

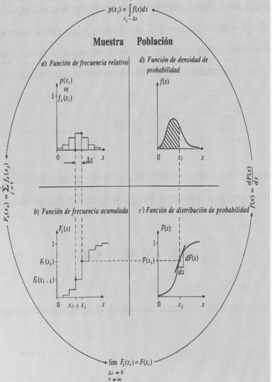
1. ***INTRODUCCIÓN***

El fenómeno del cambio climático genera muchas inquietudes acerca del comportamiento de algunas variables; es por esto que, variables hidrológicas como las precipitaciones y los caudales, muestran un papel fundamental en este aspecto, debido a que son variables directamente influenciables por fenómenos de amplio espectro temporal y espacial.

Por otra parte, para solucionar los problemas que implican el diseño de obras y la planificación hidrológica, se debe recurrir al estudio de la probabilidad, dado que dichos problemas se refieren a eventos que se podrían producir en el futuro, sin base en una estimación real, (Linsley *et. al.,* 1988). Sin embargo no solamente se debe estimar la magnitud del diseño, también se debe indicar la probabilidad de excedencia, con el fin de dar un cierto grado de seguridad a la obra, o bien el riesgo de falla, (Muñoz, 2004). De esta forma es necesario construir un modelo probabilístico, en donde se debe contar con una función de distribución de probabilidad (FDP), que represente la variable hidrológica de interés.

Las crecidas corresponden a fenómenos de concentración. Según Ollero (1996), son procesos naturales, sin periodicidad, constituidos por un incremento importante y repentino del caudal en un sistema fluvial, el cual lleva consigo un ascenso del nivel de la corriente, que puede desbordar el cauce menor para ocupar progresivamente el cauce mayor, hasta alcanzar un máximo, o caudal-punta y descender a continuación. Según Paoli *et al.* (1998), las crecidas que se presentan en términos hidrológicos, poseen un grado de riesgo y una probabilidad de excedencia diferente, según el caudal máximo, el volumen o la duración que se considere. Como consecuencia de ello las obras y medidas no estructurales que se disponen, estarán sometidas a un nivel de riesgo diferente según la variable utilizada, para determinar valores de diseño correspondientes a un determinado período de retorno.

En este marco, el presente documento técnico basado en los contenidos del segundo curso del proyecto FONDEF D08I1054, tiene como fin realizar un aporte al conocimiento de las funciones de distribución de probabilidad que mejor ajustan al comportamiento de eventos extremos y en particular de la variable caudales máximos instantáneos (caudales punta), a través de un ejemplo práctico, en donde se utilizarán las 4 funciones más utilizadas en estudios de variables hidrológicas, a saber Gumbel, Goodrich, Log-Normal y Pearson Tipo III, adicionalmente y a través de pruebas de bondad de ajuste como son el el coeficiente de determinación R cuadrado y el test de Kolmogorov Smirnov, se podrá realizar un análisis comparativo y una proyección de posibles eventos extremos asociados a diferentes períodos de retorno de los caudales punta, obtenidos de la información fluviométrica proporcionada por la Dirección General de Aguas.

1. ***FUNCIONES DE DISTRIBUCIÓN DE PROBABILIDAD (FDP)***
   1. ***Definición de las FDP***

Según Chow *et al.* (1994), al conjunto de observaciones *x1, x2, . . . , xn,* de la variable aleatoria, se denomina *muestra.* Una muestra es sacada de una población hipotéticamente infinita, que posee propiedades estadísticas constantes. Las propiedades de una muestra pueden cambiar de una muestra a otra y, el conjunto de todas las muestras posibles que puede extraerse de una población, se conoce como *espacio muestral*, en donde un evento es un subconjunto muestral. Si las observaciones de una muestra están idénticamente distribuidas, éstas pueden ordenarse para formar un *histograma de frecuencia.* Ahora bien, si el número de observaciones *ni*, en el intervalo *i* que cubre un cierto rango, se divide por el número total de observaciones *n*, el resultado se conoce como *frecuencia relativa.* Asimismo, la suma de los valores de la frecuencia relativa hasta un punto dado, es la *función de frecuencia acumulada*, y en su límite, cuando n→∞ y ∆χ→0, se denomina *función de distribución de probabilidad* (F.D.P.). De igual forma, la derivada o incremento finito de la F.D.P., se conoce como función de densidad de probabilidad (f.d.p.).

**3.3.- Formas de Determinar la Probabilidad**

Pizarro y Novoa (1986), afirman que para conseguir definir la probabilidad implícita es preciso consignar dos conceptos previos, que son el período de retorno y la probabilidad de excedencia.

Período de retorno: Se define como el tiempo que transcurre entre dos sucesos iguales. Sea ese tiempo T.

Probabilidad de excedencia: Es la probabilidad asociada al período de retorno.

P (x>X) = 

En otras palabras la probabilidad de que la variable aleatoria tome un valor igual o inferior a cierto número X, está dada por la función de distribución de probabilidad F(X).

considerando que 



Luego, la probabilidad de que x sea mayor que X, viene a estar dada por la función complementaria.



De esta forma es importante definir las Funciones de Distribución de Probabilidad que mejor se ajustan al comportamiento de las variables hidrológicas, donde no se considerará para este ejemplo la función Normal, la cuál si bien es el modelo más utilizado y con mayor importancia en el campo de la estadística, su uso es muy limitado en hidrología, dado que las variables raramente se comportan de esta forma (Varas y Bois, 1998).

Por otra parte existen otros modelos que representan de mejor forma el comportamiento de variables hidrológicas, un ejemplo de esto es la Ley de ***Distribución de Gumbel*** la cuál ha demostrado poseer una adecuada capacidad de ajuste, a valores máximos de caudales, (Pizarro y Novóa 1986). Así mismo y según Aparicio 1997, la función de distribución de probabilidad de Gumbel se comporta de la siguiente forma:

**Distribución Gumbel:**

(Ecuación 1)

Donde:

χ: Representa el valor a asumir por la variable aleatoria

e: Constante de Neper.

μ y d: Parámetros

Los parámetros de la distribución de una muestra de tamaño infinito, tienden a los siguientes valores, en base a la media aritmética y la desviación estándar de la muestra:

 ; 

Otra FDP conocida en hidrología es la función de ***Distribución de Goodrich***, la que posee la cualidad de eliminar valores extremos, es decir aquellos cuya probabilidad de ocurrencia es muy pequeña. Por lo mismo, consigue suprimir las distorsiones que pueda provocar un sólo valor anómalo. Posee la siguiente función de distribución de probabilidad (Pizarro *et. al.,* 1993).

**Distribución de Goodrich:**

(Ecuación 2)

Para X1< X ≤ ∞

En tanto los parámetros se determinan a partir del siguiente sistema de ecuaciones:

 ;

 ;



Donde:

m3 : Momento central de orden tres, 

S3 : Desviación típica al cubo.

P(p): Función auxiliar de Goodrich.

S2 : Varianza muestral.

Г : Función Gamma.

 : Media muestral.

e : Constante de Neper

Finalmente, despejando la variable aleatoria x de la función de distribución de probabilidad de Goodrich, se obtiene lo siguiente:



Una tercera FDP importante de considerar importante es la ***Distribución Log-Normal***, la cuál tiene la desventaja sobre la distribución normal de que está limitada a (X>0) y también que la transformación Log tiende a reducir la asimetría positiva comúnmente encontrada en la información hidrológica, debido a que al tomar los logaritmos, se reduce una proporción mayor de los números grandes en relación a los pequeños (Chow *et. al.* 1994). Presenta la siguiente función de distribución de probabilidad:

**Distribución Log-Normal:**

(Ecuación 3)

Donde los parámetros existentes que se basan en los logaritmos de la variable aleatoria, están definidos de la siguiente forma:





Donde:

χ: Representa el valor a asumir por la variable aleatoria

*α, β:* Parámetros

e: Constante de Neper

En el mismo caso que la distribución normal, se asigna z como una variable estandarizada:



Y la probabilidad se encuentra en la tabla Normal, donde el valor de la variable x es:

 (Ecuación 3 modificada)

La última Función a estudiar es la ***Distribución Pearson Tipo III***, la cuál se aplicó por primera vez en la hidrología por Foster (1924), para poder describir la probabilidad de caudales máximos anuales. Cuando la información es muy asimétrica positivamente, se utiliza una transformación Log para reducir su asimetría, la cual se presenta de la siguiente forma (Chow *et. al.*, 1994):

**Distribución Pearson Tipo III:**

 (Ecuación 4)

Donde los parámetros de la distribución pueden ser estimados en función del promedio () y desviación estándar (S) de la muestra, por medio de las siguientes expresiones:

 ;  ; 

Donde:

 : Coeficiente de sesgo

e : Constante de Neper

*α,* *β* y *δ*: parámetros

S: Desviación típica

: media aritmética

Asimismo la variable estandarizada y se presenta a continuación:



Posteriormente, el ajuste, se realiza a través de la tabla chi-cuadrado, donde:





Por lo tanto, el valor que asume la variable aleatoria x a partir de lo anteriormente señalado, se define como:

**** (Ecuación 4 modificada)

Y la probabilidad es obtenida a través de los valores presentes en la tabla de percentiles de la distribución x2, con n grados de libertad.

Los resultados del estudio de Kroll y Vogel (2002, en 1505 estaciones de Estados Unidos, determinan que la función de Pearson Tipo III, es la que mejor representa a las series de caudales mínimos intermitentes, donde se presentan descargas con valores cero.

Al contar con los ajustes necesarios de estas 4 FDP se desea representar en este trabajo cuál es la que mejor representa la tendencia y comportamiento de los caudales punta, para esto se presenta la siguiente metodología de trabajo:

1. ***METODOLOGÍA***

El ejemplo a considerar contempla la estación ***Río Maipo en El Manzano*** (Latitud Sur: 33°35’; Longitud Oeste: 70°22’), ubicada en la Región Metropolitana y para ejemplificar se considerará un período de registro que está comprendido entre los años 1993 – 2007 (15 datos anuales de caudales máximos instantáneos).



**Figura 1: Estación Fluviométrica**

***3.1. Tratamiento inicial de la información y cálculo de Estadígrafos***

Para cada serie de datos, se determinaron por una parte, los estadígrafos de posición, también llamados de tendencia central, para indicar alrededor de qué valor se agruparon los datos obtenidos, como es el caso de la media, y por otra parte, los estadígrafos de dispersión. Estos estadígrafos o estadísticos, extraen información de una muestra, indicando las características de la población:

* **Media µ**: Es el valor esperado de la variable misma o primer momento respecto al origen. Muestra la tendencia central de la distribución y su valor estimado a partir de la muestra es:



* **Desviación Estándar**: Es una medida de la variabilidad, ya que es la raíz cuadrada, de los cuadrados de las diferencias y, su valor estimado, se denota como:



En la siguiente tabla se entrega la información fluviométrica de la estación a trabajar y los estadígrafos media y desviación estándar necesarias para comenzar a ajustar las funciones:

**Tabla N°1:** **Información de Caudales Punta Estación Rio Maipo en El Manzano, válida para el cálculo de las FDP.**



***3.2. Cálculo de la FDP de Gumbel***

Posteriormente, con el resultado de los estadísticos media y desviación estándar, se procede a ajustar las series anuales de caudales máximos anteriormente presentados. Para esto es preciso determinar los parámetros de la primera función considerada, (Gumbel). Como antecedente, este ejemplo puede ser acotado para otras variables hidrológicas de igual forma, como por ejemplo son las precipitaciones máximas.

*Determinación de* parámetros μ y *d*:

**S**  = desviación estándar de la muestra

 = media de la muestra.

**Tabla N°2**: **Parámetros Gumbel**

|  |
| --- |
| D |
| **0,004590323** |
| μ |
| **306,0642297** |

Una vez determinados los parámetros de Gumbel, se debe ajustar la FDP para esto, primero se debe ordenar la variable aleatoria en forma creciente, posteriormente se debe determinar la frecuencia observada acumulada Fn(x) y la frecuencia teórica acumulada F(x). Las frecuencias observadas se calculan ordenando en forma ascendente la siguiente expresión de Weibull:

 (Ecuación 5)

Donde:

**Fn(X)** = Frecuencia Observada Acumulada.

**n** = Número del dato n.

**N** = Número total de datos.

Por su parte, la frecuencia teórica acumulada de la función de Gumbel se obtiene a través de la siguiente expresión, donde se consideran los parámetros y la variable ordenada que corresponda para los 15 datos ejemplificados:



Para finalizar con el ajuste, se debe comprobar la calidad del ajuste, para esto se contrastaran 2 pruebas **el Test de Kolmogorov-Smirnov** (test K-S) y el **Coeficiente de Determinación (R2).**

El primero se calcula mediante la obtención del supremo de las diferencias (ver tabla 3), que consiste en determinar el valor absoluto de la máxima diferencia entre las frecuencias observadas y acumuladas. Esta diferencia se denomina por la letra Dc y su expresión es la siguiente:

(Ecuación 6)

Donde:

Dc = Supremo de las Diferencias.

Fn(X)i = Frecuencia Observada Acumulada.

F(X)i = Frecuencia Teórica Acumulada.

Una vez obtenido el supremo de las diferencias, se compara con el valor de la tabla Kolmogorov-Smirnov. Si el valor obtenido de la tabla K-S (Dt), (Tabla Anexa N°1) es mayor que el supremo de las diferencias (Dc), se puede aceptar la hipótesis nula (Ho) que indicaría que se está en presencia de un buen ajuste con el nivel de confianza asumido (Dt > Dc).

Por otra parte para calcular el Coeficiente de Determinación (R2) se debe considerar la ecuación 7. Dicho Coeficiente es un indicador que mide cuál proporción o porcentaje de la variación total de la variable dependiente, es explicada por el modelo de regresión (Gujarati, 1992):

**** (Ecuación 7)

Donde:

R2 : Coeficiente de determinación; 0≤R2≤1

: Media de las frecuencias observadas acumuladas

: Frecuencia observada

: Frecuencia teórica acumulada

A continuación, se presenta en la **Tabla N°3** el cálculo de la frecuencia observada acumulada Fn(x), la frecuencia teórica acumulada y el Supremo de las diferencias que se explica con posterioridad:

**Tabla N°3. Ejemplo para el ajuste de la FDP de Gumbel**



Por otra parte, en la Tabla 4 se presentan los resultados de las pruebas de bondad del ajuste (Test Kolmogorov-Smirnov y Coeficiente de Determinación), para el ejemplo de la función de Gumbel desarrollado.

**Tabla N°4. Resultados tests de bondad del ajuste Gumbel Gumbel**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **FDP Ajustada** | **Dc** | **Dt** | **Ajuste K-S** | **R2** |
| Gumbel | 0,139 | 0,338 | Acepta Ho | 0,951 |
|  |  |  |  |  |

Donde:

**A** = Se acepta el ajuste.

**Dc** = Supremo de las diferencias.

**Dt** = Valor tabla K-S.

**R2** = Coeficiente de determinación.

Los resultados de esta tabla indican que el valor Dc=0,139, es menor que el valor Dt=0,338, obtenido de la tabla K-S con un 95% de confianza ( = 0,05 con n = 15), con lo cual se cumple la condición de un buen ajuste K-S.

Por otra parte el Coeficiente de Determinación R2 alcanza un 95,1%, lo que también indica un buen ajuste del modelo, por lo que se puede indicar que la FDP Gumbel es una Distribución que se ajusta bien a los datos de caudales presentados en el ejemplo.

La forma de calcular el R2, para que quede mejor ejemplificado y de acuerdo a la Tabla 3 es la siguiente:

****

Como último punto relacionado con la FDP Gumbel y para ejemplificar de mejor forma el ejercicio, es necesario conocer el valor a asumir por la variable aleatoria (caudal, precipitación etc.) para cierto período de retorno asociado, de tal manera de poder predecir posibles eventos futuros y poder tomar decisiones de gestión. Para esto se pretende despejar dicha variable y calcular su valor asociado a 10, 20 y 50 años. El despeje se presenta a continuación:



Al despejar el valor de la variable aleatoria de la función original, se obtiene lo siguiente:

****

Donde:

χ: Representa el valor a asumir por la variable aleatoria

e: Constante de Neper

µ y d: Parámetros

***3.3. Cálculo de la Probabilidad de Excedencia***

En la Tabla N°5, se entregan los valores de caudales máximos asociados a los períodos de retorno nombrados anteriormente, de acuerdo a los valores despejados de la variable x y a los parámetros asociados. Para esto se debe considerar que la probabilidad de excedencia: Es la probabilidad asociada al período de retorno, y viene dada por la siguiente expresión:

.

P (x>X) = 

En otras palabras la probabilidad de que la variable aleatoria tome un valor igual o inferior a cierto número X, está dada por la función de distribución de probabilidad F(X).



Por ejemplo para un período de retorno (T) de 30 años:



El valor de la variable aleatoria asumiendo ese período debiera ser:

****

Lo que indicaría que considerando dicha serie de datos y un período de retorno de 30 años, el valor del caudal punta podría adquirir un valor igual o inferior a:

1039 

Por lo tanto considerando los períodos de retorno 10, 20 y 50 años para el ejercicio se presenta la siguiente tabla de resultados:

**Tabla N°5: Probabilidad de Caudales Máximos para los Distintos Períodos de Retorno según Gumbel.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Período de** | **Probabilidad** | **Probabilidad** | **Gumbel** |
| **retorno** | **de excedencia** | **de no excedencia** |  |
| 10 | 0,100 | 0,900 | 796 |
| 20 | 0,050 | 0,950 | 953 |
| 50 | 0,020 | 0,980 | 1156 |

De esta forma se ha trabajado con la FDP Gumbel, por lo tanto se procede a realizar todo el ejercicio, desde la obtención de parámetros, los ajustes, la comprobación de la calidad del ajuste y el valor estimado de la variable asociada a diversos períodos de retorno, para las 3 FDP restantes:

***3.4. Cálculo de la FDP de Goodrich***

Con la serie de datos original pertenecientes a la Tabla N°1, la cual se presenta a continuación se procede a calcular los parámetros de esta FDP.

**Tabla N°1:** **Información de Caudales Punta Estación Rio Maipo en El Manzano, válida para el cálculo de las FDP.**



Para poder realizar los ajustes respectivos de la FDP Goodrich:



Es necesario contar con los parámetros,  y  , de los cuáles su forma de obtener se detalla a continuación:







S3 : Desviación típica al cubo.

P(p): Función auxiliar de Goodrich.

S2 : Varianza muestral.

Г : Función Gamma.

 : Media muestral.

Para obtener , primero es necesario conocer (momento de orden tres), el cual se define como la sumatoria de los desvíos de la serie de datos, con respecto a su respectiva media, elevado al cubo y cuya sumatoria se divide por el número total de datos:



Para este ejemplo dicha sumatoria se debe calcular de la siguiente forma:

 Es necesario obtener , para esto se divide por (Desviación estándar a cubo):



Con el valor de  se obtiene en la Tabla denominada Valores de la Función Auxiliar de Goodrich que se presenta al final de este documento, y corresponde a la Tabla Anexa N°2, se obtiene, interpretado por el valor ubicado en la derecha de la Tabla.

Para el caso del ejercicio el valor obtenido es necesario interpolarlo ya que los valores de  se encuentran entre 0,737553 y 0,764037.

El segundo parámetro a calcular es , el cual se obtiene de:



Para esto se debe consideran S2, que corresponde a la Varianza muestral y (Función Gamma), esta última se obtiene teniendo presente lo siguiente:

( *p* ) = ( *p* −1)!

 ( *p* +1) = *p*!= *p*( *p* )

 ( *p* −1) = ( *p* − 2) ( *p* − 2)

Se utiliza la tabla de valores de la función Gamma, que también se encuentra al final del capítulo, en este caso es la la Tabla Anexa N°3, donde se debe entrar por la parte izquierda, y ubicando el segundo decimal, se determina la fila y columna que expresa el valor a asumir por la función.

Por lo tanto, en algunas ocasiones no será posible aplicar directamente la tabla, por lo cual se debe hacer una transformación, la que en términos generales puede plantearse como:

( *p* + m) = ( *p* + m −1)( *p* + m − 2)( *p* + m − 3)............( *p* + m − n)r( *p*)

con 1 *p*  1.99

Un ejemplo para aclarar cómo se obtienen los valores , que es la parte más compleja dentro de la expresión, es el siguiente:

Si :

Para  entonces al buscar en la tabla de valores de la función Gamma () de 1,6, se obtiene , por lo tanto .

Y para este caso entonces como el valorde 2,2 no se encuentra en la tabla se procede a descomponer de esta forma: 

Por lo tanto para despejar la ecuación final queda de la siguiente forma:



Como último punto para despejarsolo se debe considerar los valores ya obtenidos anteriormente e incluir : Media muestral del total de los datos.



A continuación se presentan losparámetros , y .

**Tabla N° 6: Parámetros Goodrich**



De acuerdo a la **ecuación 2**, y considerando los parámetros calculados recientemente la frecuencia Teórica queda expresada de la siguiente forma:



La frecuencia relativa se obtiene de igual forma que en el ejercicio anterior, considerando:

 (Ecuación 5)

Donde:

**Fn(X)** = Frecuencia Observada Acumulada.

**n** = Número del dato n.

**N** = Número total de datos

A continuación, en la Tabla 7 se ejemplifica el cálculo de la frecuencia observada acumulada Fn(x), frecuencia teórica acumulada F(x), y el Supremo de las diferencias Dc, tal como en el ejemplo anterior.

**Tabla N°7. Ejemplo para el ajuste de la FDP de Goodrich**



En la Tabla 8 se presentan los resultados de las pruebas de bondad del ajust (Test Kolmogorov-Smirnov y Coeficiente de Determinación). Los cuales se obtienen de las ecuaciones 7 y 8.

Para la calidad del ajuste en la tabla 8 se observa el **test de Kolmogorov-Smirnov** (test K-S) donde debe ser comparado el Dc máximo con el Dt igual como en el ejemplo de Gumbel a través de las ecuación 6, esto se realiza al nivel de confianza del 95% y con un n observado de 15 datos.

Por otra parte, para conocer el porcentaje de los datos observados que son explicados por el modelo en este caso el **Coeficiente de Determinación (R2)** se utiliza la ecuación 7, de igual forma que el ejercicio anterior.

**Tabla 8. Resultados tests de bondad del ajuste Goodrich**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **FDP Ajustada** | **Dc** | **Dt** | **Ajuste K-S** | **R2** |
| Goodrich | 0,116 | 0,338 | Acepta Ho | 0,97 |
|  |  |  |  |  |

Los valores obtenidos de K-S muestran un Dc=0,116, menor que el valor Dt=0,338, aceptándose la Hipótesis Nula de un buen ajuste. Para el caso del Coeficiente de Determinación R2 se presenta un 97%, lo que también indica un buen ajuste del modelo. Por lo que el ajuste de Goodrich es óptimo al igual que el caso de Gumbel. Los valores mencionados anteriormente se presentan a continuación:

Por otra parte para conocer el valor asumido por la variable aleatoria asociado a los períodos de retorno estudiados 10, 20 y 50 años, se realiza el despeje de la siguiente FDP:



Finalmente, despejando la variable aleatoria x de la función de distribución de probabilidad de Goodrich, se obtiene lo siguiente:



Donde:

χ: Representa el valor a asumir por la variable aleatoria.

,  y : Parámetros obtenidos anteriormente.

F(X): Probabilidad de no excedencia asociada al período de retorno:



***3.5. Cálculo de la Probabilidad de Excedencia***

Se procede a entregar la Tabla N°9, con los valores de caudales máximos asociados a los períodos de retorno, de acuerdo al despeje de la variable x, junto con los parámetros asociados a la función de Goodrich.

**Tabla N° 9: Probabilidad de Caudales Máximos para los Distintos Períodos de Retorno según Goodrich.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Período de** | **Probabilidad** | **Probabilidad** | **Goodrich** |
| **retorno** | **de excedencia** | **de no excedencia** |  |
| 10 | 0,100 | 0,900 | 812 |
| 20 | 0,050 | 0,950 | 947 |
| 50 | 0,020 | 0,980 | 1105 |

Estos valores son obtenidos del despeje anterior y se asocian a los períodos de retorno nombrados. Como referencia se puede considerar el ejemplo de Gumbel, donde se muestra como al despejar la variable aleatoria con la probabilidad asociada se puede estimar dichos valores de caudales para el futuro.

A continuación se presentan los resultados de la tercera Función considerada Log-Normal:

**3.6. Cálculo de la FDP Log-Normal**

,

En primera instancia se presenta nuevamente la Tabla 1, la cual conserva los estadígrafos media y desviación estándar por ser la base con los mismos 15 datos para el ejercicio, los cuáles son considerados para calcular los nuevos parámetros.

**Tabla N°1:** **Información de Caudales Punta Estación Rio Maipo en El Manzano, válida para el cálculo de las FDP.**



A continuación se presenta la FDP Log-Normal correspondiente a la ecuación 3:



Los parámetros necesarios para este caso son los que se basan en los logaritmos de la variable aleatoria, los cuales son los siguientes:



Para el caso de  cada valor de la variable ordenada se le calcula el ln de cada dato, se sumay se divide por el n° de datos que es en este caso 15, por lo tanto el ejercicio partiría de la siguiente forma:



El segundo parámetro de esta FDP es:



El cálculo se realiza de forma similar pero esta vez se considera la diferencia los cuadrados de los ln y (obtenido anteriormente) dividida por n (15), en sumatoria, al final a todo este valor se le aplica raíz o se eleva a ½.

Por ejemplo si = 4,5 en este caso el cálculo sería de la siguiente forma:



A continuación se presentan los valores obtenidos de los parámetros yen el ejemplo de la FDP Log-Normal.

**Tabla N° 10: Parámetros Log-Normal**

|  |  |
| --- | --- |
|  | 5,852 |
|  | 0,689 |

De acuerdo a la **ecuación 3**, considerando los parámetros obtenidos la F(X) es la siguiente:



Pero de acuerdo a la complejidad de la función que considera una integral se debe estandarizar la expresión quedando expresada de la siguiente forma:



Para este caso incluyendo los parámetros de la función se obtiene lo siguiente:



Donde el cálculo de la frecuencia teórica F(X), se obtiene del valor z, el cual se busca en la tabla normal estándar (Tabla Anexa N°4). Si el z es positivo la Frecuencia Teórica es el valor directo que se entrega en dicha tabla, pero si z es negativo el valor F(X) a asumir es 1 – el valor de z, pero dicho z debe ir en valor absoluto. Un ejemplo de lo anterior es lo siguiente:

Si z = -1,95, entonces F(X) = 1 – z(1,95), donde z(1,95) =.0,9744, por lo tanto F(X) = 1- 0,9744 = 0,0256.

Por otra parte la frecuencia relativa se obtiene de igual forma que en los ejercicios anteriores, considerando:

 (Ecuación 5)

Donde:

**Fn(X)** = Frecuencia Observada Acumulada.

**n** = Número del dato n.

**N** = Número total de datos.

A continuación, en la Tabla N°11 se presenta el cálculo de la frecuencia observada acumulada Fn(x), la frecuencia teórica el Dc con el supremo de dichas diferencias, para la Función Log-Normal, de igual forma como los ejercicios anteriores.

**Tabla N°11. Ejemplo para el ajuste de la FDP Log-Normal**



Posteriormente la Tabla 12 presenta los resultados de las pruebas de bondad del ajuste (Test Kolmogorov-Smirnov y Coeficiente de Determinación). Estos se calculan de la misma forma que en los ejercicios anteriores considerando las ecuaciones 7 y 8, a partir de los datos obtenidos de la Tabla 11.

**Tabla 12. Resultados tests de Bondad de Ajuste Log-Normal**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **FDP Ajustada** | **Dc** | **Dt** | **Ajuste K-S** | **R2** |
| Log-Normal | 0,130 | 0,338 | Acepta Ho | 0,938 |
|  |  |  |  |  |

En la comparación del **test de Kolmogorov-Smirnov** (Test K-S) el Dc máximo con el Dt al nivel de confianza del 95% y con un n observado de 15 datos para el caso de Log-Normal se obtiene lo siguiente: Dc = 0,116 es menor que el Dt = 0,338, por lo que el ajuste es aceptable.



Por otra parte el porcentaje de los datos observados explicados por el modelo a través del **Coeficiente de Determinación (R2)** entregan un 93,8%, lo que es un valor alto y aceptable también.

Para conocer el valor asumido por la variable aleatoria asociado a los 3 períodos de retorno estudiados es necesario conocer el despeje de la variable aleatoria x, la cuál de acuerdo a la complejidad de la Función original esta se realiza de la estandarización de z, obteniéndose la siguiente fórmula:

 (Despeje de X en la Ecuación 3 estandarizada)

Donde:

, = Parámetros obtenidos

= Constante de neper

Y donde Z = es el valor que se obtiene de la Tabla Anexa N°4, el cual se aproximo para este ejercicio, en este caso se ingresa con una probabilidad y se obtiene Z, por ejemplo:

Si se considera un período de retorno T = 30 años, entonces:

Por lo tanto al ingresar a la tabla Z, con una probabilidad de 0,96 se asume un Z = 0,9608 se obtiene un valor de Z aproximado de 1,76.

Entonces con todos estos valores se despeja la variable aleatoria, la cual se presenta a continuación.

***3.7. Cálculo de la Probabilidad de Excedencia***

Con la expresión anterior se procede a entregar la tabla N°13, que incluyen los valores de caudales máximos asociados a los períodos de retorno para la FDP Log-Normal, de acuerdo al despeje de la variable x.

**Tabla N° 13: Probabilidad de Caudales Máximos para los Distintos Períodos de Retorno según Log-Normal.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Período de** | **Probabilidad** | **Probabilidad** | **Log-** |
| **Retorno** | **de excedencia** | **de no excedencia** | **Normal** |
| 10 | 0,100 | 0,900 | 846 |
| 20 | 0,050 | 0,950 | 1085 |
| 50 | 0,020 | 0,980 | 1439 |

Estos valores se asocian a los períodos de retorno considerados de acuerdo a la cantidad de datos y al ajuste obtenido a través de Log-Normal

Finalmente se presentan los resultados de acuerdo a la FDP Pearson Tipo III.

***3.8. Cálculo de la FDP Pearson Tipo III***

Para ejemplificar de forma más clara la obtención de los parámetros de esta FDP, es necesario nuevamente considerar la Tabla 1, que representa la base original de la información.

**Tabla N°1:** **Información de Caudales Punta Estación Rio Maipo en El Manzano, válida para el cálculo de las FDP.**



Los parámetros necesarios para ajustar la distribución Pearson Tipo III, son,  y , los cuales pueden ser estimados en función del promedio () y desviación estándar (S) de la muestra,

pero en primera instancia se requiere calcular del coeficiente de sesgo, el cual presenta la siguiente fórmula:

 : Coeficiente de sesgo

Por lo tanto a partir del coeficiente de sesgo, se pueden calcular los restantes parámetros, ahora se ejemplifica la forma de calcularlo de acuerdo a los datos que se tienen considerados:



Con el valor de  se obtienen los parámetros a finales los cuales se entregan en las siguientes expresiones:

*α,*  y *δ*: Parámetros directos de la función

S: Desviación estándar

: Media aritmética

A continuación se presentan los valores obtenidos de los parámetros yen el ejemplo de la FDP Log-Normal.

**Tabla N° 14: Parámetros Pearson Tipo III**

|  |  |
| --- | --- |
|  | 104,841 |
|  | 7,102 |
|  | -312,766 |

De acuerdo a la **ecuación 4**, considerando los parámetros obtenidos se obtiene la F(X) siguiente:



Para hacer menos compleja la FDP anterior se presenta estandarizada de la siguiente forma:



En este caso incluyendo los parámetros seria:



Posteriormente con los valores de **** obtenidos de la expresión anterior, se procede a calcular la frecuencia teórica del ajuste, la cual corresponde a , con  grados de libertad, obtenidas de la Tabla de percentiles, de la Distribución Chi-Cuadrado (Tabla Anexa N°5). De no encontrarse el valor exacto, se aproxima al más cercano, de lo contrario se procede a interpolar.





Un ejemplo de cómo obtener dicha frecuencia teórica acumulada y considerando los datos originales es el siguiente:

Para X = 90,99 m3/s (primer dato de la serie ordenada), entonces:



Entonces el valor

Ahora

Por lo tanto se busca el valor en la Tabla anexa N°5 dey 14 grados de libertad, de lo que se obtiene que el valor de probabilidad exacto se presenta entre 6,57 correspondiente a un 0,05 y 7,79 correspondiente a 0,10, luego interpolando dicho valor se obtiene que para X = 90,99 m3/s la frecuencia teórica acumulada es 0,096. De esta forma se construye completamente el F(X) Pearson Tipo III.

Para el caso de la frecuencia observada o frecuencia relativa se conserva la fórmula:

 (Ecuación 5)

Donde:

**Fn(X)** = Frecuencia Observada Acumulada.

**n** = Número del dato n.

**N** = Número total de datos.

A continuación, en la **Tabla N°15** se presentan los valores de la frecuencia observada acumulada Fn(x), la frecuencia teórica el Dc con el supremo de dichas diferencias, tal como en el ejemplo anterior:

**Tabla N°15. Ejemplo para el ajuste de la FDP Pearson Tipo III**



Posteriormente la Tabla 16 presenta los resultados de las pruebas de bondad del ajuste (Test Kolmogorov-Smirnov y de).

Para esto nuevamente se calculan las ecuaciones 7 y 8, a partir de los datos obtenidos en este caso de la Tabla 11 y siguiendo el primer ejemplo.

**Tabla N°16. Resultados tests de bondad del ajuste Pearson Tipo III PearGoodrich**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **FDP Ajustada** | **Dc** | **Dt** | **Ajuste K-S** | **R2** |
| Pearson Tipo III | 0,118 | 0,338 | Acepta Ho | 0,965 |
|  |  |  |  |  |
|  |  |  |  |  |

Se comprueba la calidad del ajuste, con las ecuaciones 6 y 7 consideradas para las 2 pruebas ya utilizadas obteniéndose para **el test de Kolmogorov-Smirnov** (test K-S) un Dc = 0,118 menor el mismo Dt = 0,338 obtenido para todos ejemplos realizados, por lo que se acepta Ho.

Mientras que para el **Coeficiente de Determinación (R2)** se alcanza un 96,5 %, lo que referencialmente es bastante bueno también. Los resultados se presentan a continuación:

Para conocer el valor asumido por la variable aleatoria asociado a los 3 períodos de retorno estudiados es necesario conocer el despeje de X, la que también al ser una FDP compleja se le realiza un despeje pero con los valores de  obtenidos de la expresión asociada a la ecuación presentada anteriormente la cual representa a x de la siguiente forma:

**** (Despeje de la Ecuación 4 modificada)

Los valores de fueron los parámetros obtenidos anteriormente, mientras que el caso de se obtiene de la ecuación:

Por ejemplo si se considera se considera un período de 40 años:



A lo que le correspondería según la Tabla Anexa N°5:



Asumiendo los 14 grados de libertad, por lo tanto el valor de:



De esta forma se obtienen los parámetros para poder calcular el valor a asumir por la variable aleatoria que para la probabilidad de excedencia se consideran los períodos de retorno 10, 20 y 50 años al igual que en los ejercicios anteriores.

***3.9. Cálculo de la Probabilidad de Excedencia***

Con la expresión anterior se procede a entregar la tabla N°17, que incluyen los valores de caudales máximos asociados a los períodos de retorno para la FDP Pearson Tipo III, de acuerdo al despeje de la variable x, según la ecuación anterior y considerando las ecuaciones , 

**Tabla N° 17: Probabilidad de Caudales Máximos para los Distintos Períodos de Retorno según Pearson Tipo III.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Período de** | **Probabilidad** | **Probabilidad** | **Pearson** |
| **retorno** | **de excedencia** | **de no excedencia** | **Tipo III** |
| 10 | 0,100 | 0,900 | 793 |
| 20 | 0,050 | 0,950 | 930 |
| 50 | 0,020 | 0,980 | 1108 |

Para finalizar se presenta como cuadro resumen los Testsde Bondad de Ajuste de las 4 FDP y se procede a elegir la que mejor representa la tendencia de los datos considerados en el ejercicio:

**Tabla N°18: Resumen FDP y test de bondad de ajuste**



Por lo tanto la Función que entrega una diferencia mayor entre el valor de tabla y el valor crítico supremo en la prueba **K-S**, es Goodrich, lo que estaría indicando una leve inclinación a un mejor ajuste de los datos presentados, sobre todo si también se presenta el mayor **R2** de las 4 Funciones.

Estos resultados demuestra lo que algunos estudios indican como por ejemplo Blazkova y Beven (2003), determinan la probabilidad de excedencia asociada para los caudales a través de Goodrich, encontrando mayor claridad para la evaluación del grado de seguridad en el funcionamiento de obras para mitigar eventos extremos. Sin embargo otros autores siguen proponiendo el uso de las 4 funciones para ajustar valores extremos, dado que va a depender del espacio geográfico donde se esté estimando, la variabilidad, la cantidad de datos con que se cuente, entre otros.

**REFERENCIAS BIBLIOGRÁFICAS**

Aguilera, M.A. 2007. Análisis de las funciones de distribución de probabilidad para caudales máximos en la Región del Maule. Talca, Chile. Universidad de Talca, Facultad de Ciencias Forestales. 86 p.

Aparicio, F. 1997. Fundamentos de Hidrología de Superficie. 11 ed. México. Editorial Limusa S.A. 303 p.

Ashakar, F.; T.B.M.J Ouarda, R. Roy and B. Bobée. 1993. Robust estimators in hydrologic frequency analysis, in Engineering Hydrology. Edited by C.Y, Am. Soc. Civ. Eng. 347-352 p.

Blazkova, S.; Beven, K.; 2003. Flood frequency estimation by continuos simulation of subcatchment rainfalls and discharges with the aim improving dam safety assessment in a large basin in the Czech Republic. Journal of hydrology. 294 (1-4) 153-172. Consultado 03 Junio 2008. Disponible en base de datos Science Direct.

Casassa, G.; Echelmeyer, K.; Neumman, T.; Raymond, Ch.; Rignot, E.; Rivera, A. 2005. Reatreat of Glaciar Tyndall, Patagonia, over the last half-century. Journal of Glaciology. 51 (1) 239-247. Consultado 28 Jul. 2008. Disponible en [**http://www.glaciologia.cl/textos/raymd.pdf**](http://www.glaciologia.cl/textos/raymd.pdf)

Chavarri, E. 2004. Curso IA-4026, Hidrología Aplicada. (en línea). Consultado 19 nov. 2007. Disponible en [http://tarwi.lamolina.edu.pe/ ~echavarri/echv\_7.html](http://tarwi.lamolina.edu.pe/%20~echavarri/echv_7.html)

Chow, V.; Maidment, D.; Mays, L. 1994. Hidrología aplicada. Colombia. Editorial McGraw-Hill. Interamericana S.A. 584 p.

Combes, S.; Prentice, M.; Hansen, L.; Rosentrater, L. 2002. Climate change and global glacier decline. (en línea). Consultado 06 diciembre 2007. Disponible en: <http://www.wwf.fi/wwf/www/uploads/pdf/glaciersraportti.pdf>

Foster, H. 1988. Theoretical frequency curves and their application to engineering problems. Trans. Am. Soc. Civ. Eng. 87 (2) 142-173.

Kroll, C.; Vogel, R. 2002. The probability distribution of low streamflow series in the Unit United States. Journal of Hydrologic Engineering. EEUU. ASCE, 7(2), 137-146.

Linsley, R.; Kohler, M.; Paulus, J. 1988. Hidrología para ingenieros. 2 ed. México. Editorial McGraw-Hill Interamericana S.A. 386 p.

Llamas. 1993. Hidrología General. Edición española. Servicio Editorial Universidad del País Vasco. 635 p.

Miller N., Bashford K. ,Strem E. 2003. Potential Impacts of Climate Change on California hydrology Journal of the American Water Resources Association 39 (4), 771–784. Consultado 13 junio 2008. Disponible en base de datos Dialnet.

Morales, C. M. 2005. Análisis de las escorrentías mensuales y anuales de la cuenca del Lontué y la potencial influencia glaciar en la producción de agua. Talca, Chile. Universidad de Talca, Facultad de Ciencias Forestales. 127 p.

Muñoz, M.F. 2004. Análisis de algunas de las variables hidrológicas y su ajuste a funciones de distribución de probabilidad, en tres cuencas de la Región del Maule. Talca, Chile. Universidad de Talca, Facultad de Ciencias Forestales. 86 p.

Ollero, A.1996. El curso medio del Ebro: geomorfología fluvial, ecogeografía y riesgos. (en línea). Consultado 26 Jun. 2008. Disponible en http://www.ingeba.euskalnet.net/lurralde/lurranet/lur20/ 200oller/ollero20.htm

Önoz, B.; Bayazit, M. T. 2001. Effect of the occurrence process of the peaks over threshold on the flood estimates. Journal of hydrology. 224 (1-2) 86-96.

Consultado 06 Abril 2008. Disponible en base de datos Sicence Direct.

Paoli C.; Cacik P.; Bolzicco J. 1998. Análisis de riesgo conjunto en la determinación de crecidas de proyectos de régimenes complejos. (en línea). Consultado 24 nov. 2007. Disponible en [https://upcommons.upc.edu/revistes/bitstream/2099/2986/1/52arti cle.pdf](https://upcommons.upc.edu/revistes/bitstream/2099/2986/1/52arti%20cle.pdf)

Pizarro, R.; González, P.; Wittersheim, M.; Saavedra, J.; Soto, C. 1993.

Elementos técnicos de Hidrología III. Proyecto regional sobre el uso y conservación del recurso hídrico en áreas de América Latina y el Caribe. Talca, Chile. Editorial Universidad de Talca. 136 p.

Pizarro, R.; Novoa, P. 1986. Elementos técnicos de Hidrología I. Instructivos técnicos. La Serena, Chile. Corporación Nacional Forestal, Ministerio de Agricultura. 57 p.

Pizarro, R.; Morales, C; Moran, L.; Vargas, J.; Sanguesa, C.; Godoy, C.; 2006. Informe País. Estado del Medio Ambiente en Chile 2005. Santiago. Chile. Capítulo 2. Aguas Continentales. Instituto de Asuntos Públicos, Universidad de Chile. Editorial LOM. 371 p.

Sullivan, A.; Ternan, J.; Williams, A. Land use change and hydrological response in the Camel catchment, Cornwall. Applied Geography. 24 (2) 119-137. Consultado 06 diciembre 2007. Disponible en base de datos Science Direct.

Varas, E.; Bois, P. 1998. Hidrología probabilística. Santiago. Chile. Editorial Universidad Católica de Chile. 156 p.

Waylen, P.; Woo , M. Prediction of Annual Floods Generated by Mixed Processes. Water Resources Research. 18 (4) 1283-1286. Consultado 03 Junio 2008. Disponible en base de datos CSA.

Yue, S.; Ouarda, T.; Bobée, B.; 2006, Legendre; P.; Bruneau, P. The Gumbel mixed model for flood frequency analysis. Journal of hydrology. 226 (1-2) 88-100. Consultado 06 diciembre 2007. Disponible en base de datos Sicence Direct.