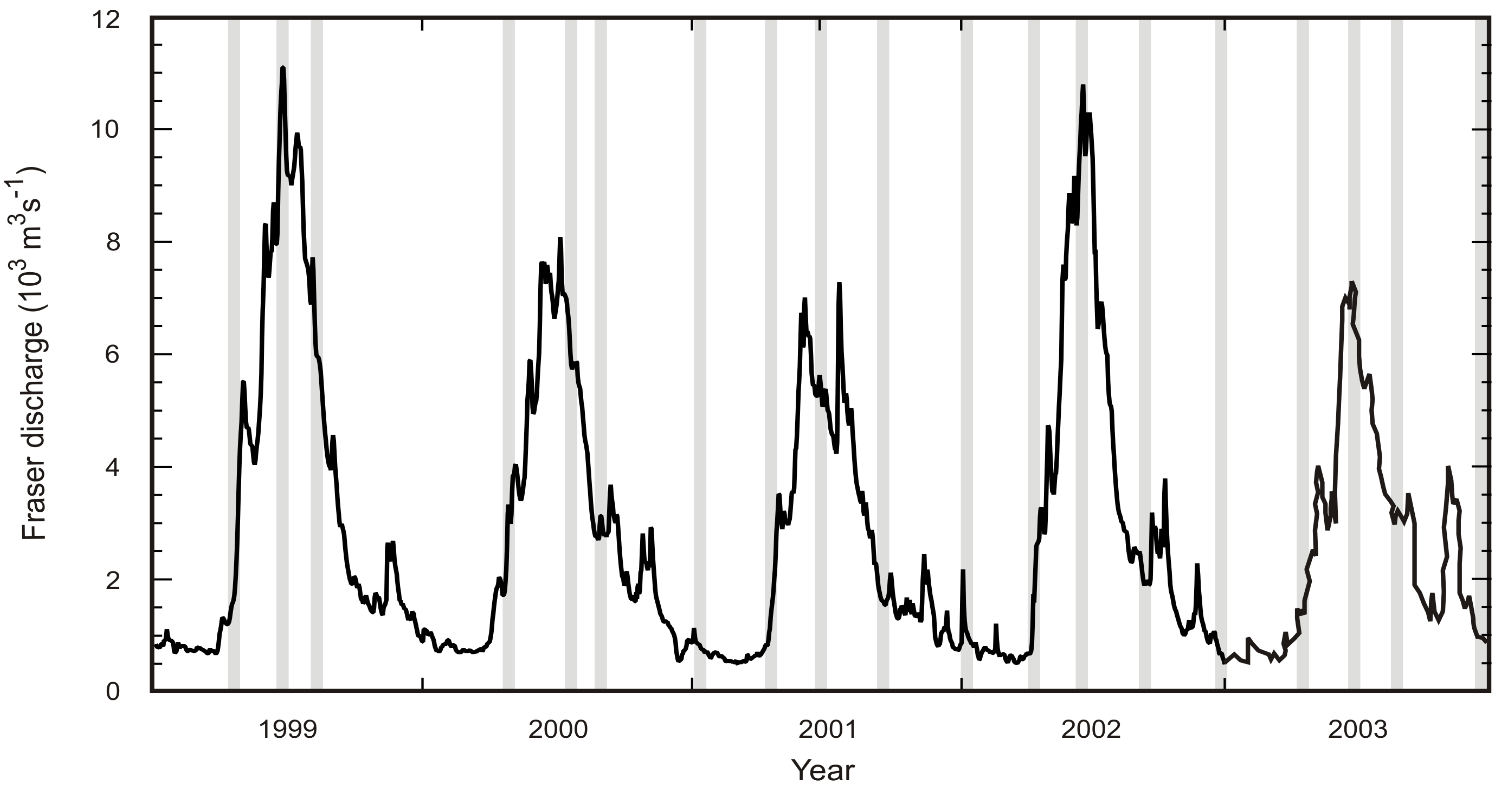
3. Frequency Analysis



Data analysis

– What do data tell us ?

– Can we generalize the results ?

• Future predictions

Prepared by: Birhanu W.

Cont…

- Water resource systems must be planned for future hydrologic events like flooding for which no exact time of occurrence can be forecasted.
- hydrologist must give a statement of the probability of the stream flows will equal or exceed a specified value.
- probabilities are important to the economic and social evaluation of a project.
- The planning goal is not to eliminate all floods but to reduce the frequency of flooding, and hence the resulting damages.
- For major projects, the failure of which seriously threatens human life, a more extreme event or the probable maximum flood, has become the standard for designing the spillway.

3.1 FREQUENCY ANALYSIS: is the hydrologic term used to describe the probability of occurrence of a particular hydrologic event

- e.g. Rainfall, Flood, Drought, etc…
- The primary objective of frequency analysis is to relate the magnitude of extreme events to their frequency of occurrence through the use of probability distributions
- The data are assumed to be independent and identically distributed.
- Hydrology mainly deal with forecasting and prediction which can only be achieved by carrying out statistical and probability analysis of historical data.
- probabilistic analysis deal with prediction of chance from the collected sample
- statistic deal with the computation of parameter with the objective to extract essential data from the set of data

- Therefore, basic knowledge about *probability* (e.g. distribution functions, plotting position) and *statistics* (e.g. measure of location, measure of spread, measure of skewness, etc) is *essential*.
- ***Statically method*** are used for collecting ,organizing and Summarizing and Analyze quantitative information.
- ***Probability*** is the scale of measurement that is used to describe likelihood of an event

N.B

Frequency analysis usually requires recorded hydrological data.

More than two statistical parameters are require to define the distribution

Distributions (probability distribution)

- **1.Disccret distribution** a) Binomial b) Poisson
- **2. Continuous distribution**
- Normal family

e.g Normal, lognormal, lognormal III

- Generalize extreme value family

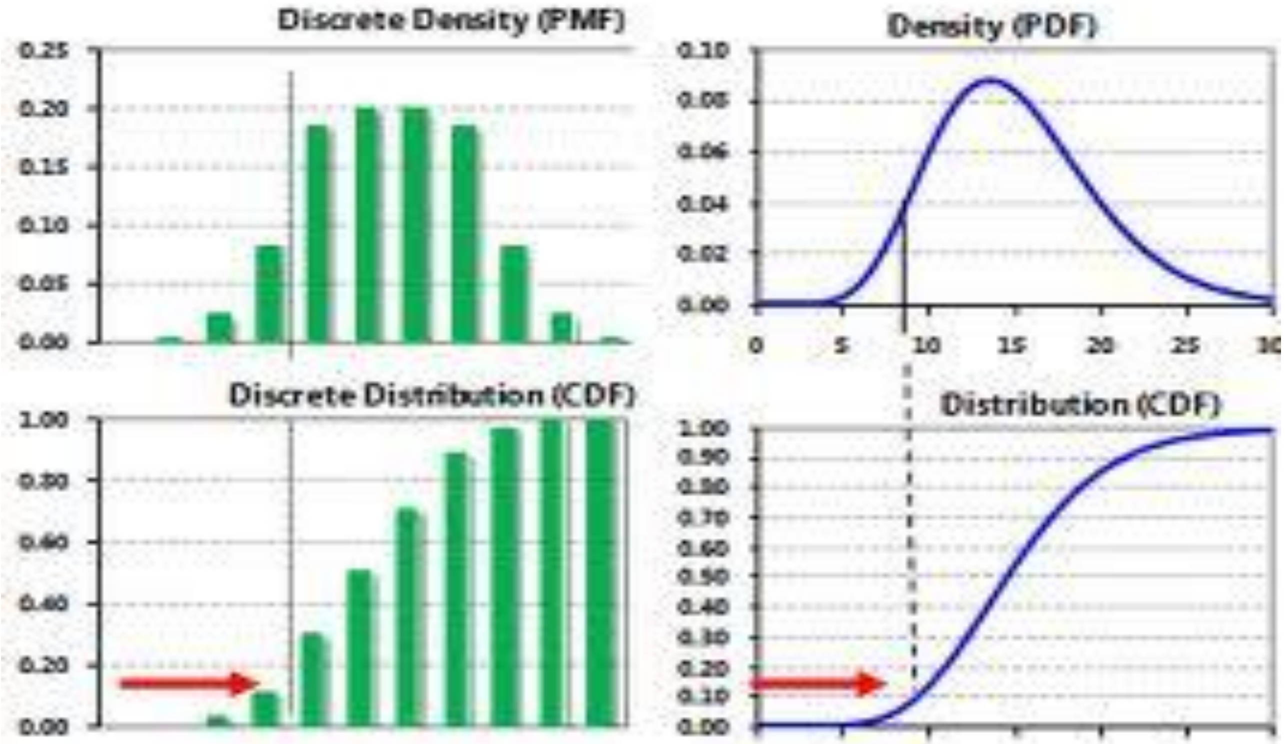
e.g EVI (Gumbel),EVII and EVIII (Weibul)

- Exponential /Pearson type family

e.g Exponential , Pearson type II ,Logpearson type III

PDF and CDF

Cumulative Distribution Function & Probable Distribution Function



$$\Pr(X \leq x_o) = \int_{-\infty}^{x_o} f(x)\, dx$$

How to use a PDF/CDF

a. ***Selection of reasonable & simple distribution.***
b. ***Estimation of parameters in distribution.***
c. ***Assessment of risk with reasonable accuracy.***

Statistical Parameters

- **Sample**: observed values of x (variate) for a finite number of years.
- **Population**: consists of the values of the variate from time immemorial to eternity.
- **Random variable**: If the value of one variate is independent of any other….
- A **statistic** is a numerical value computed from a sample. Its value may differ for different samples.

  e.g. sample mean, sample standard deviation s,
- A **parameter** is a numerical value associated with a population.

Statistical parameter

- *Describe the statistical distribution characteristics of a sample*
- *Measurement of location/ Central tendency/*

Mean, median, mode

The 1st moment about the origin

It represent the concentration of the distribution about the central value.

MEAN

$$(\bar{x} \text{ or } \mu) = \int_{-\infty}^{\infty} x f(x) dx$$

For continuous data

$$(\bar{x} \text{ or } \mu) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

For Discrete data

- *Measure of spread*

Standard deviation , variance

- The second moment about the mean it represent the variability of data

Cont-

For Discrete data

$$(S^2 \text{ or } \sigma^2) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2$$

For Continuous data

$$(S^2 \text{ or } \sigma^2) = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$$

Measure of skewness

The third moment about the mean. it represent the symmetry of the distribution.

For Continuous data

$$(g \text{ or } \gamma) = \int_{-\infty}^{\infty}(x-\mu)^3 f(x)dx$$

For Discrete data

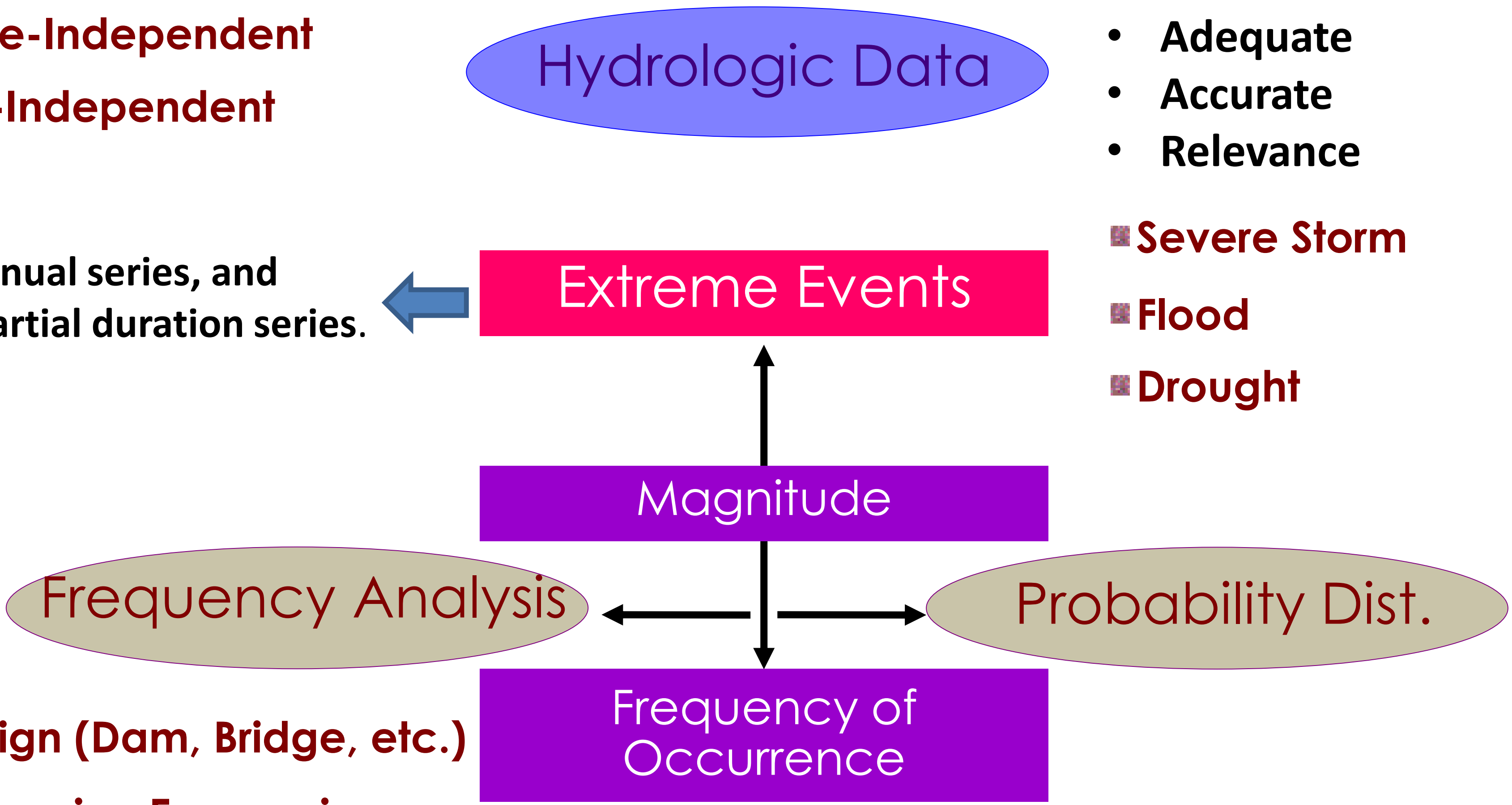
$$(g \text{ or } \gamma) = \sum_{i=1}^{n}(x_i - \mu)^3 f(x_i)$$

Kurtosis

The fourth moment of the data about the mean. It represent congestion or grouping at the central place.

Frequency Analysis & Data

- Space-Independent
- Time-Independent



- Adequate
- Accurate
- Relevance

- Severe Storm
- Flood
- Drought

(i) The annual series, and
(ii) The partial duration series.

- Design (Dam, Bridge, etc.)
- Determine Economic Value

Why Frequency Analysis?

The objective of frequency analysis of hydrologic data is to relate the magnitude of extreme events to their frequency of occurrence through the use of probability distributions.

$$\text{Magnitude} \propto \frac{1}{\text{Frequency of occurence}}$$

The results of flood flow frequency analysis can be used for many engineering **purposes**;

- for the design of dams, bridges, culvert, and flood control structures.
- to delineate flood plains.
- to determine the economic value of flood control projects.
- to determine the effect of encroachments on the flood plain.

3.2 Flow Frequency

- FDC that is the relationship between any given discharge and the percentage of time that the discharge is exceeded.
- The FDC only applies for the period for which it was derived. If this is a long period, say more than 10 to 20 years, the FDC may be regarded as a ***probability curve or flow frequency curve,*** which may be used to estimate the percentage of time that a specified discharge will be equaled or exceeded in the future
- *The shape of the FDC gives a good indication of a catchment's* *An initially steeply sloped curve results from a very variable discharge, usually from small catchments with little storage where the stream flow reflects directly the rainfall pattern.
- FDC that have very flat slope indicate little variation in flow regime or the resultant of the damping effects of large storages.

Discharge (m3/s)	Descending Order	Rank	%Exceeded or Equaled (m / (N·))
(a)	(b)	(c)	(d)
106.70	1200	1	8.33%
107.10	964.7	2	16.67%
148.20	497	3	25.00%
497.00	338.6	4	33.33%
1200.00	177.6	5	41.67%
964.70	148.2	6	50.00%
338.60	142.7	7	58.33%
177.60	141	8	66.67%
141.00	141	9	75.00%
141.00	126.6	10	83.33%
142.70	107.1	11	91.67%
126.60	106.7	12	100.00%

3.3 Flood Probability

Assumptions

- – Independence
- – Stationary with time
- – Population parameters from sample

Test on Hydrological Data

✓ Wald-Wolfowitz-for independence and Stationarity

✓ Mann-Whitney -Tests Homogeneity and Stationarity

If probability analysis is to provide reliable answers, it must start with a data series that is

- ➢ relevant, means
- ➢ adequate, means
- ➢ accurate means

- There are two data series of floods:

(i) **The annual series, and**

(ii) **The partial duration series**.(Peak Over Threshold POT)

Cont…

- **Relevance** implies that the data must deal with the problem. If the problem is to know period of time for which a highway (road) adjacent to a stream or water way is likely to be flooded, then the data series should represent duration of flows in excess of some critical value.
- **Adequacy** refers to length of stream flow record. Note: The output (flood probability) obtained from so many years data is more reliable than that obtained by using smaller number of stream flow records.
- **Accuracy** refers primarily to the problem of homogeneity or the similarity between available data and actual stream flow record.

The annual series: The annual series constitutes the data series of a single maximum daily/monthly/annually discharge record in each year so that the number of data values equals the record length in years.

The partial duration series: The partial duration series constitutes the data series with those values that exceed some arbitrary level. All the peaks above a selected level of discharge (a threshold) are included in the series and hence the series is often called the Peaks over Threshold (POT) series.

Cont…

Plotting Positions

- Flood peaks do not occur with any fixed pattern in time or magnitude. Time intervals between floods vary.
- The definition of return period Is the average of these inter-event times between flood events.
- It is a plot of peak flow and return period.
- To obtain the probability of each flood peak in the series, being equaled or exceeded, first the data is arranged in descending order, then the probability of each event is calculated.
- For small return periods or where limited extrapolation is required, a simple best-fitting curve through plotted points can be used as the probability distribution. A logarithmic scale for is often advantageous.
- However, when larger extrapolations of are involved, theoretical probability distributions (e.g. Gumbel Extreme Value, Log-Pearson Type III, and Log Normal Distributions) have to be used

Plotting Position

Plotting Position	Formula	
	T	$F = 1 - 1/T$
Weibull	$\frac{N+1}{m}$	$\frac{i}{N+1}$
Gringorton	$\frac{N+0.12}{m-0.44}$	$\frac{i-0.44}{N+0.12}$
Hazen	$\frac{N}{m-0.5}$	$\frac{i-0.5}{N}$
Blom	$\frac{N+0.25}{m-0.375}$	$\frac{i-0.375}{N+0.25}$
Cunnane	$\frac{N+0.2}{m-0.4}$	$\frac{i-0.4}{N+0.2}$
California	$\frac{N}{m}$	$\frac{i-1}{N}$
Chegodayev	$\frac{N+0.4}{m-0.3}$	$\frac{i-0.3}{N+0.4}$
Adamowski	$\frac{N+0.5}{m-0.24}$	$\frac{i-0.26}{N+0.5}$
EWSD	$\frac{N+1-\alpha}{m-\alpha}$	$\frac{i}{N+1-\alpha}$

Note: i is the rank in ascending order $= N - m + 1$; m is the rank in descending order $= N - i + 1$; N is the number of observations.

From Cunnane (1989).

Order No. m	Flood magnitude m^3/s	EX. Potting position(P)		T (1/P)
		Weibull $\frac{N+1}{m}$	California $\frac{N}{m}$	
1	160	51	50	1/51
2	135	25.5	25	1/25.5
3	128	17	17	1/17
4	116	12.75	12.5	
⋮ :	⋮ :	⋮ :	⋮ :	⋮ :
49	65	1.04	1.02	1/1.04
50	63	1.02	1	1/1.02

3.4 Probability Distribution

- Chow has shown that most frequency-distribution functions applicable in hydrologic studies can be expressed by the following equation known as the general equation of hydrologic frequency analysis

$$X_T = X + K\boldsymbol{\sigma}$$

- K =frequency factor which depends upon the return period, T and the assumed frequency distribution

General form of a probability distribution

- For estimation of large return period, extreme event probability distributions are applicable.

General form of a probability distribution:

$$x_T = \bar{x} + K\sigma$$

Where X_T = value of the variate X of a random hydrologic series with a return period T,

$\bar{X}$ = mean of the variate,

σ = standard deviation of the variate,

K = frequency factor which depends upon the return period, T and the assumed frequency distribution.

Cot…

Procedures

- Select distribution & estimate parameters
- Choose a distribution
  - X^2 test
  - Kolmogrov-Simirnov test
  - Coefficient of Skewness and Kurtosis
  - Moment ratio tests
- Use selected distribution to estimate T year event

Distribution	Parameter est.
Normal	Least squares
Gamma	MOM
Pearson type III	ML
Exponential	Maximum entropy
Pareto	PWM
Logistic	
EV I, II, III, GEV	
Wakeby	
Kappa	

Normal Distribution

- The normal distribution is used in frequency analysis for fitting empirical distributions to hydrological data, and in simulation of data.
- As many statistical parameters are approximately normally distributed, the normal distribution is often used for statistical inferences.
- The probability density function of a normally distributed variable x is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Lognormal Distribution

- If the random variable Y= log X is normally distributed, then x- is said to be log normally distributed .
- The lognormal distribution has been found to describe the hydraulic conductivity in porous media, the distribution of raindrop sizes in a storm and other hydrologic variables.

Exponential Distribution

- Some sequences of hydrologic events such as the occurrence of precipitation, may be considered Poisson process, in which events occur instantaneously and independently on a time horizon, or along a line.
- The time between such events is described by the exponential distribution whose parameter λ is the mean rate of occurrence of events.
- The probability density function of exponential distribution given by:

$$f(x) = \lambda e^{-\lambda x} \quad \text{for } x \geq 0 \text{ and } \quad \lambda = \frac{1}{\bar{x}}$$

Gamma Distribution

- The Gamma distribution has a smoothly varying form like the typical probability density function and is useful for describing skewed hydrologic variables without the need for log transformation.
- Thus, the probability density function of Gamma distribution is given by:

$$f(x) = \frac{\lambda^{\beta} x^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)} \quad \text{where } \Gamma \text{ is Gamma function and } x \geq 0 \quad \text{and} \quad \lambda = \frac{\bar{x}}{s_x^{\,2}}$$

Extreme Value Distribution

- Extreme values are selected maximum or minimum values of sets of data.
- Distributions of extreme values selected from sets of samples of any probability distribution may fit into one of the forms of Extreme value distributions (Type I, II and III).

The probability distribution function for the GEV is:

$$F(x) = exp\left[-\left(1-k\frac{x-u}{\alpha}\right)^{1/k}\right]$$ where k, u *and* α are parameters to be determined.

Three limiting cases are available where in each case α is assumed to positive.

Case- 1: when k=0 , the GEV reduces to the ***Extreme Value Type I*** (***EV I***) distribution or **Gumbel distribution** whose probability distribution function of is given by:

$$f(x) = \frac{1}{\alpha} exp\left[\frac{x-u}{\alpha} - exp\left(-\frac{x-u}{\alpha}\right)\right] \qquad -\infty < x < \infty$$ and $$\alpha = \frac{\sqrt{6}\, s_x}{\pi}$$

$$u = \bar{x} - 0.5772\alpha$$

Case-2: For k < 0 , it reduces to **GEV II (Frechet distribution)** where $u + \frac{\alpha}{k} \leq x \leq \infty$

Case-3: For k >0 , it reduces to **GEV III (Weibull distribution)** where $-\infty \leq x \leq u + \frac{\alpha}{k}$

Parameter Estimation Techniques

i) Normal distribution (MOM)

T-year flood X_T is given by:

$$X_T \equiv \overline{X} + K_T . S_x$$

where $\overline{X}$ =sample mean

S_x =sample standard deviation

K_T = frequency factor corresponding to probability of exceedance =1/T and C_s =0.0 (K_T is obtained from standard tables)

ii) Log- Normal Distribution (MOM)

$$X_T = e^{\left(\bar{y} + K_T \, . \, S_y\right)}$$

where $\bar{y}$ *and* S_y are the sample mean and standard deviations are for log transformed series respectively.

The frequency factor can be expressed as:

$$K_T = \frac{X_T - \mu}{\sigma} = Z$$

This is similar to the standard normal variable Z. The value of Z corresponding to an exceedance probability of p $\left(P = 1/T\right)$ can be calculated by finding the value of an intermediate variable , w.

$$w = \left[ln \left(\frac{1}{p^2} \right) \right]^{0.5}$$ for $0 < p \leq 0.5$, then Z is calculated using:

$$Z = w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3}$$

When $p > 0.5$, $1 - p$ is substituted for p to compute w and the computed value of Z is multiplied by (-1).

Pearson Type-II (MOM)

$$X_T = \bar{X} + K_T \, . \, S_x$$

where K_T = frequency factor corresponding to C_s of original series and probability of exceedance= $\frac{1}{T}$.

Log-Pearson Type-III (MOM)

$$X_T = e^{\left(\bar{y} + K_T \, . \, S_y\right)}$$

where K_T = frequency factor corresponding to C_s of log transformed series and probability of exceedance= $\frac{1}{T}$.

When $C_s = 0$, the frequency factor is equal to standard normal variate, Z.

When $C_s \neq 0$, K_T is approximated by the following equation (Kite, 1977).

$$K_T = Z + \left(Z^2 - 1\right)k + \frac{1}{3}\left(Z^3 - 6Z\right)k^2 - s = \left(Z^2 - 1\right)k^3 + Zk^4 + \frac{1}{3}k^5$$

$$where \quad k = \frac{C_s}{6}$$

$$C_S = \frac{n \sum_{i=1}^{n} (x_i - \bar{x})^3}{(n-1)(n-2)\, S^3}$$

Table for normal distribution

TABLE 11.2.1
Cumulative probability of the standard normal distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
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	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	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	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	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	0.9418	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

Cont…

ν	0·995	0·990	0·975	0·950	0·900	0·10	0·05	0·025	0·010	0·005
1	0·000039	0·00016	0·00098	0·0049	0·0158	2·71	3·84	5·02	6·63	7·788
2	0·0100	0·0201	0·0506	0·103	0·211	4·16	5·99	7·38	9·21	10·60
3	0·0717	0·115	0·216	0·253	0·584	6·25	7·81	9·35	11·34	12·84
4	0·207	0·297	0·484	0·711	1·06	7·78	9·49	11·14	13·28	14·86
5	0·412	0·554	0·831	1·15	1·61	9·24	11·07	12·83	15·09	16·75
6	0·676	0·872	1·24	1·64	2·20	10·64	12·59	14·45	16·81	18·55
7	0·989	1·24	1·69	2·17	2·83	12·02	14·07	16·01	8·48	20·28
8	1·34	1·65	2·18	2·73	3·49	13·36	15·51	17·53	20·09	21·96
9	1·73	2·09	2·70	3·33	4·17	14·68	16·92	19·02	21·67	23·59
10	2·16	2·56	3·25	3·94	4·87	15·99	18·31	20·48	23·21	25·19
11	2·60	3·05	3·82	4·57	5·58	17·28	19·68	21·92	24·73	26·76
12	3·07	3·57	4·40	5·23	6·30	18·55	21·03	23·34	26·22	28·30
13	3·57	4·11	5·01	5·89	7·04	19·81	22·36	24·74	27·69	29·82
14	4·07	4·66	5·63	6·57	7·79	28·06	23·68	26·12	29·14	31·32
15	4·60	5·23	6·26	7·26	8·55	22·31	25·00	27·49	30·58	32·80
16	5·14	5·81	6·91	7·96	9·31	23·54	26·30	28·85	32·00	34·27
17	5·70	6·41	7·56	8·67	10·08	24·77	27·59	30·19	33·41	35·72
18	6·26	7·01	8·23	9·39	10·86	25·99	28·87	31·53	34·81	37·16
19	6·84	7·63	8·91	10·12	11·65	27·20	30·14	32·85	36·19	38·58
20	7·43	8·26	9·59	10·85	12·44	28·41	31·41	34·17	37·57	40·00
21	8·03	8·90	10·28	11·59	13·24	29·62	32·67	35·48	38·93	41·40
22	8·64	9·54	10·98	12·34	14·04	30·81	33·92	36·78	40·29	42·80
23	9·26	10·20	11·69	13·09	14·85	32·01	35·17	38·08	41·64	44·18
24	9·89	10·86	12·40	13·85	15·66	33·20	36·42	39·36	42·98	45·56
25	10·52	11·52	13·12	14·61	16·47	34·38	37·65	40·65	44·31	46·93
26	11·16	12·20	13·8	15·38	17·29	35·56	38·88	41·92	45·64	48·29
27	11·81	12·88	14·57	16·15	18·11	36·74	40·11	43·19	46·96	49·64
28	12·46	13·56	15·31	16·93	18·94	37·92	41·34	44·46	48·28	50·99
29	13·12	14·26	16·05	17·71	19·77	39·09	42·56	45·72	49·59	52·34
30	13·79	14·95	16·79	18·49	20·60	40·26	43·77	46·98	50·89	53·67

For larger degrees of freedom ($\nu > 30$) χ_{ν}^2 is approximately normal with mean$=\nu$ and variance$=2\nu$.

Extreme –Value Type I (Gumblel's) Distribution

- It is one of the most widely used probability distribution functions for extreme values in hydrologic and meteorological studies for prediction of *flood peaks, maximum rainfalls, and maximum wind speed, etc*.
- Therefore, this extreme value theory of ***Gumbel is only applicable to annual extremes***
- In Gumbel's method, the data are ranked in ***ascending*** order and it makes use of the probability of ***non-exceedence*** q=1-P (the probability that the annual maximum flow is less than a certain magnitude). The return period T is therefore given by T = 1 / P = 1 / (1-q).

Probability Distribution Function

Extreme Value Type I (Gumbel's Distribution)-MOM

Cumulative Distribution Function(CFD)

$$F(x) = \exp[-\exp(-\frac{x-\mu}{\alpha})]$$

$$-\infty \le x \le \infty$$

Parameters :

$$\alpha = \frac{\sqrt{6}s}{\pi}$$

$$\mu = \bar{x} - 0.5772\alpha$$

For Extreme Value I Distribution:

$$K_T = -\frac{\sqrt{6}}{\pi}\left\{0.5772 + ln\left[ln\left(\frac{T}{T-1}\right)\right]\right\}$$

Reduced Variate, y :

$$y = \frac{x-\mu}{\alpha}$$

CDF :

$$F(x) = \exp[-\exp(-y)]$$

$$\frac{1}{Tr} = P(X \ge X_{Tr}) = 1 - P(X < X_{Tr}) = 1 - F(X_{Tr})$$

$$\Rightarrow F(X_{Tr}) = \frac{Tr-1}{Tr}$$

Exceedence probability = 1/Tr and F(x) becomes probability of non exceedence

EV I

The general format EV I

$$x_T = \bar{x} + K\sigma_{n-1}$$

Where σn-1 = standard deviation of the sample $\sigma_{n-1} = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N-1}}$

K = frequency factor $K = \frac{y_T - \bar{y}_n}{S_n}$

In which y_T = reduced variate, a function of T and is given by $y_T = -\left[\ln(\ln\left(\frac{T}{T-1}\right))\right]$

y_n = reduced mean, a function of sample size N and is given in Table

S_n = reduced standard deviation, a function of sample size N and is given in Table

To estimate Flood magnitude corresponding to a given return period based on annual flood series..

- Assemble the discharge data and note the sample size N. Here the annual flood value is the variate X.

Find x and σ_{n-1} for the given data.

- Using Tables 3.4 and 4.5 on your hand out
- determine Y_n and S_n appropriate to given N.
- Find Y_T for a given T .
- Find K .
- Determine the required X_T

Example(EVI)

Annual maximum recorded floods in a certain river, for the period 1951 to 1977 is given below. Verify whether the Gumbel extreme-value distribution fit the recorded values. Estimate the flood discharge with return period of (i) 100 years and (ii) 150 years by graphical extrapolation.

Year	Max. flood (m^3/s)	Year	Max. flood (m^3/s)	Year	Max. flood (m^3/s)
1951	2947	1960	4798	1969	6599
1952	3521	1961	4290	1970	3700
1953	2399	1962	4652	1971	4175
1954	4124	1963	5050	1972	2988
1955	3496	1964	6900	1973	2709
1956	2947	1965	4366	1974	3873
1957	5060	1966	3380	1975	4593
1958	4903	1967	7826	1976	6761
1959	3757	1968	3320	1977	1971

Example

$$T_P = \frac{N+1}{m} = \frac{28}{m}$$

Order number M	Flood discharge x (m^3/s)	T_P (years)	Order number M	Flood discharge x (m^3/s)	T_P (years)
1	7826	28.00	15	3873	1.87
2	6900	14.00	16	3757	1.75
3	6761	9.33	17	3700	1.65
4	6599	7.00	18	3521	1.56
5	5060	5.60	19	3496	1.47
6	5050	4.67	20	3380	1.40
7	4903	4.00	21	3320	1.33
8	4798	3.50	22	2988	1.27
9	4652	3.11	23	2947	-
10	4593	2.80	24	2947	1.17
11	4366	2.55	25	2709	1.12
12	4290	2.33	26	2399	1.08
13	4175	2.15	27	1971	1.04
14	4124	2.00			
N = 27 years, $\bar{x}$ = 4263 m^3/s, σ_{n-1} = 1432.6 m^3/s					

Cont…

Table: Reduced mean $\bar{y}_n$ in Gumbel's Extreme Value distribution, N = sample size

N	0	1	2	3	4	5	6	7	8	9
10	0.4952	0.4996	0.5035	0.5070	0.5100	0.5128	0.5157	0.5181	0.5202	0.5220
20	0.5236	0.5252	0.5268	0.5283	0.5296	0.5309	0.5320	0.5332	0.5343	0.5353
30	0.5362	0.5371	0.5380	0.5388	0.5396	0.5402	0.5410	0.5418	0.5424	0.5430
40	0.5436	0.5442	0.5448	0.5453	0.5458	0.5463	0.5468	0.5473	0.5477	0.5481
50	0.5485	0.5489	0.5493	0.5497	0.5501	0.5504	0.5508	0.5511	0.5515	0.5518
60	0.5521	0.5524	0.5527	0.5530	0.5533	0.5535	0.5538	0.5540	0.5543	0.5545
70	0.5548	0.5550	0.5552	0.5555	0.5557	0.5559	0.5561	0.5563	0.5565	0.5567
80	0.5569	0.5570	0.5572	0.5574	0.5576	0.5578	0.5580	0.5581	0.5583	0.5585
90	0.5586	0.5587	0.5589	0.5591	0.5592	0.5593	0.5595	0.5596	0.5598	0.5599
100	0.5600									

Table: Reduced standard deviation S_n in Gumbel's extreme value distribution, N = sample size

N	0	1	2	3	4	5	6	7	8	9
10	0.9496	0.9676	0.9833	0.9971	1.0095	1.0206	1.0316	1.0411	1.0493	1.0565
20	1.0628	1.0696	1.0754	1.0811	1.0864	1.0915	1.0961	1.1004	1.1047	1.1086
30	1.1124	1.1159	1.1193	1.1226	1.1255	1.1285	1.1313	1.1339	1.1363	1.1388
40	1.1413	1.1436	1.1458	1.1480	1.1499	1.1519	1.1538	1.1557	1.1574	1.1590
50	1.1607	1.1623	1.1638	1.1658	1.1667	1.1681	1.1696	1.1708	1.1721	1.1734
60	1.1747	1.1759	1.1770	1.1782	1.1793	1.1803	1.1814	1.1824	1.1834	1.1844
70	1.1854	1.1863	1.1873	1.1881	1.1890	1.1898	1.1906	1.1915	1.1923	1.1930
80	1.1938	1.1945	1.1953	1.1959	1.1967	1.1973	1.1980	1.1987	1.1994	1.2001
90	1.2007	1.2013	1.2020	1.2026	1.2032	1.2038	1.2044	1.2049	1.2055	1.2060
100	1.2065									

Example

32.6)

For N=27 Y_n S_n is 0.5332 & 1.0014 respectively

$$Y_T = -\left[\ln\left(\ln\left(\frac{T}{T-1}\right)\right)\right] = \left[\ln\left(\ln\left(\frac{10}{10-1}\right)\right)\right] = 2.25037$$

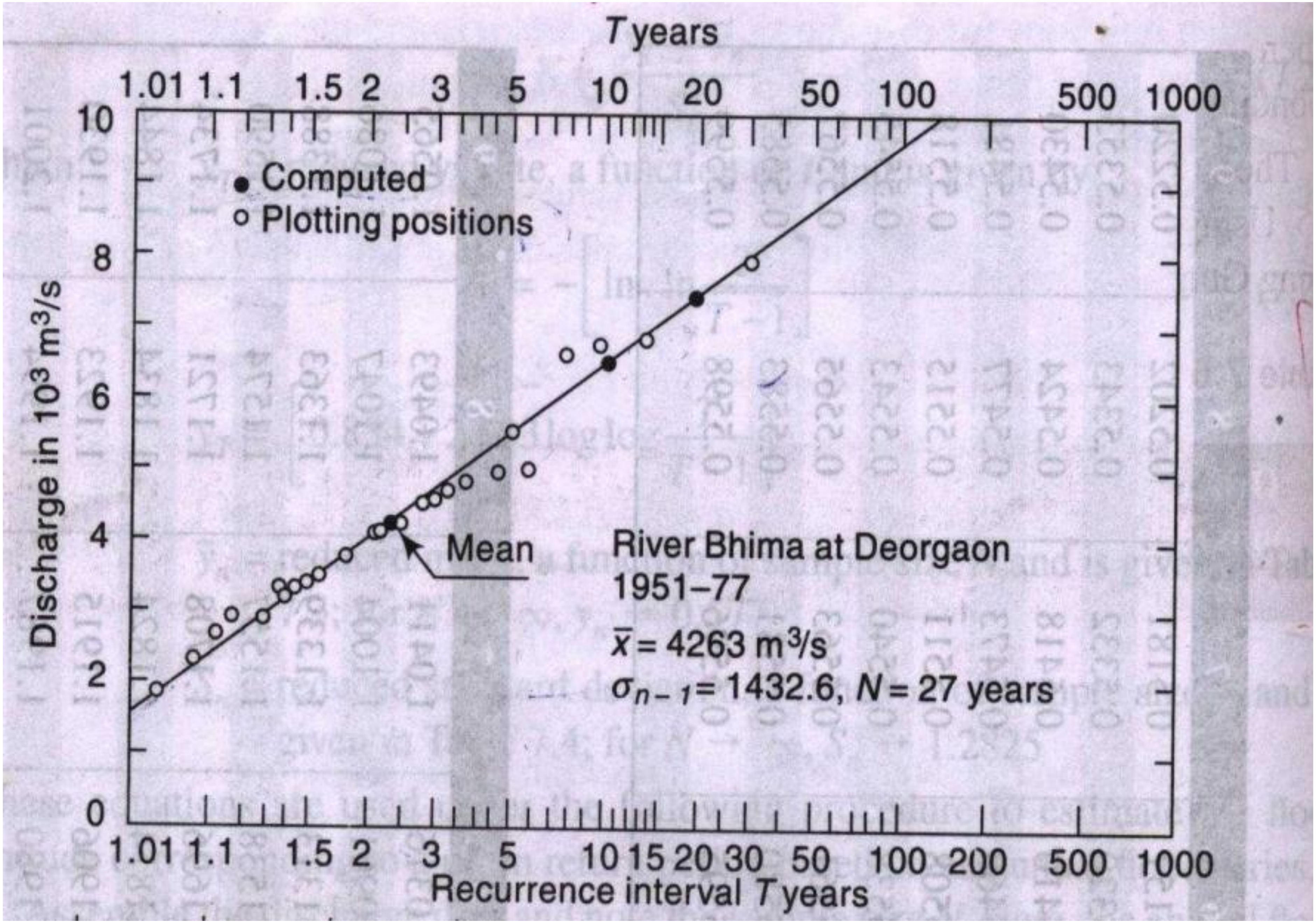
$$K = \frac{y_T - \bar{y}_n}{S_n}$$

$$K = \frac{2.25037 - 0.5332}{1.1004} = 1.56$$

$$x_T = \bar{x} + K\sigma_{n-1}$$

Return Period	X_T
5	5522
10	6498
20	7436

$$X_T = 4263 + (1.56x1432.6)$$ =6498


T years
1.01 1.1 1.5 2 3 5 10 20 50 100 500 1000
10
8
6
4
2
0
Discharge in 10³ m³/s
Computed
Plotting positions
Mean
River Bhima at Deorgaon
1951–77
x̄ = 4263 m³/s
σn−1 = 1432.6, N = 27 years
1.01 1.1 1.5 2 3 5 10 15 20 30 50 100 200 500 1000
Recurrence interval T years

Log-Pearson Type-III

This distribution is widely used in USA. In this distribution the variate is first transformed into logarithmic form (base 10) and the transformed data is then analyzed. If X is the variate of a random hydrologic series, then the series of Z variates

Where:

$Z = \log x$

$$z_T = \bar{z} + K_z \sigma_z$$

Where K_z = a frequency factor which is a function of recurrence interval T and the coefficient of skew C_s

σ_z = standard deviation of the Z variate sample = $\sqrt{\dfrac{\sum (z-\bar{z})^2}{N-1}}$

and C_s = coefficient of skew of variate Z = $\dfrac{N\sum (z-\bar{z})^3}{(N-1)(N-2)(\sigma_z)^3}$

Cs = 0, the Log-Pearson Type III distribution reduces to *Log-normal distribution*

Table 3.5 K_z=F(Cs,T) for use in Log-Pearson Type III Distribution

Coef. of skew, C_s	Return Period T in years						
	2	10	25	50	100	200	1000
3.0	-0.396	1.180	2.278	3.152	4.051	4.970	7.250
2.5	-0.360	1.250	2.262	3.048	3.845	4.652	6.600
2.2	-0.330	1.284	2.240	2.970	3.705	4.444	6.200
2.0	-0.307	1.302	2.219	2.912	3.605	4.298	5.910
1.8	-0.282	1.318	2.193	2.848	3.499	4.147	5.660
1.6	-0.254	1.329	2.163	2.780	3.388	3.990	5.390
1.4	-0.225	1.337	2.128	2.706	3.271	3.828	5.110
1.2	-0.195	1.340	2.087	2.626	3.149	3.661	4.820

1.0	-0.164	1.340	2.043	2.542	3.022	3.489	4.540
0.9	-0.148	1.339	2.018	2.498	2.957	3.401	4.395
0.8	-0.132	1.336	1.998	2.453	2.891	3.312	4.250
0.7	-0.116	1.333	1.967	2.407	2.824	3.223	4.105
0.6	-0.099	1.328	1.939	2.359	2.755	3.132	3.960
0.5	-0.083	1.323	1.910	2.311	2.686	3.041	3.815
0.4	-0.066	1.317	1.880	2.261	2.615	2.949	3.670
0.3	-0.050	1.309	1.849	2.211	2.544	2.856	3.525
0.2	-0.033	1.301	1.818	2.159	2.472	2.763	3.380
0.1	-0.017	1.292	1.785	2.107	2.400	2.670	3.235
0.0	**0.000**	**1.282**	**1.751**	**2.054**	**2.326**	**2.576**	**3.090**
-0.1	0.017	1.270	1.716	2.000	2.252	2.482	2.950
-0.2	0.033	1.258	1.680	1.945	2.178	2.388	2.810
-0.3	0.050	1.245	1.643	1.890	2.104	2.294	2.675
-0.4	0.066	1.231	1.606	1.834	2.029	2.201	2.540
-0.5	0.083	1.216	1.567	1.777	1.955	2.108	2.400
-0.6	0.099	1.200	1.528	1.720	1.880	2.016	2.275
-0.7	0.116	1.183	1.488	1.663	1.806	1.926	2.150
-0.8	0.132	1.166	1.448	1.606	1.733	1.837	2.035
-0.9	0.148	1.147	1.407	1.549	1.660	1.749	1.910
-1.0	0.164	1.128	1.366	1.492	1.588	1.664	1.880
-1.4	0.225	1.041	1.198	1.270	1.318	1.351	1.465
-1.8	0.282	0.945	1.035	1.069	1.087	1.097	1.130
-2.2	0.330	0.844	0.888	0.900	0.905	0.907	0.910
-3.0	0.396	0.660	0.666	0.666	0.667	0.667	0.668

year	X	z=log(x)	Year	X	z=log(x)
1951	2947	3.46938	1965	4366	3.640084
1952	3521	3.546666	1966	3380	3.528917
1953	2399	3.38003	1967	7826	3.89354
1954	4124	3.615319	1968	3320	3.521138
1955	3496	3.543571	1969	6599	3.819478
1956	2947	3.46938	1970	3700	3.568202
1957	5060	3.704151	1971	4175	3.620656
1958	4903	3.690462	1972	2988	3.475381
1959	3751	3.574147	1973	2709	3.432809
1960	4798	3.68106	1974	3873	3.588047
1961	4290	3.632457	1975	4593	3.662096
1962	6552	3.816374	1976	6761	3.830011
1963	5050	3.703291	1977	1971	3.294687
1964	6900	3.838849			

$$\sigma_z = 0.1427 \qquad C_s = \frac{27 \times 0.0030}{(26)\,(0.1427)^3\,(25)}$$

$$\bar{z} = 3.6071 \qquad = 0.043$$

T	Kz	Kz$\boldsymbol{\sigma}$	Zt	XT (antilog)
100	2.358	0.3365	3.9436	8782
200	2.616	0.3733	3.9804	9559

The value of K from table for each return period and skewness coefficient of 0.043

Confidence Limits for the Fitted Data

- Since the value of the variate for a given return period, x_T determined by a given distribution can have errors due to the limited sample data used; an estimate of the confidence limits of the estimate is desirable.
- The confidence interval indicates the limits about the calculated value between which the true value can be said to lie with a specific probability based on sampling errors only.

For a confidence probability c, the confidence interval of the variate x_T is bound by value x_1 and x_2 given by

$$x_{1/2} = x_T \pm f(c)\, S_e$$

Where f(c) = function of the confidence probability c determined by using the table of normal variate

C in per cent	50	68	80	90	95	99
f(c)	0.674	1.00	1.282	1.645	1.96	2.58

$$S_e = probable\ error = b\frac{\sigma_{n-1}}{\sqrt{N}} \quad b = \sqrt{1 + 1.3K + 1.1k^2}$$

K= frequency factor

$$K = \frac{y_T - \overline{y}_n}{S_n}$$

Example

Data covering a period of 92 years for a certain river yielded the mean and standard deviation of the annual flood series as 6437 and 2951 m^3/s respectively. Using Gumbel's method, estimate the flood discharge with a return period of 500 years. What are the (a) 95% and (b) 80% confidence limits for this estimate?

Solution: From Table 4.4 and 4.5 for N = 92 years, $\bar{y}_n$ = 0.5589, and S_n = 1.2020. Then

$y_{500} = -[\ln((\ln(500/499))] = 6.21361$

$K_{500} = (6.21361 - 0.5589)/1.2020 = 4.7044$,

Hence, $x_{500} = 6437 + 4.7044*2951 = 20320 m^3/s$.

From Eq.(6.16a), $b = \sqrt{1 + 1.3(4.7044) + 1.1(4.7044)^2} = 5.61$

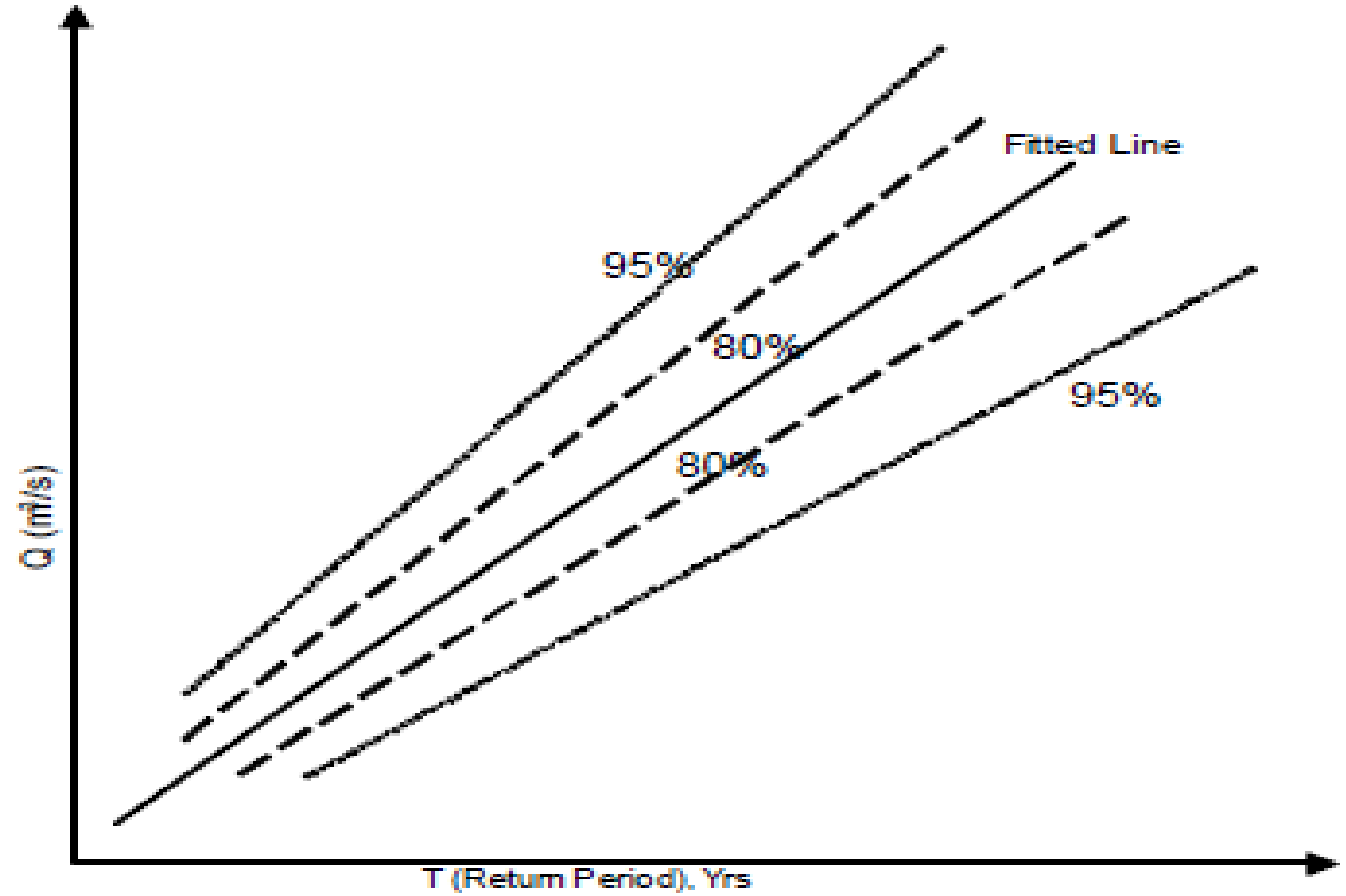
$$S_e = \text{probable error} = 5.61 * \frac{2951}{\sqrt{92}} = \underline{1726}$$

(a) For the 95% confidence probability *f(c)* = 1.96 and by Eq.(4.16) $x_{1/2}$ = $20320 x_T \pm (1.96*1726)$, which results in $x_1 = 23703 m^3/s$ and $x_2 = 16937 m^3/s$. Thus the estimated discharge of $20320 m^3/s$ has a 95% probability of lying between 23700 and $16940 m^3/s$.

(b) For 80% confidence probability, f(c) = 1.282 and by Eq.(4.16) $x_{1/2}$ = $20320 x_T \pm (1.282*1726)$, which results in $x_1 = 22533 m^3/s$ and $x_2 = 18107 m^3/s$. Thus the estimated discharge of $20320\ m^3/s$ has an 80% probability of lying between 22533 and $18107 m^3/s$.

For the data of Example 4.2, the values of x_T for different values of T are calculated and can be shown plotted on a Gumbel probability paper.

Confidence Interval



Testing the Goodness of Fit

The validity of probability distribution function proposed to fit the empirical frequency distribution of a given sample may be tested using graphical and analytical methods.

Analytical tests:

- **i)** Chi-square test
- **ii)** Kolmogorov-Smirnov test
- **iii)** D-index test

Chi-square (χ^2) Test

The sample value of the relative frequency of interval, i is :

$$f_s(x_i) = \frac{n_i}{n}$$ .... relative frequency function

The theoretical value is :

$$p(x_i) = F(x_i) - F(x_{i-1})$$ ... incremental probability function

The $\chi_c{}^2$ test statistic given by :

$$\chi_c{}^2 = \sum_{i=1}^{m} \frac{n\,[f_s(x_i) - p(x_i)]^2}{p(x_i)}$$

where m is the number of intervals. It may be noted that $n\, f_s(x_i) = n_i$, the observed number of occurrences in interval i ; $np(x_i)$ is the corresponding expected number of occurrences in interval i .

To describe the χ^2 test, the χ^2 probability distribution must be defined. A χ^2 distribution with ν degree of freedom is the distribution for the sum of squares of ν independent standard normal random variables Z_i; the sum is the random variable $\chi_\nu^2 = \sum_{i=1}^{\nu} Z_i^2$. The χ^2 distribution is tabulated and given as standard tables. In the χ^2 test $\nu = m - p - 1$ where m is the number of intervals and p is the number of parameters used in fitting the proposed distribution. The confidence level is chosen for the test and often expressed as $1-\alpha$ where α is termed as the significance level ($\chi^2{}_{critical}\ at\ \alpha\,\%\ S.L = \chi^2{}_{(1-\alpha),(m-p-1)}$). The *null hypothesis* for the test is that the *proposed probability distribution fits the data adequately*.

If $\chi^2{}_{comp} < \chi^2{}_{critical}$, then the distribution can be assumed to fit well (the null hypothesis is accepted) otherwise rejected. For $\chi^2{}_{critical}$ standard tables can be consulted.

Example

Year	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920
P ,in	39.9	31	42.3	42.1	41.1	28.7	16.8	34.1	56.4	48.7
Year	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930
P ,in	44.1	42.8	48.4	34.2	32.4	46.4	38.9	37.3	50.6	44.8
Year	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940
P ,in	34	45.6	37.3	43.7	41.8	41.1	31.2	35.2	35.1	49.3
Year	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950
P ,in	44.2	41.7	30.8	53.6	34.5	50.3	43.8	21.6	47.1	31.2
Year	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960
P ,in	27	37	46.8	26.9	25.4	23	56.5	43.4	41.3	46
Year	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970
P ,in	44.3	37.8	29.6	35.1	49.7	36.6	32.5	61.7	47.4	33.9
Year	1971	1972	1973	1974	1975	1976	1977	1978	1979	
P ,in	31.7	31.5	59.6	50.5	38.6	43.4	28.7	32	51	

Solution:

Step 1: Divide the given data into suitable ranges (Column 1)

Step 2: Determine the frequency of occurrence (the number of data in the ranges)

Step 3: compute the relative frequency function as $f_s(x_i) = \frac{n_i}{n}$. *For example* $f_s(x_4) = \frac{14}{69} = 0.203$ (column 3)

Step 4: Compute the cumulative frequency function , $p(X \le x)$(summation of consecutive relative frequencies) (column 4)

Step 5: Compute the standard normal variate , Z as $Z = \frac{x-\mu}{\sigma}$, use the lower range as x-value (column 5)

Step 6: find the corresponding value the cumulative probability function using eqn. or standard tables (column 6)

Step 7: Find the incremental Probability function:

eg. $P(x_4) = p(30 \le X \le 35) = F(35) - F(30) = 0.301 - 0.144 = 0.158$.

Step 8: Compute the χ^2 test statistic using $\chi_c^{\,2} = \sum_{i=1}^{m} \frac{n\left[f_s(x_i) - p(x_i)\right]^2}{p(x_i)}$. For example, for i=4

$$\chi_c^{\,2} = \frac{69*(0.2029-0.1577)^2}{0.1577} = 0.891$$

Table: Chi-square computation

Column	1	2	3	4	5	6	7	8
Interval, i	Range	n_i	$f_s(x_i)$	$F_s(x_i)$	Z_i	$F(x_i)$	$P(x_i)$	x_c^2
1	<20	1	0.014	0.014	-2.157	0.015	0.015	0.004
2	20-25	2	0.029	0.043	-1.611	0.053	0.038	0.147
3	25-30	6	0.087	0.13	-1.065	0.144	0.091	0.008
4	**30-35**	**14**	**0.203**	**0.333**	**-0.52**	**0.301**	**0.157**	**0.891**
5	35-40	11	0.159	0.493	0.026	0.51	0.209	0.805
6	40-45	16	0.232	0.725	0.571	0.716	0.206	0.222
7	45-50	10	0.145	0.87	1.117	0.868	0.152	0.019
8	50-55	5	0.072	0.942	1.662	0.952	0.084	0.114
9	55-60	3	0.043	0.986	2.208	0.986	0.034	0.163
10	>60	1	0.014	1.000	2.753	1.00	0.014	0.004
	Total	**69**	**1.000**				**1.000**	**2.377**
mean	39.77							
stdev.	9.17							

At 5 % significance level, the $\chi^2_{7,0.95} = 14.1$ (Note: $\nu = m - p - 1 = 10 - 2 - 1 = 7$).

Since $\chi^2_{comp} < \chi^2_{critical}$ i.e., 2.377<14.1 , the ***null hypothesis is accepted***, the ***data fits well the normal distribution***.

3.5 Low Flow Analysis

- The objective of low-flow analysis is to estimate the frequency or probability with which stream flow in a given reach will be less than various levels.

Thus the flow-duration curve(is an important tool of low-flow analysis)

From FDC one can readily determine the flow associate with any exceedence or non-exceedence probability.

Low Steam Flows

Most of the time, the flow exceeded 95% of the time, q95, is a useful index of water availability that is often used for design purposes.

Cont…

Why low flow analysis is required ?

- Conditions of drought upset the ecological balance.
- adequacy of a stream to receive wastes,
- To supply municipal and industrial water requirements,
- To meet supplemental irrigation demand,
- To maintain aquatic life.

Lowflow analysis in gauged sites

- **Frequency Analysis:** This analysis is similar to peak-flow analysis.
- **Analysis of runs:** The deficit in flow with respect to the base flow is counted along with the duration (of the drought).

Cont…

Lowflow analysis for ungauged catchments

- **Partial-record length:** A few **baseflow** measurements at a site are related to the **concurrent discharges at a neighbouring station** for which a low flow frequency curve is available.
- **Seepage Runs:** During a period of baseflow , **discharge** at intervals along a channel reach are measured to identify the **loss or gain in the flow** along a river channel.
- **Interpolation:** By plotting the **low flow characteristics at gage sites against the channel distance** , the flows at intermediate points are interpolated.

Cont…

- Low-flow quantize values are cited as "dQp," where p is now the annual non-excedence probability (in percent) for the flow averaged over d-days

a. Low flow analysis at gauged stations:

For gauged reach, low flow analysis involves development of a time series of annual d-day low flows, where d is the averaging period.

The following steps are usually applied:

i. List days in first column of your table and enter average daily flows for each year in the second column

ii. Compute overlapping d day averages for the averaging periods of interest. For each value of d, this creates[$365-(-1)$]) values of consecutive d-day averages for each year.

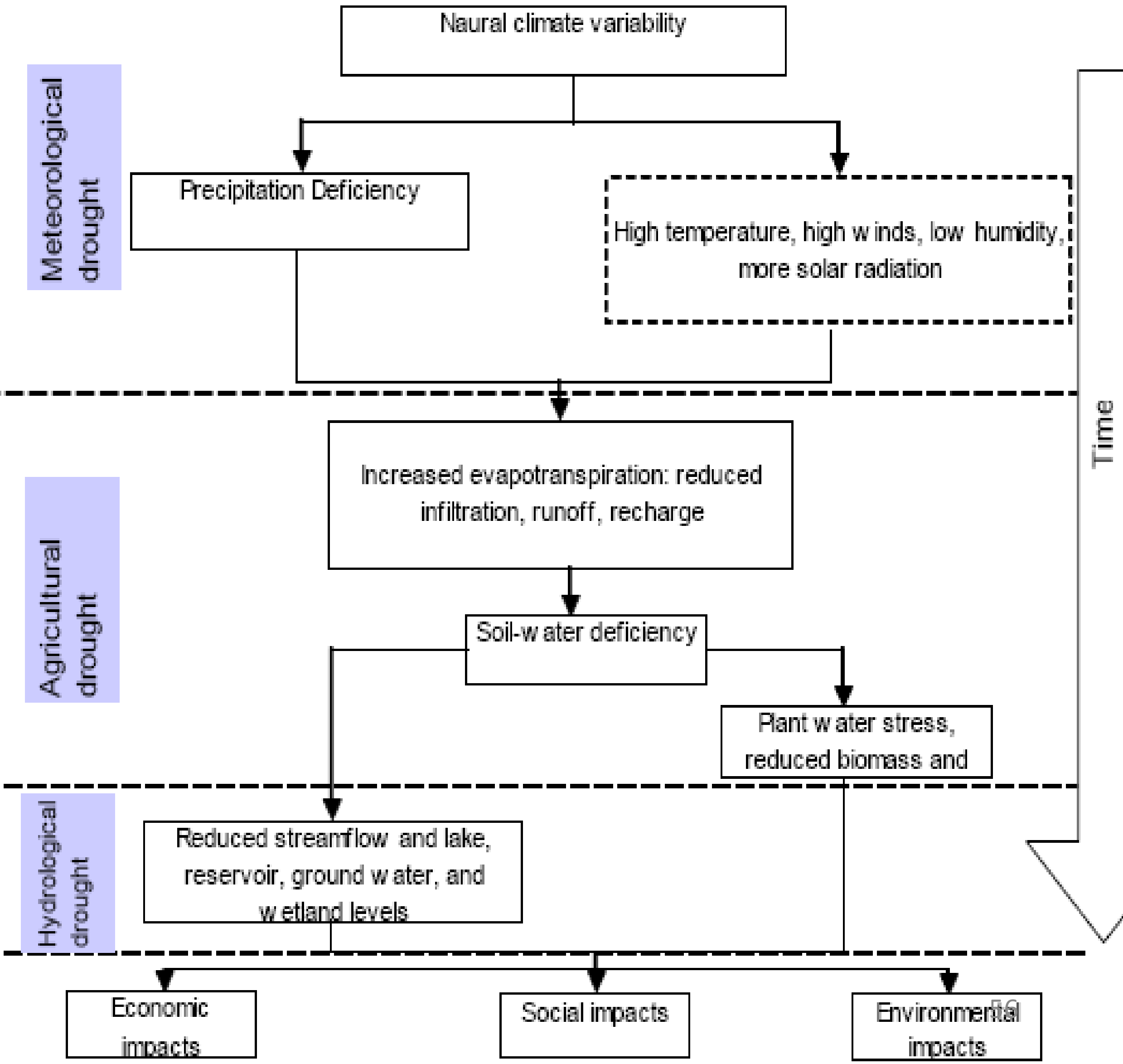
iii. Smallest of these values (among[$365-(-1)$]) is then selected to produce an annual time series of minimum d-day flows and it is this time series that is then subjected to frequency analysis to estimate the quantiles of the annual d-day flows.

Drought Types

Type of drought

- **Meteorological**
- **Agricultural**
- **Hydrological**

The objective of drought analysis is to **characterize** the magnitude, duration, and severity of meteorological, agricultural, or hydrological drought in a region of interest.



Drought Analysis

- ***Drought definition:-*** A time series of a selected quantity, X (e.g., **precipitation**, **stream flow**, **ground water level**), averaged over an appropriate dt. The quantitative definition of drought is determined by the truncation level, X_0, selected by the analyst: Values of $X < X_0$ are defined as droughts.

$$X_0 = \bar{X} \,(\text{mean value of } X)$$

$$X_0 = X_{50} \,(\text{median of } X) \text{ or}$$

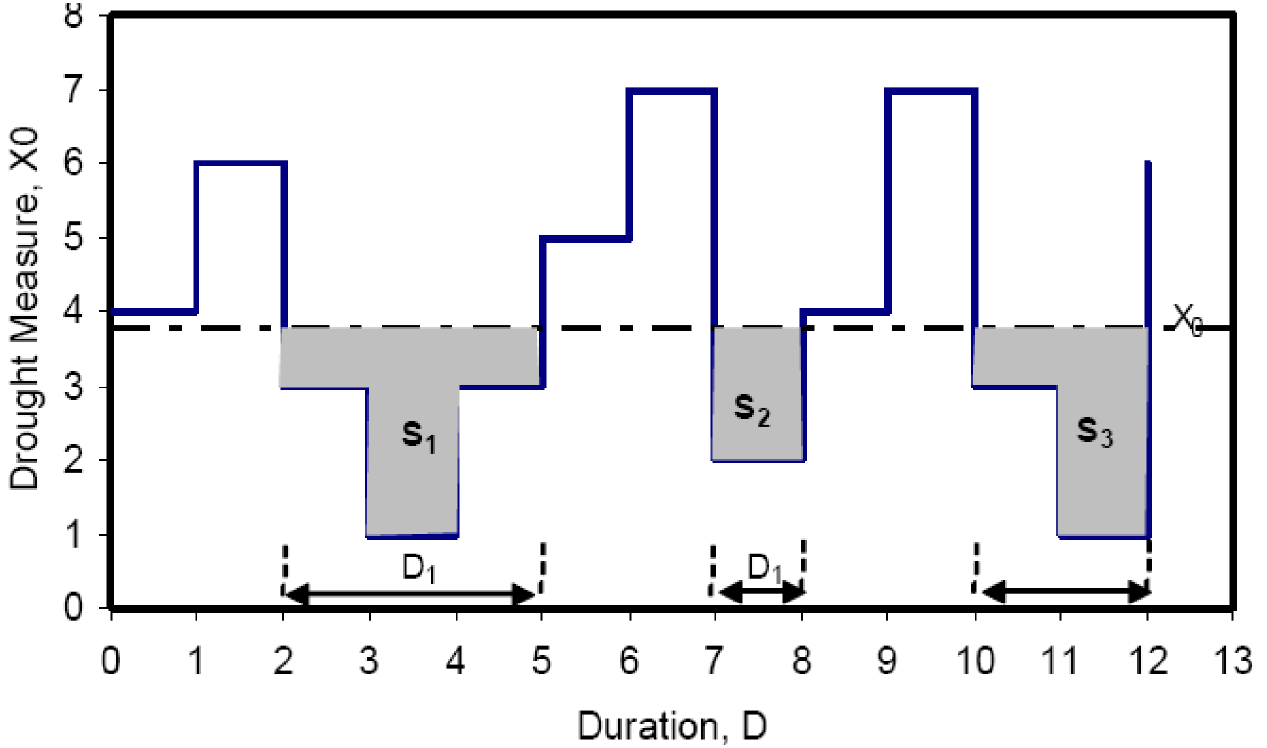
$$X_0 = \bar{X} - \sigma_{n-1} \,(\text{mean minus one standard deviation})$$

What type of drought of interest?

Averaging Period:-drought analysis requires selection of an averaging period (dt).

Drought Characteristics

- Once the severities, durations, and intensities of "drought" have been determined for a given time series, the magnitude-frequency characteristics of each of those quantities can be analyzed.



Low Flow Analysis

Estimation of 7-day Average 10-yr Low Flow

Many times, flow diversions from streams are permissible only if flows equal to or more than the 7-day, 10-yr (7Q10) low flows of the stream are left in the stream to meet instream flow requirements. The 7Q10 low flow is also used to evaluate the impacts of wastewater discharges on the water quality of the receiving stream. The 7-day average implies an average of all successive combinations of 7 days of flows in a year, i.e., average of flows from day 1 to day 7, day 2 to day 8, day 3 to day 9, etc. This is also known as the 7-day moving average. The lowest 7-day average flow is computed for each year for which records are available. The mean,

standard deviation, and skew coefficient of the resulting lowest annual 7-day average low flows are computed, as described in the section of this chapter entitled "Statistical Analysis of Available Data." Using normal, log-normal, or log Pearson Type III distributions, the *T*-year low flow is computed using the equation,

$$Q_T = \overline{X} - K s \qquad (2\text{-}21)$$

Find the 7 day 10yr low flow?
(Log Pearson Type III)

Table 2-22(b). Computation of 7-day average low flow

Daily flow (m^3/s)	7-day moving average (m^3/s)
14.39	
33.74	
10.80	
8.93	
8.15	
11.06	
6.58	13.379
10.98	12.891
9.35	9.407
—	—
—	—
—	—
5.14	12.864
4.41	11.009
3.53	8.774
3.93	6.720
2.66	5.500
2.82	4.376
4.01	3.786
6.00	3.909
9.41	4.623
—	—
—	—
—	—

Table 2-22(a). 7-day average annual low flows

Year	Flow (m^3/s)
1	4.111
2	3.212
3	5.313
4	6.005
5	2.891
6	2.976
7	3.505
8	3.987
9	3.831
10	5.892
11	4.905
12	4.303
13	4.550
14	3.525
15	3.761
16	*

Compute of d-consecutive day averages for low flow analysis given below. Where'd' repres
7and 15.

Given Data		Solution		
Days	1-Day Average (m^3/s)	3-Day Average (m^3/s)	7-Day Average (m^3/s)	15-Day Average (m^3/s)
1	5.75			
2	5.55	*5.59		
3	5.47	**5.45		
4	5.32	5.34	***5.35	
5	5.24	5.24	5.24	
6	5.15	5.13	5.29	
7	4.98	5.04	5.52	
8	4.98	5.29	5.70	5.60
9	5.89	5.97	5.78	5.58
10	7.05	6.52	5.82	5.63
11	6.63	6.48	5.87	5.73
12	5.78	5.95	5.93	5.77
13	5.44	5.51	5.87	5.78
14	5.32	5.39	5.76	5.77
15	5.41	5.41	5.81	5.76
16	5.49	5.71	5.85	5.75
17	6.23	6.25	5.83	5.67
18	7.02	6.42	5.78	5.54
19	6.00	6.12	5.70	**5.50**
20	5.32	5.43	5.61	5.52
21	4.96	5.04	5.39	5.53
22	4.84	4.89	**5.12**	6.08
23	4.87	**4.81**	5.13	6.79
24	**4.73**	4.91	5.22	
25	5.13	5.30	5.31	
26	6.03	5.71	6.55	
27	5.97	5.86	8.15	
28	5.58	8.37		
29	13.56	11.73		
30	16.06			

Note well: Values in **bold** are minimum for the period shown

Sample calculation:

$$* \, 5.59 = \frac{5.75 + 5.55 + 5.47}{3} \quad \textit{and} \; ** \, 5.45 = \frac{5.55 + 5.47 + 5.32}{3}$$

Sample calculation:

$$* \, 5.59 = \frac{5.75 + 5.55 + 5.47}{3} \quad \textit{and} \; ** \, 5.45 = \frac{5.55 + 5.47 + 5.32}{3}$$

Return Period

- The Rxn b/n probability of occurrence of flood and its return period can be justified as : a given flood of with a return period T may be exceeded once in T year Hence the probability of execeedence F(X)
- Suppose that an extreme event is defined to have occurred if a random variable X is greater than or equal to some level x_{Tr}. $X \geq x_{Tr}$

- The recurrence interval, T is the time between occurrences of $X \geq x_{Tr}$
- The probability of occurrence of the event $X \geq X_T$ in any observation is **p**

Hence, $$p = P(X \geq x_{Tr}) = \frac{1}{Tr}$$ ⇨ **T is inverse of P**

3.6 Risk Analysis

- Natural inherent risk of failure the probability of occurrence of an event ($x \geq xT$) at least once over a period of n successive years is called the risk, $\bar{R}$. Thus the risk is given by $\bar{R}$ = 1 -(probability of non-occurrence of the event $x \geq xT$ in n years)

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

$$\bar{R} = 1 - \left[1 - \frac{1}{T}\right]^n$$

The reliability R_e, is defined as

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

n ⇒could be the design period of the structure or economical life

R↓ T↑ Q↑

- Consideration of Risk
  - Structure may fail if event exceeds T–year design magnitude
  - R = P(event occurs at least once in *n* years)

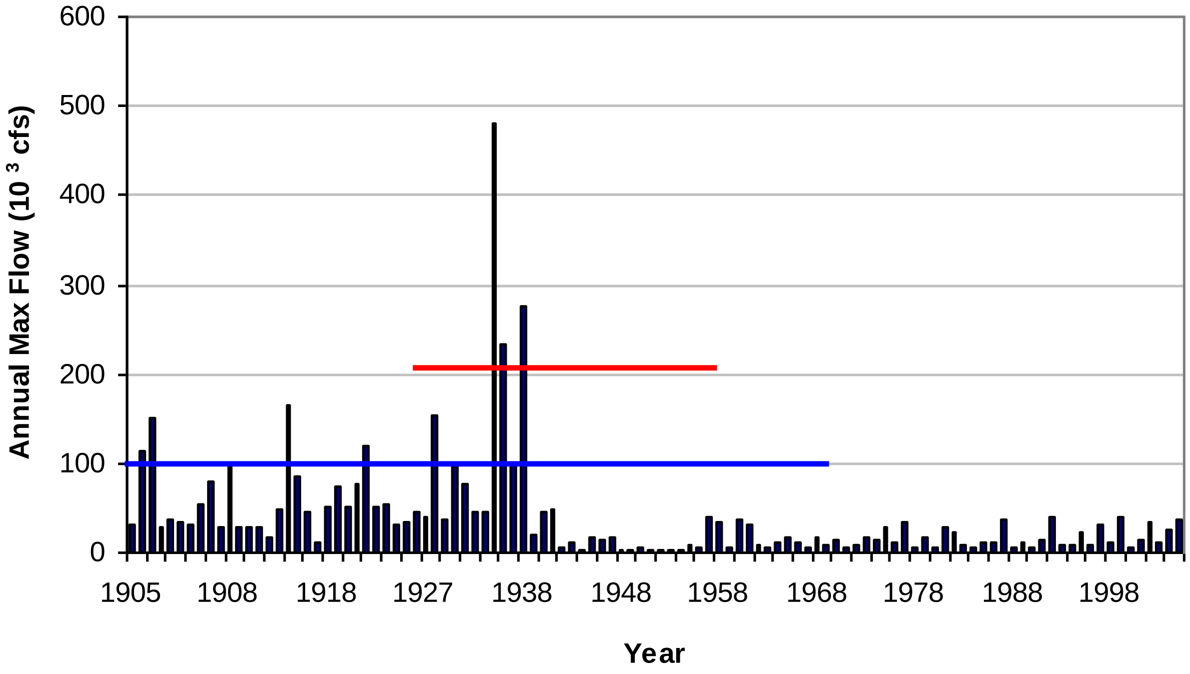
- **Safety Factor:** addresses uncertainties associated with structural, constructional, operational and environmental causes as well as from non-technological considerations such as economic, sociological and political causes.
- Thus , any water resource development project will have a safety factor for a given hydrological parameter M as defined below.

Safety factor (Sm)

$$= \frac{\text{Actual value of the parameter M adopted in the design of the project}}{\text{Value of the parameter M obtained from hydrological considerations only}}$$

Example Risk

- Dataset – annual maximum discharge for 106 years on a River



x_T = 200,000 cfs

No. of occurrences = 3

2 recurrence intervals in 106 years

T = 106/2 = 53 years

If x_T = 100, 000 cfs

7 recurrence intervals

T = 106/7 = 15.2 yrs

P(X ≥ 100,000 cfs at least once in the next 5 years) = $1- (1-1/15.2)^5 = 0.29$

Example

- Expected life of culvert = 10 yrs
- Acceptable risk of 10 % for the culvert capacity
- Find the design return period

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

$$0.10 = 1 - \left(1 - \frac{1}{T}\right)^{10}$$

$$T = 95 \ \ yrs$$

- What is the chance that the culvert designed for an event of 95 yr return period **will not** have its capacity exceeded for 50 yrs?

The risk associated with failure of culvert when the flow exceed 95 yr flood in the next 95 years is:

$$R = 1 - \left(1 - \frac{1}{95}\right)^{50}$$

$$R = 0.41$$

The chance that the capacity will not be exceeded during the next 50 yrs is

1-0.41 = 0.59

Example

A bridge has an expected life of 25 years and is designed for a flood magnitude of return period 100 years. (a) What is the risk of this hydrological design? (b) If 10% risk is acceptable, what return period will have to be adopted?

(a) The risk $\bar{R}$ for n = 25 years and T = 100 years is: $\bar{R} = 1 - \left(1 - \frac{1}{100}\right)^{25} = 0.222$. Hence the inbuilt risk in this design is 22.2%.

(b) If $\bar{R}$ =10% = 0.10, $0.10 = 1 - \left(1 - \frac{1}{T}\right)^{25} \Rightarrow \left(1 - \frac{1}{T}\right)^{25} = 0.90$ and T = 238 years

(c) say 240 years. Hence to get 10% acceptable risk, the bridge will have to be designed for a flood of return period T = 240 years.

Return Period

Plotting Positions

Mostly applicable for **short return period** (T) extrapolations

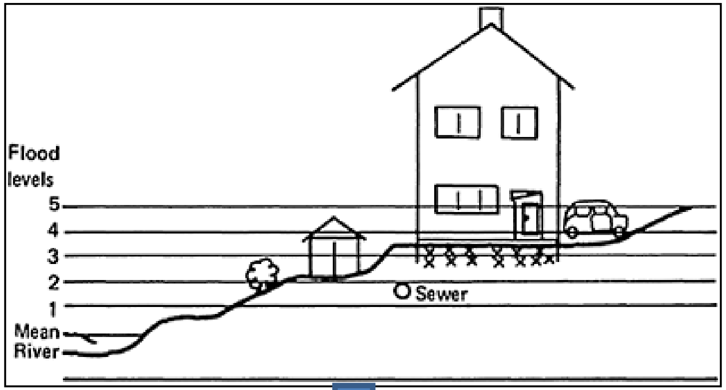
Method	P
California	m / N
Hazen	(m - 0.5) / N
Weibull	m / (N+1)
Chegodayev	(m - 0.3) / (N+0.4)
Gringorten	(m - 3/8) / (N + 1/4)

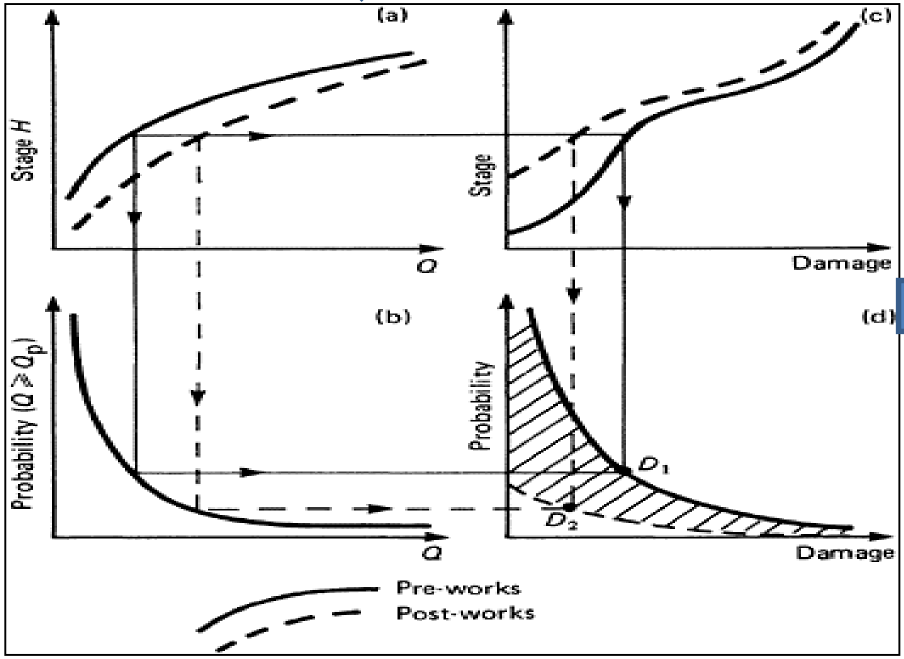
What is the probability that a Tr-year return period event will occur at least once in N years?

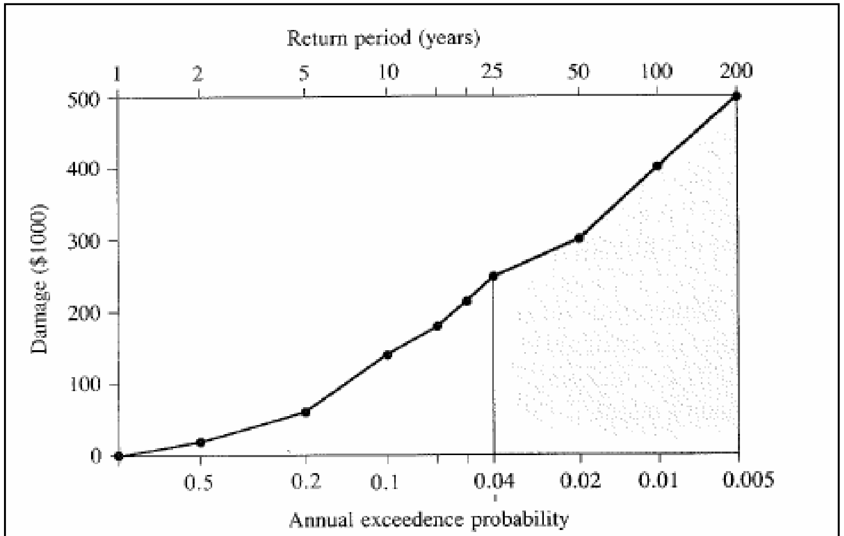
$P(X \geq x_{Tr}$ at least once in N years$) = 1-(1-p)^N$ or

$P(X \geq xTr$ at least once in N years$) = 1-[1-(1/Tr)]^N$

Economic Design







(a) Damages for events of various return periods.

(b) Hydroeconomic analysis.

Home study

1. A bridge has an expected life of 25 years and is designed for a flood magnitude of return period 100 years. (a) What is the risk of this hydrological design? (b) If 10% risk is acceptable, **what return period** will have to be adopted?
2. Analysis of annual flood series of a river yielded a sample mean of 1000 and standard deviation of 500 m3/s. Estimate the design flood of a structure(XT) on this river to provide 90% assurance that the structure **will not fail** in the next 50 years. Use Gumbel method and assume the sample size to be very large

   $\bar{y}_n = 0.577$ & $S_n = 1.2825$

Thank You !!!