5 HIDROLOGIA ESTATÍSTICA

5 - HIDROLOGIA ESTATÍSTICA

5.1 - Considerações Iniciais

Séries de variáveis hidrológicas como precipitações, vazões, evaporação e outras, quando observadas ao longo do tempo, apresentam variações sazonais. Estas variações não são entretanto absolutamente regulares. A observação de séries longas de dados hidrológicos revelará a ocorrência de extremos (máximos e mínimos) e diferentes sequências de valores, que caracterizam as variáveis hidrológicas como **aleatórias**. Variáveis hidrológicas estarão sempre associadas portanto a uma probabilidade de ocorrência. Consequentemente as obras hidráulicas devem sempre dimensionadas para um determinado "risco" de falha.

O objetivo da estatística é o de extrair informações significativas de uma dada massa de dados. As técnicas utilizadas em estatística aplicadas à hidrologia permitem avaliar a probabilidade de ocorrência de um fenômeno hidrológico com determinada magnitude.

5.2 - Conceito de Período de Retorno e Risco Permissível

O inverso da probabilidade de ocorrência é denominado em Hidrologia de **período de retorno** ou **intervalo de recorrência**.

Assim se uma determinada grandeza hidrológica tem a probabilidade de ser igualada ou excedida igual a 5% (p = 0.05) seu período de retorno será:

$T = l / p = l / 0.05 = 20$ anos

O período de retorno é expresso em anos. assim se uma cheia é igualada ou excedida **em média a cada 20 anos** terá um período de retorno T = 20 anos. Em outras palavras, diz-se que esta cheia tem 5% de probabilidade de ser igualada ou excedida **em qualquer ano**.

Se uma obra hidráulica for projetada para durar somente 1 ano (uma ensecadeira por exemplo) o risco de que ela seja ultrapassada por uma cheia é igual a probabilidade desta cheia.

Obras que devam durar vários anos, expõe-se **todo ano** a um risco igual à probabilidade de ocorrência de vazão de projeto.

O risco de a obra **falhar uma ou mais** vezes ao longo da sua vida útil pode ser deduzido dos conceitos fundamentais da teoria das probabilidades e é igual a:

$$R = 1 - \left(1 - \frac{1}{T}\right)^n$$

onde:

T é o período de retorno em anos

n é a vida útil da obra em anos

R é o risco permissível

Por exemplo o risco de que a canalização do rio Tamanduateí falhe uma ou mais vezes considerando que o projeto foi efetuado para T = 500 anos e sua vida útil é de 50, será:

$$R = 1 - \left(1 - \frac{1}{500}\right)^{50} = 0.1 = 10\%$$

Um problema interessante é responder à seguinte questão:

Qual a probabilidade do evento de T anos ocorrer em T anos?

Resposta:

Este problema pode ser solucionado lembrando da Distribuição Binomial. A probabilidade de não ocorrer o evento em T anos é:

probabilidade de nenhum sucesso em T anos = $\binom{T}{0} p^0 (1-p)^T$

onde p é a probabilidade de ocorrer (sucesso)

e 1-p é a probabilidade de não ocorrer (fracasso)

portanto a probabilidade de pelo menos sucesso 1 sucesso é igual a

$$1 - \binom{T}{0} p^0 (1-p)^T = 1 - (1-p)^T = 1 - \left(1 - \frac{1}{T}\right)^T \approx 1 - \frac{1}{e} \approx 0.632$$

ou seja a probabilidade de que o evento de T anos venha a ocorrer em T anos é 63.2%, ao contrário de muitos que julgam ser 100 % !

5.3 - Sumários Estatísticos

Considere-se as descargas médias anuais de um determinado rio "X". Estas informações podem ser agrupadas conforme mostradas na Tabela 5.1 ou de forma gráfica na Figura 5.1 chamado de **histograma**. Pode-se também representar descargas médias anuais em termos acumulativos, como indicado na Tabela 5.2 e na Figura 5.2 .

TABELA 5.1 - Números de vazões médias anuais em cada classe

Intervalo de vazões (m^3/s)	0 a 40	40 a 80	80 a 120	120 a 160	160 a 200	200 a 240	240 a 280
Número de ocorrências	3	5	4	3	4	1	3

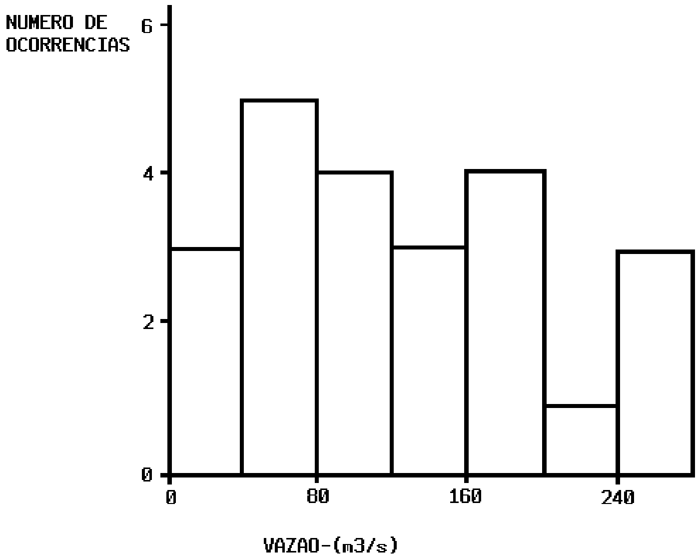

FIGURA 5.1 - Histograma das vazões médias anuais do rio "X"

TABELA 5.2 - Distribuição acumulada das vazões médias anuais do rio "X"

Vazões (m^3/s)	0	40	80	120	160	200	240	280
Número de vezes que é igualada ou excedida	23	20	15	11	8	4	3	0

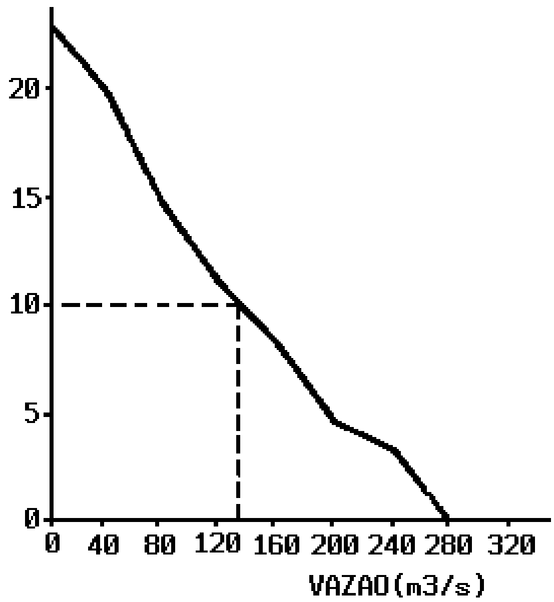

FIGURA 5.2 - Gráfico da distribuição acumulada das vazões médias anuais do rio "X"

Esses dados constituem uma pequena amostra das vazões que ocorrem normalmente. Se forem disponíveis registros muito longos, é possível diminuir os intervalos de classe e obter curvas mais suaves.

Entende-se por intervalo de classe aos intervalos em que se agrupam os dados para transformá-lo num conjunto de elementos discretos, de modo a ser possível o traçado do histograma.

Isto pode ser notado nas Figuras 5.3 e 5.4. Os registros disponíveis constituem apenas uma parte do universo de valores que definem uma curva suave. Se a forma exata dessa curva fosse conhecida, **não haveria necessidade de coletar grande quantidade de dados hidrológicos**; entretanto, essa curva não é conhecida procura-se determiná-la a partir da análise estatística das amostras observadas, como será visto a seguir.

FIGURA 5.3 - Distribuição de vazões de uma série muito longa

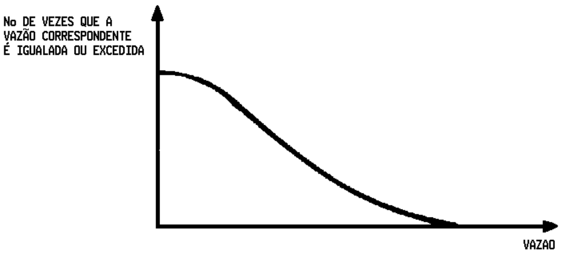

FIGURA 5.4 - Distribuição acumulada de vazões de uma série muito longa

Usualmente seleciona-se uma família de curvas que possa ajustar-se aos **dados da melhor forma possível**. Em seguida **define-se a curva particular** de melhor aderência. Por exemplo, se esta curva for um reta, basta estimar dois parâmetros, a inclinação e o intercepto, para o melhor ajuste.

As curvas usadas em estatística são normalmente caracterizadas pelos **sumários estatísticos**, dentre eles destacam-se:a média, o devio padrão e o coeficiente de assimetria. Tais parâmetros são estimados a partir das amostras pelas seguintes expressões:

Sendo x_i = o valor do evento i
n = número total de eventos,

temos:

$$Média = \overline{X} = \frac{\sum x_i}{n}$$

A média está no centróide da área sob o histograma

$$Desvio\ Padrão = S = \sqrt{\frac{\sum (x_i - \overline{X})^2}{n-1}}$$

ou

$$S^2 = \frac{\sum (x_i^2) - (\sum x_i)^2}{n-1}$$

O desvio padrão é equivalente ao raio de giração da área sob o histograma; sendo assim, é uma forma de medir o grau de dispersão em relação à média, para cada massa de dados.

$$Coeficiente\ de\ Assimetria = g = \left[\frac{\sum (x_i - \overline{X})^3}{n}\right]$$

ou

$$g = \left[\frac{n^2(\sum x_i^3) + 3n\sum(x_i)\sum(x_i^2) + 2(\sum x_i)^3}{n(n-1).(n-2)S^3}\right]^{1/3}$$

Outros parâmetros de interesse são:

Mediana = valor acima do qual situam-se de metade dos dados, consequentemente a outra metade estará abaixo dela (percentil 50%)

Moda = valor mais frequente.

Esses sumários estatísticos são utilizados na estimação dos parâmetros das **Distribuições de Probabilidades**. As distribuições são empregadas para o ajuste dos histogramas amostrais, do tipo que se descreveu acima.

Em Hidrologia as Distribuições de Probabilidades são escolhidas em função do tipo de amostra que se dispõe, isto é chuvas intensas, vazões máximas, vazões mínimas, etc.

A seguir apresenta-se algumas Distribuições Contínuas e suas aplicações. As Distribuições para Variáveis Discretas não serão tratadas neste curso.

5.4 - Distribuições de probabilidades

a) Distribuição Normal

A distribuição Normal ou Curva de Gauss é uma das mais utilizadas pelos estatísticos, principalmente pela facilidade de seu emprego. A Função Densidade de Probabilidades (FDP) teórica é dada por:

$$f_x(x) = [1/(\sigma\sqrt{2\pi})] \exp [-(x-\mu)^2/(2\sigma^2)$$

μ e σ são os parâmetros da Distribuição Normal.

É bom lembrar como se utiliza a função densidade para calcular probabilidades, ou seja:

$$P(a \leq x \leq b) = \int_a^b f_x(x).dx$$

A figura 5.5 apresenta a FDP para a Distribuição Normal:

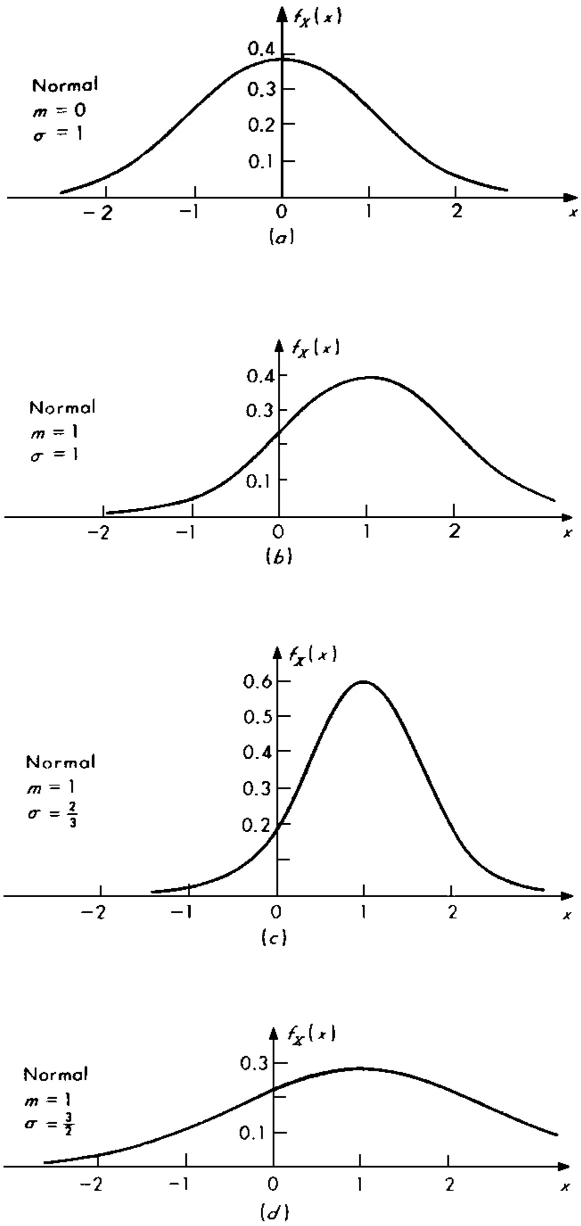

FIGURA 5.5 - Função Densidade Normal

A normal é definida por apenas dois parâmetros: o desvio padrão e a média da população. Esses parâmetros são estimados por $\overline{X}$ 10 e S, de acordo com as seguintes expressões:

$$\mu = \overline{X} \quad e \quad \sigma = S$$

Algumas de suas propriedades são:

1. $f_x(x) \to 0$ quando x $\to \forall$ 4.
2. $f_x(x)$ = máx quando x = $\overline{X}$.
3. A Distribuição é simétrica em relação à média (tem a forma de um sino).
4. O coeficiente de assimetria é igual a zero.

Em problemas de estatística é usual o emprego da chamada Função Acumulada de Probabilidade (FAP) ou seja a integral da expressão $f_x(x)$.

$$F_x(x_0) = \int_{-\infty}^{x_0} f_x(x).dx = P(x \leq x_0)$$

A distribuição normal é usada, às vezes, para análise de vazões médias anuais. Raramente proporciona bom ajuste para vazões de enchente (vazões máximas).

Uma forma muito simples de aplicar a Distribuição Normal e outras Distribuições é através da chamada expressão do **fator de frequência** :

Nesta equação a variável de interesse (vazão, chuva, etc.) é expressa em função da média, do desvio padrão e do **fator de frequência K**. Ou seja:

$$P(Q \geq Q_T) = \frac{1}{T}$$

Q_T é dado por :

$$Q_T = \overline{Q} + S_Q \cdot K_T$$

Q_T = variável de interesse (vazão, chuva, etc.) para o período de retorno T
$\overline{Q}$ = média amostral
S_Q = desvio padrão amostral
K_T = fator de frequência, tabelado conforme a Distribuição de Probabilidades em função do período de retorno T.

A série de vazões anuais do rio "X", por exemplo, permite calcular os parâmetros:

$\overline{Q}$ = 123,1 m^3/s
S_Q = 79,7 m^3/s

Se admitirmos que a distribuição das vazões médias anuais pode ser ajustada pela distribuição normal, o valor de vazão que será igualado ou superado com uma probabilidade de 10% será função de K_{10} 15. A tabela 5.3 fornece os valores de K para a distribuição normal. Então Q_{10} 16 pode ser calculado da seguinte forma:

Q_{10} = 123,1 + 1,282 (79.7)

ou

Q_{10} = 225,2 m^3/s

Para ilustrar, vamos demonstrar como se determina o fator K para a Distribuição Normal. Para as demais Distribuições serão apresentadas simplesmente as tabelas de K em função de T.

Tem-se que se Q é normal, então:

$$\xi = \frac{Q - \overline{Q}}{S_Q}$$

é Normal, com:

$$\overline{\xi} = 0 \text{ e } S_\xi = 1$$

ξ é tabelado. Portanto para obter Q_T deve-se:

$$P(Q \geq Q_T) = 1 - P(Q \leq Q_T) = 1 - P(\xi \leq \xi_T)$$

$P(\xi \leq \xi_T)$ é tabelado

donde pode-se calcular o valor de ξ_T. Portanto :

$$\xi_T = \frac{Q_T - \overline{Q}}{S_T}$$

então:

$$Q_T = \xi_T S_T + \overline{Q}$$

neste caso : $K_T = \xi_T$

TABELA 5.3 - Valores da Distribuição Normal

Probabilidade de Exceder	K	Probabilidade de Exceder	K
0.0001	3.719	0.500	0.000
0.0005	3.291	0.550	-0.126
0.001	3.090	0.600	-0.253
0.005	2.576	0.650	-0.385
0.010	2.326		
0.025	1.960	0.700	-0.524
0.050	1.645	0.750	-0.674
		0.800	-0.842
0.100	1.282	0.850	-1.036
0.150	1.036	0.900	-1.282
0.200	0.842		
0.250	0.674	0.950	-1.645
0.300	0.524	0.975	-1.960
		0.990	-2.326
0.350	0.385	0.995	-2.576
0.400	0.253	0.999	-3.090
0.450	0.126	0.9995	-3.291
0.500	0.000	0.9999	-3.719

b) Distribuição Log-Normal

O emprego da Distribuição Log-Normal é muito simples pois ela possui a seguinte propriedade:

se Q é log-normal então $\ln Q$ é normal

ou seja, ao invés de se trabalhar com a série de Q trabalha-se com a série dos logarítmos neperianos de Q. Portanto a sequência de cálculo é a mesma da Distribuição Normal.

c) Distribuição Log Pearson Tipo III

Essa Distribuição é bastante utilizada para vazões máximas anuais. Para sua utilização é necessário determinar o coeficiente de assimetria "g", já definido. A distribuição log-normal, anteriormente vista, é um caso particular da Log Pearson Tipo III quando g=0. Os valores de K estão na Tabela 5.4. e dependem de g e T.

TABELA 5.4 - Valores de (K) para Coeficiente de Assimetria (g) e Períodos de Retorno

Coeficiente de Assimetria	Período de retorno em anos							
	2	5	10	25	50	100	200	1000
	Porcentagem de probabilidade de ocorrência							
	50	20	10	4	2	1	0.5	0.1
3.0	-0.396	0.420	1.180	2.278	3.152	4.051	4.970	7.250
2.5	-0.360	0.518	1.250	2.262	3.048	3.845	4.652	6.600
2.2	-0.330	0.574	1.284	2.240	2.970	3.705	4.444	6.200
2.0	-0.307	0.609	1.302	2.219	2.912	3.605	4.298	5.910
1.8	-0.282	0.643	1.318	2.193	2.848	3.499	4.147	5.660
1.6	-0.254	0.675	1.329	2.163	2.780	3.388	3.990	5.390
1.4	-0.225	0.705	1.337	2.128	2.706	3.271	3.828	5.110
1.2	-0.195	0.732	1.340	20.87	2.626	3.149	3.661	4.820
1.0	-0.164	0.758	1.340	2.043	2.542	3.022	3.489	4.540
0.9	-0.148	0.769	1.339	2.018	2.498	2.957	3.401	4.395
0.8	-0.132	0.780	1.336	1.998	2.453	2.891	3.312	4.250
0.7	-0.116	0.790	1.333	1.967	2.407	2.824	3.223	4.150
0.6	-0.099	0.800	1.328	1.939	2.359	2.755	3.132	3.960
0.5	-0.083	0.808	1.323	1.910	2.311	2.686	3.041	3.815
0.4	-0.066	0.816	1.317	1.880	2.261	2.615	2.949	3.670
0.3	-0.050	0.824	1.309	1.849	2.211	2.544	2.856	3.525
0.2	-0.033	0.830	1.301	1.818	2.159	2.472	2.763	3.380
0.1	-0.017	0.836	1.292	1.785	2.107	2.400	2.670	3.235
0	0.000	0.842	1.282	1.751	2.054	2.326	2.576	3.090

TABELA 5.4 - Valores de (K) para Coeficiente de Assimetria (g) e Períodos de Retorno (continuação)

Coeficiente de Assimetria	Período de retorno em anos							
	2	5	10	25	50	100	200	1000
	Porcentagem de probabilidade de ocorrência							
	50	20	10	4	2	1	0.5	0.1
-0.1	0.017	0.836	1.270	1.716	2.000	2.252	2.482	2.950
-0.2	0.033	0.850	1.258	1.680	1.945	2.178	2.388	2.810
-0.3	0.050	0.853	1.245	1.643	1.890	2.104	2.294	2.675
-0.4	0.066	0.855	1.231	1.606	1.834	2.029	2.201	2.540
-0.5	0.083	0.856	1.216	1.567	1.777	1.955	2.108	2.400
-0.6	0.099	0.857	1.200	1.528	1.720	1.880	2.016	2.275
-0.7	0.116	0.857	1.183	1.488	1.663	1.806	1.926	2.150
-0.8	0.132	0.856	1.166	1.448	1.606	1.733	1.837	2.035
-0.9	0.148	0.854	1.147	1.407	1.549	1.660	1.749	1.910
-1.0	0.164	0.852	1.128	1.366	1.492	1.588	1.664	1.800
-1.2	0.195	0.844	1.086	1.282	1.379	1.449	1.501	1.625
-1.4	0.225	0.832	1.041	1.198	1.270	1.318	1.351	1.465
-1.6	0.254	0.817	0.994	1.116	1.166	1.197	1.216	1.280
-1.8	0.282	0.799	0.945	1.035	1.069	1.087	1.097	1.130
-2.0	0.307	0.777	0.895	0.959	0.980	0.990	0.995	1.000
-2.2	0.330	0.752	0.844	0.888	0.900	0.905	0.907	0.910
-2.5	0.360	0.711	0.771	0.793	0.798	0.799	0.800	0.802
-3.0	0.396	0.636	0.660	0.666	0.666	0.667	0.667	0.668

A seguir, apresenta-se um roteiro de cálculo para aplicação da Distribuição:

ROTEIRO DE CÁLCULO

1 - Transformar os N valores da variável hidrológica

$Y_1, Y_2, Y_3, ...Y_1...Y_N$ nos correspondentes logarítimos:

$X_1, X_2, X_3, ...X_1...X_N$

2 - Calcular a média dos logarítimos:

$$M_x = \frac{\sum_{i=1}^{N} X_i}{N}$$

3 - Calcular o desvio padrão dos logarítimos:

$$S_x = \sqrt{\frac{\sum_{i=1}^{N} (x_i - M_x)^2}{(N-1)}}$$

4 - Calcular o coeficiente de assimetria:

$$g = \frac{N^2}{(N-1)(N-2)} . (Mx_3 - 3 Mx_2 . Mx + 2 Mx_3)$$

$$Mx_3 = \frac{\sum_{i=1}^{N} X_i^3}{N} \quad e \quad Mx_2 = \frac{\sum_{i=1}^{N} X_i^2}{N}$$

5 - O fator K é determinado através da tabela (5.4) para o valor de g calculado e considerando-se também o período de retorno (T) desejado.

6 - Calcular os logarítimos dos valores correspondentes a determinados períodos de retorno, através da expressão log Y = Mx + K Sx (expressão do fator de frequência).

O log Y é o logarítimo do valor procurado.

7 - Achar o antilog do log Y, para determinar o valor de Y procurado

EXEMPLO DE APLICAÇÃO

O exemplo mostra a aplicação do método Log-Pearson tipo III à série histórica de 34 anos de vazões máximas observadas no Rio Jaguari, num determinado local. Deseja-se conhecer a vazão máxima correspondente ao período de retorno 100 anos.

A Tabela 5.5 apresenta os valores de vazões máximas observadas, dispostos em ordem decrescente.

TABELA 5.5 - Vazões Máximas Anuais Observadas no Rio Jaguari

m	Q (m3/s)	m	Q (m3/s)
1	490	18	167
2	425	19	165
3	314	20	163
4	302	21	153
5	289	22	139
6	250	23	135
7	244	24	123
8	240	25	121
9	237	26	116
10	225	27	113
11	212	28	109
12	212	29	102
13	206	30	96
14	205	31	95
15	182	32	93
16	171	33	76
17	169	34	52

Estes dados foram tabulados para desenvolver os cálculos anteriormente descritos. Os resultados estão na Tabela 5.6:

TABELA 5.6 - Determinação dos X_i

m	X_i=LOG Q	X_i^2	X_i^3	m	X_i=LOG Q	X_i^2	X_i^3
1	2,69	7,24	19,46	18	2,22	4,93	10,94
2	2,63	6,92	18,19	19	2,22	4,93	10,94
3	2,50	6,25	15,62	20	2,21	4,88	10,79
4	2,48	6,15	15,25	21	2,18	4,78	10,36
5	2,46	6,05	14,89	22	2,14	4,58	9,80
6	2,40	5,76	13,82	23	2,13	4,53	9,66
7	2,30	5,29	12,17	24	2,09	4,37	9,13
8	2,38	5,66	13,48	25	2,08	4,33	9,00
9	2,37	5,62	13,31	26	2,06	4,24	8,74
10	2,35	5,52	12,98	27	2,05	4,20	8,61
11	2,33	5,43	12,65	28	2,03	4,12	8,36
12	2,33	5,43	12,65	29	2,00	4,00	8,00
13	2,31	5,34	12,33	30	1,98	3,92	7,76
14	2,31	5,34	12,33	31	1,97	3,88	7,64
15	2,26	5,11	11,54	32	1,97	3,88	7,64
16	2,23	4,97	11,10	33	1,88	3,53	6,64
17	2,23	4,97	11,10	34	1,72	2,96	5,09

Seguindo o roteiro de cálculo indicado:

1 - Transformar Yi = Qi 6 Xi = log Qi (ver Tabela 5.6)

2 - Calcular a média dos logarítimos 6 Mx = 2,23

3 - Calcular o desvio padrão dos logarítimos 6 Sx = 0,2113

4 - Calcular o coeficiente de assimetria 6 g = 0,55 (da Tabela 5.6 6 Mx2 =4,97 e Mx3 = 11,25)

5 - Determinar K_T 6 da Tabela 5.4 entrando com g = 0,55 (coeficiente de assimetria) e T = 100 anos (período de retorno). K_T = 2,72

6 - Calcular log Q

$$\mathrm{Log}(Q) = Mx + K_T \bullet S_x$$

= 2,23 + 2,72 x 0,2113
log Q = 2,80

7 - Calcular Q_{100}
Q_{100} é o antilog do log Q. Logo $Q_{100} = 10^{2,80} = 631\ m^3/s$

OBS.: Para calcular Q para outros T, consultar a Tabela 5.4 entrando com g = 0,55; obtendo-se assim, valores de Q para T de interesse.

d) Distribuição de Gumbel

Outra distribuição utilizada com bons resultados para análise de máximos é a chamada distribuição de GUMBEL; expressa pela seguinte fórmula:

$$P\,(X \geq x) = 1 - e^{-e^{-y}} = \frac{1}{T}$$

ou seja:

$$y_T = -\,\mathrm{Ln}[-\,\mathrm{Ln}(\frac{T-1}{T})]$$

onde:

P = probabilidade de um valor extremo X ser maior ou igual a um dado valor x
T = período de retorno
y_T = variável reduzida Gumbel
e = base neperiana

A relação entre y_T e Q_T é dado por:

$$y_T = \frac{Q_T - \overline{Q} + 0{,}45\, S_Q}{0{,}7797\, S_Q}$$

onde:

$\overline{Q}$ = média da amostra
S_Q = desvio padrão da amostra

5.5 - Papéis de probabilidade e Análise de Frequência

Até agora foram vistos alguns métodos para ajustar analiticamente uma curva de probabilidades a um conjunto de dados. A questão agora é: como escolher a melhor Distribuição de Probabilidades? A Teoria Estatística resolve este problema com os chamados Testes de Hipóteses. Uma forma simples de se testar o ajuste de uma Distribuição é através do papel de probabilidades. É o chamado ajuste gráfico.

O Papel de Probabilidades nada mais é do que um papel em escala especial para transformar a FAP numa reta.

Na Figura 5.6 ilustra-se a construção do papel de probabilidades Normal. Na parte (a) está representada a FAP da Distribuição Normal num papel em escala linear. Na parte (b) a escala aparece alterada por anamorfose, para transformar a curva em forma de S em uma reta.

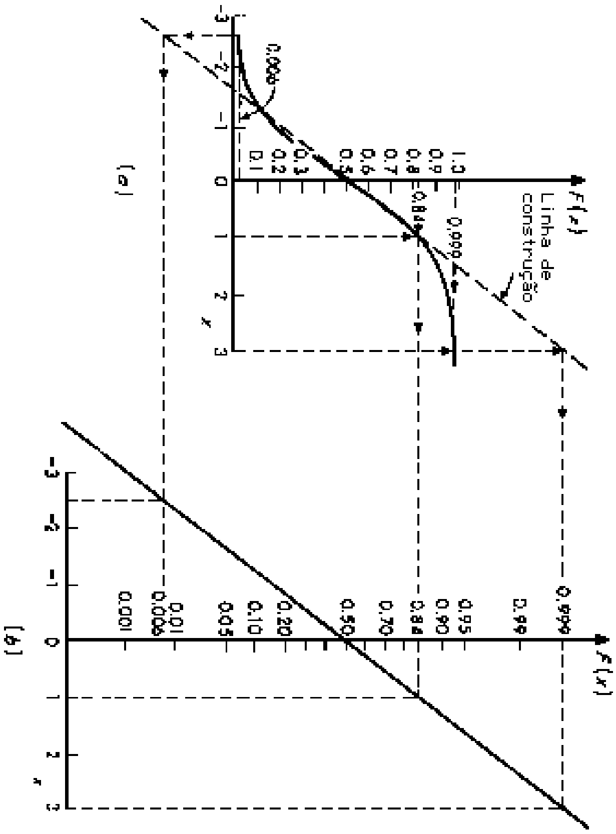

FIGURA 5.6 - Construção do papel de probabilidades Normal. a) Escala Linear. b) Papel de probabilidades.

Em anexo podem ser vistos os papéis normal, log-normal e Gumbel. Os papéis de probabilidades só são úteis para Distribuições que possuem no máximo dois parâmetros.

Se uma série de dados se ajustar a uma Distribuição Normal, as frequências desses dados deverão estar alinhadas (numa reta) no papel Normal. O mesmo ocorre para os demais papéis, Log-Normal, Gumbel, etc.

Portanto, o uso do papel de probabilidades exige que se faça a chamada **análise de frequência**:

Esta análise será descrita para o caso de vazões de enchente, o método é semelhante para qualquer outra variável hidrológica:

Suponha que se deseja calcular a vazão de cheia (máxima diária) anual para um dado período de retorno T. A primeira atividade é obter a amostra de trabalho: neste caso extrai-se dos dados de vazões médias diárias os valores máximos de cada ano hidrológico (que vai de outubro a setembro do ano seguinte no hemisfério sul), obtem-se assim a chamada **série anual**. Alguns hidrólogos preferem escolher as vazões acima de um certo valor (patamar), é a chamada **série parcial**, entretanto considera-se aqui somente a série anual.

Em seguida esses dados são tabelados, procedendo-se a uma ordenação decrescente dos mesmos. A cada valor atribui-se um número de ordem. Ao maior valor atribui-se ordem 1, ao segundo ordem 2, e assim sucessivamente até a ordem N para o menor valor. N é o tamanho da amostra.

A tabela 5.5 apresenta um exemplo deste tipo de ordenação, m é número de ordem.

A partir desta tabela pode-se estimar as frequências (probabilidades). Isto pode ser feito raciocinando da seguinte forma:

Quantos valores são maiores do que a vazão de ordem 1? A resposta é zero, pois ordenamos os valores em ordem decrescente. Agora quantos são iguais? A resposta é 1, ele próprio. Portanto pode-se estimar :

$$P(\text{vazao} \geq \text{vazao de ordem } 1) \approx \frac{1}{N}$$

Fazendo o mesmo raciocínio para a de ordem 2 chega-se a:

$$p(\text{vazao} \geq \text{vazao de ordem } 2) \approx \frac{2}{N}$$

para ordem m :

$$p(\text{vazao} \geq \text{vazao de ordem } m) \approx \frac{m}{N}$$

para se evitar o valor unitário de probabilidade para a vazão de ordem N soma-se o valor 1 ao denominador da expressão acima, obtendo-se então o estimador da frequência (ou posição de plotagem como alguns autores costumam indicar):

$$P(\text{vazao} \geq \text{vazao de ordem } m) \approx \frac{m}{N+1}$$

Portanto, para cada valor da tabela pode-se atribuir uma probabilidade de ocorrência.

Os valores de vazão e as correspondentes probabilidades (veja que o inverso das probabilidades são os períodos de retorno estimados!) são grafados nos papéis de probabilidade. Pode-se ajustar uma reta a estes pontos e extrapolá-la para estimar as vazões para os períodos de retorno desejados. A escolha da Distribuição vai depender do papel. Escolhe-se aquela Distribuição para a qual se obtem o melhor ajuste da reta.

Se ao invés de séries de valores máximos estivermos trabalhando com séries de mínimas, basta inverter a ordem de ordenação na tabela de frequência, ou seja ordena-se do menor para o maior. Os demais cálculos são idênticos.

5.6 - Curvas de Permanência

A avaliação das vazões de um rio pode ser feita através da sua curva de permanência.

A curva de permanência indica a porcentagem do tempo em que qualquer descarga foi igualada ou excedida.

A forma da curva de permanência reflete as características do regime de descargas do rio. Uma curva abatida ou achatada indica que o rio apresenta cheias reduzidas e grande potencial hídrico subterrâneo o que resulta em vazões mínimas elevadas; uma curva com formato inclinado, ao contrário, indica uma maior potencialidade de cheias e vazões mínimas mais reduzidas.

Uma das principais aplicações das curvas de permanência é o estudo do potencial hidroenergético de um rio.

Na Tabela 5.7 e mostrado um exemplo de cálculo de curva de permanência de vazões médias mensais de um curso d'água. A Figura 5.7 apresenta a curva de permanência calculada.

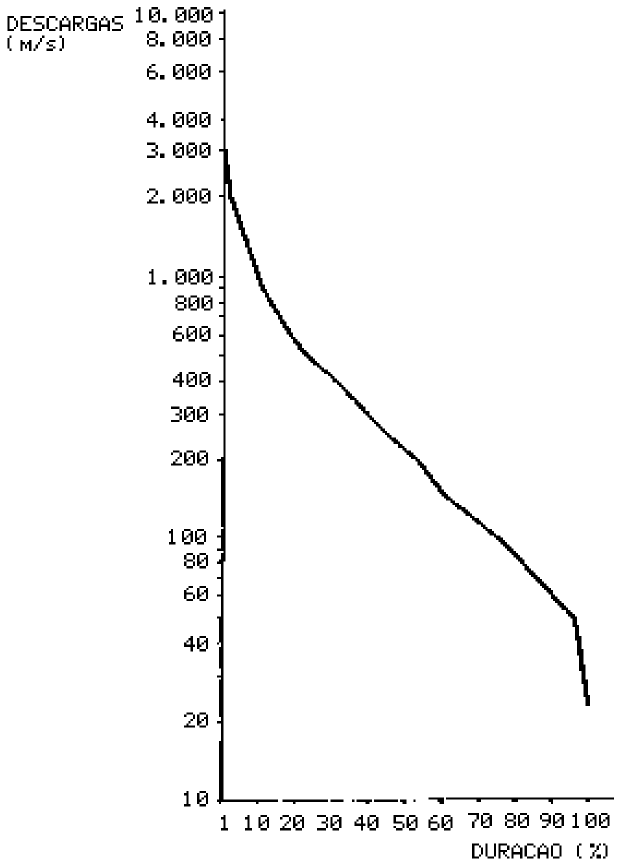

FIGURA 5.7 - Curva de permanência de vazões médias mensais

TABELA 5.7 - Dados da curva de permanência calculada

DESCARGAS (m^3/s)	OCORRÊNCIAS	OCORRÊNCIAS ACUMULADAS	DURAÇÃO (%)
23-49	10	252	100,0
50-99	54	242	96,0
100-149	38	188	74,6
150-199	16	150	59,5
200-249	20	134	53,7
250-299	14	114	45,2
300-349	10	100	39,7
350-399	9	90	35,7
400-499	23	87	32,1
500-599	11	58	23,0
600-699	8	47	18,7
700-799	6	39	15,5
800-899	5	33	13,1
900-999	4	28	11,1
1000-1999	20	24	9,5
2000-2999	4	9	1,6
	Total = 252		

5.7 - Correlação Linear Simples

A correlação linear simples é uma técnica muito utilizada em Hidrologia. Muitas análises se baseiam nesta estatística. Por exemplo, num mesmo rio em que se disponha de dois ou mais postos pluviométricos séries vazões, com abrangências diferentes, mas que tenham pelo menos um período comum de observações é possível compará-las através de correlação.

Chamando de X e Y as observações disponiveis, o Modelo de Correlação Linear é dado por:

$$Y = A + B\,X$$

onde A e B são os coeficientes de regressão linear estimados por:

$$B = \frac{\sum X_i Y_i - N\overline{X}\,.\overline{Y}}{\sum X_i^2 - N(\overline{X}\,)^2} = \frac{\sum A\,X_i\,.\,A\,Y_i}{\sum (A\,X_i\,)^2}$$

sendo:

$$A X_i = X_i - \overline{X}$$

$$A Y_i = Y_i - \overline{Y}$$

N = número de pares (X_i, Y_i)

"A" será dado por:

$$A = \overline{Y} - B\overline{X} = \overline{Y} - \overline{X} \cdot \left[\frac{\sum A X_i \cdot A Y_i}{\sum (A X_i)^2} \right]$$

A qualidade da correlação é medida pelo coeficiente de correlação "r", dado pela relação:

$$r = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum (X_i - \overline{X})^2 \cdot \sum (Y_i - \overline{Y})^2}} = \frac{\sum X_i Y_i - N \overline{X} \cdot \overline{Y}}{\sqrt{(\sum X_i^2 - N\overline{X^2})(\sum Y_i^2 - N\overline{Y^2})}}$$

onde *x e *y, são os desvios padrão de X e Y respectivamente.

O coeficiente "r" pode variar de -1,0 a 1,0; quanto mais próximo for de 1,0 ou de -1,0 melhor será a qualidade ou representatividade da correlação. O valor zero indica ue a correlação linear não é um bom modelo para a análise desejada.

Os estudos de correlação são extremamente úteis para teste de consistência de dados, assim como para o preenchimento de falhas e extensão de séries de dados em postos com séries curtas.