GLOSSARY AND NOTATION

Glossary

The terms used in this guide include definitions taken from references listed in the Bibliography or from "Nomenclature for Hydraulics," Manual 43, American Society of Civil Engineers, 1962, and from definitions especially prepared for this guide. For more technical definitions of statistical terms, see "Dictionary of Statistical Terms" by M. G. Kendall and W. R. Buckland, Hafner Publishing Company, New York, 1957.

TERM	Definition
Annual Flood	The maximum momentary peak discharge in each year of record. (Sometimes the maximum mean daily discharge is used.)
Annual Flood Series	A list of annual floods.
Annual Series	A general term for a set of any kind of data in which each item is the maximum or minimum in a year.
Array	A list of data in order of magnitude; in flood-frequency analysis it is customary to list the largest value first, in a low-flow frequency analysis the smallest first.
Broken Record	A systematic record which is divided into separate continuous segments because of deliberate discontinuation of recording for significant periods of time.

Coefficient of Skewness	A numerical measure or index of the lack of symmetry in a frequency distribution. Function of the third moment of magnitudes about their mean, a measure of asymmetry. Also called "coefficient of skew" or "skew coefficient."
Confidence Limits	Computed values on both sides of an estimate of a parameter that show for a specified probability the range in which the true value of the parameter lies.
Distribution	Function describing the relative frequency with which events of various magnitudes occur.
Distribution-Free	Requiring no assumptions about the kind of probability distribution a set of data may have.
Exceedance Frequency	The percentage of values that exceed a specified magnitude, 100 times exceedance probability.
Exceedance Probability	Probability that a random event will exceed a specified magnitude in a given time period, usually one year unless otherwise indicated.
Expected Probability	The average of the true probabilities of all magnitude estimates for any specified flood frequency that might be made from successive samples of a specified size.
Generalized Skew Coefficient	A skew coefficient derived by a procedure which integrates values obtained at many locations.
Homogeneity	Records from the same populations.

Incomplete Record

A streamflow record in which some peak flows are missing because they were too low or high to record or the gage was out of operation for a short period because of flooding.

Level of Significance

The probability of rejecting a hypothesis when it is in fact true. At a "10-percent" level of significance the probability is 1/10.

Mean-Square Error

Sum of the squared differences between the true and estimated values of a quantity divided by the number of observations. It can also be defined as the bias squared plus the variance of the quantity. *

Method of Moments

A standard statistical computation for estimating the moment of a distribution from the data of a sample.

Nonparametric

The same as distribution-free.

Normal Distribution

A probability distribution that is symmetrical about the mean, median, and mode (bell-shaped). It is the most studied distribution in statistics, even though most data are not exactly normally distributed, because of its value in theoretical work and because many other distributions can be transformed into normal. It is also known as Gaussian, The Laplacean, The Gauss-Laplace, or the Laplace-Gauss distribution, or the Second Law of Laplace.

Outlier	Outliers (extreme events) are data points which depart from the trend of the rest of data.
Parameter	A characteristic descriptor, such as a mean or standard deviation.
Percent Chance	A probability multiplied by 100.
Population	The entire (usually infinite) number of data from which a sample is taken or collected. The total number of past, present, and future floods at a location on a river is the population of floods for that location even if the floods are not measured or recorded.
Recurrence Interval (Return Period, Exceedance Interval)	The average time interval between actual occurrences of a hydrological event of a given or greater magnitude. In an annual flood series, the average interval in which a flood of a given size is exceeded as an annual maximum. In a partial duration series, the average interval between floods of a given size, regardless of their relationship to the year or any other period of time. The distinction holds even though for large floods recurrence intervals are nearly the same for both series.
Sample	An element, part, or fragment of a "population." Every hydrologic record is a sample of a much longer record.
Skew Coefficient	See "coefficient of skewness."

Standard Deviation	A measure of the dispersion or precision of a series of statistical values such as precipitation or streamflow. It is the square root of the sum of squares of the deviations from the arithmetic mean divided by the number of values or events in the series. It is now standard practice in statistics to divide by the number of values minus one in order to get an unbiased estimate of the variance from the sample data.
Standard Error	An estimate of the standard deviation of a statistic. Often calculated from a single set of observations. Calculated like the standard deviation but differing from it in meaning.
Student's t Distribution (t-distribution)	A distribution used in evaluation of variables which involve sample standard deviation rather than population standard deviation.
Test of Significance	A test made to learn the probability that a result is accidential or that a result differs from another result. For all the many types of tests there are standard formulas and tables. In making a test it is necessary to choose a "level of significance," the choice being arbitrary but generally not less than the low level of 10 percent nor more than the high level of 1 percent.

Transformation	The change of numerical values of data to make later computations easier, to linearize a plot or to normalize a skewed distribution by making it more nearly a normal distribution. The most common transformations are those changing ordinary numerical values into their logarithms, square roots or cube roots; many others are possible.
Variance	A measure of the amount of spread or dispersi of a set of values around their mean, obtaine by calculating the mean value of the squares of the deviations from the mean, and hence equal to the square of the standard deviation
Weighted Means	A value obtained by multiplying each of a series of values by its assigned weight and dividing the sum of those products by the sum of the weights.

Notation

Appendix notation is described in each Appendix. While most notation is consistent, slight variations do occur.

Notation	Explanation
A	Fitting parameter used in equation 6.
a	Variate in equations 9 and 10 which depends upon the distribution (23).
B	Fitting parameter used in equation 6.
b	Variate in equation 9 which depends upon the distribution (23)
G	Skew coefficient of logarithms of annual peak discharges
$\overline{G}$	Generalized skew coefficient
$\tilde{G}$	Historically adjusted skew coefficient
G_W	Weighted skew coefficient
H	Historic record length
K_H	K value from Appendix 4 for historic period H
K	Pearson Type III deviate
K_N	K value from Appendix 4 for sample size N
$\tilde{M}$	Historically adjusted mean logarithm
MSE	Mean-square error
$MSE_{\overline{G}}$	Mean-square error of generalized skew
MSE_G	Mean-square error of station skew
m	Ordered sequence of flood values, with the largest equal to 1
N	Number of items in data set
P	Exceedance probability
Q	Peak discharge, cfs
S	Standard deviation of logarithms of annual peak discharges
$\tilde{S}$	Historically adjusted standard deviation

Notation

SE_G Standard error of sample skew coefficient, which for samples from a normal distribution can be estimated as:

$$SE_G = \sqrt{\frac{6N(N-1)}{(N-2)(N+1)(N+3)}}$$

SE_S Standard error of sample standard deviation, can be estimated as:

$$SE_S = \frac{S\sqrt{1 + 0.75\,G^2}}{\sqrt{2N}}$$

$SE_{\overline{X}}$ Standard error of sample mean, can be estimated as:

$$SE_{\overline{X}} = \frac{S}{\sqrt{N}}$$

T Recurrence interval in years

X Logarithm of peak flow

$\overline{X}$ Mean logarithm of peak flows

* X_H High outlier threshold in log units

X_L Low outlier threshold in log units