/BN(1): NEXT J
2350 FOR I=2 TO 5: II=I-1
2360 FOR J=1 TO 5: SV=0!
2370 FOR KK=1 TO II: S=0!
2380 FOR K=1 TO 5
2390 S=S+AL(I,K)*V(KK,K): NEXT K
2400 SV=SV+S*SV(KK,J): NEXT KK
2410 BL(I,J)=AL(I,J)-SV
2420 NEXT J: NEXT I
2430 FOR I=2 TO 5: BN(I)=0!
2440 FOR K=1 TO 5
2450 BN(I)=BN(I)+BL(I,K)^2: NEXT K
2460 BN(I)=SQR(BN(I))
2470 FOR J=1 TO 5
2480 V(I,J)=BL(I,J)/BN(I)
2490 NEXT J: NEXT I
2500 KE=KE+1: IF(KE-ME) i 0 THEN 1700
2510 REM INICIA IMPRESIÓN DE RESULTADOS
2520 LPRINT: LPRINT"NO.TOTAL DE ETAPAS: "; KE
2530 LPRINT"NO.DE EVALUACIONES DE LA FUNCION OBJETIVO ="; KF
2540 LPRINT"VALOR FINAL DE LA FUNCION OBJETIVO ="; FO
2550 LPRINT: LPRINT"VALORES FINALES DE LOS PARAMETROS DE AJUSTE: "
2560 LPRINT"PARAMETROS DE LA POBLACION NO CICLONICA(m3/s): "
2570 LPRINT"PARAMETRO DE UBICACIÓN a2 ="; X(3)
2580 LPRINT"PARAMETRO DE ESCALA c1 ="; X(2)
2590 LPRINT"PARAMETROS DE LA POBLACION CICLONICA (m3/s): "
2600 LPRINT"PARAMETRO DE UBICACIÓN a2 ="; X(3)

```


BIBLIOGRAFIA

BOLETÍN #17 B
SUBCOMITÉ DE HIDROLOGIA

FRECUENCIA DE FLUJO DE DILUVIO
COMITE ASESOR EN DATOS DE AGUA

COORDINACION DE DATOS DE AGUA
RESTON, VIRGINIA

MANUAL DE HIDRAULICA URBANA
TOMO 1
DIRECCION GENERAL DE CONSTRUCCION Y OPERACION HIDRAULICA
TEORIA GENERAL
1982

MANUAL DE INGENIERIA DE RIOS
ESTUDIO HIDROLOGICO PARA OBRAS DE PROTECCION
CAPITULO 3
COMISION NACIONAL DEL AGUA