7.5 ANALYTICAL PROBABILITY DISTRIBUTIONS

Numerous probability distribution functions have been used to model phenomena characterized by significant variability not deterministically explained by physical principles. Many probability distribution functions for continuous random variables can be expressed as either

$$X = \overline{X} + KS \tag{7.13}$$

Eq. 11.10 in Gupta (2017)

or

$$\log X = \overline{\log X} + KS_{\log X} \tag{7.14}$$

The sample mean $\overline{X}$ and standard deviation S are estimates of the population mean μ and standard deviation σ of the random variable X. Frequency factors K are read from published tables previously developed by integrating the appropriate probability density function. $\overline{\log X}$ and $S_{\log X}$ are the mean and standard deviation of the logarithm of the random variable X. Application of Eq. 7.14 consists of transforming the data to their logarithms and applying Eq. 7.13 to the logarithms. The normal and Pearson type III distributions are applied using Eq. 7.13. The lognormal and log-Pearson type III distributions are applied using Eq. 7.14.

A distribution function provides a probabilistic model of the phenomena represented by a particular random variable. Model parameters are computed from sample observations. The parameters mean $\overline{X}$, standard deviation S, and skew coefficient G are computed from n observations X_i with the following formulas.

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{7.15}$$

$$S = \left[\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \right]^{0.5} \tag{7.16}$$

$$G = \frac{n \sum_{i=1}^{n} (X_i - \overline{X})^3}{(n-1)(n-2)S^3} \tag{7.17}$$

(*) Textbooks and references may also present the relationship as

$$Y = \overline{Y} + KS_Y$$

where $Y = \log X$ or $\ln X$

Also K-values are determined using equations or from tables for each specific "PDF"!

SOURCE: Wurbs & James (2002)

K-values for the Normal Distribution

Exceedance* Probability	K	Exceedance* Probability	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	0.700	−0.524
0.025	1.960	0.750	−0.674
0.050	1.645	0.800	−0.842
0.100	1.282	0.850	−1.036
0.150	1.036	0.900	−1.282
0.200	0.842	0.950	−1.645
0.250	0.674	0.975	−1.960
0.300	0.524	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

*Exceedance probability is the probability that an event will be equalled or exceeded, and is equal to $1/T$ where T is the return period.

Source: Roberson et al., Hydraulic Engineering, 1998

TABLE 7.3 *K* VALUES FOR THE PEARSON TYPE III AND LOG-PEARSON TYPE III DISTRIBUTIONS

Skew	Recurrence interval, years							
Coefficient	1.0101	2	5	10	25	50	100	200
G	Exceedance frequency, percent							
	99	50	20	10	4	2	1	0.5
3.0	−0.667	−0.396	0.420	1.180	2.278	3.152	4.051	4.970
2.8	−0.714	−0.384	0.460	1.210	2.275	3.114	3.973	4.847
2.6	−0.769	−0.368	0.499	1.238	2.267	3.071	3.889	4.718
2.4	−0.832	−0.351	0.537	1.262	2.256	3.023	3.800	4.584
2.2	−0.905	−0.330	0.574	1.284	2.240	2.970	3.705	4.444
2.0	−0.990	−0.307	0.609	1.302	2.219	2.912	3.605	4.298
1.8	−1.087	−0.282	0.643	1.318	2.193	2.848	3.499	4.147
1.6	−1.197	−0.254	0.675	1.329	2.163	2.780	3.388	3.990
1.4	−1.318	−0.225	0.705	1.337	2.128	2.706	3.271	3.828
1.2	−1.449	−0.195	0.732	1.340	2.087	2.626	3.149	3.661
1.0	−1.588	−0.164	0.758	1.340	2.043	2.542	3.022	3.489
0.8	−1.733	−0.132	0.780	1.336	1.993	2.453	2.891	3.312
0.6	−1.880	−0.099	0.800	1.328	1.939	2.359	2.755	3.132
0.4	−2.029	−0.066	0.816	1.317	1.880	2.261	2.615	2.949
0.2	−2.178	−0.033	0.830	1.301	1.818	2.159	2.472	2.763
N-PD → 0.0	−2.326	0.000	0.842	1.282	1.751	2.054	2.326	2.576
−0.2	−2.472	0.033	0.850	1.258	1.680	1.945	2.178	2.388
−0.4	−2.615	0.066	0.855	1.231	1.606	1.834	2.029	2.201
−0.6	−2.755	0.099	0.857	1.200	1.528	1.720	1.880	2.016
−0.8	−2.891	0.132	0.856	1.166	1.448	1.606	1.733	1.837
−1.0	−3.022	0.164	0.852	1.128	1.366	1.492	1.588	1.664
−1.2	−3.149	0.195	0.844	1.086	1.282	1.379	1.449	1.501
−1.4	−3.271	0.225	0.832	1.041	1.198	1.270	1.318	1.351
−1.6	−3.388	0.254	0.817	0.994	1.116	1.166	1.197	1.216
−1.8	−3.499	0.282	0.799	0.945	1.035	1.069	1.087	1.097
−2.0	−3.605	0.307	0.777	0.895	0.959	0.980	0.990	0.995
−2.2	−3.705	0.330	0.752	0.844	0.888	0.900	0.905	0.907
−2.4	−3.800	0.351	0.725	0.795	0.823	0.830	0.832	0.833
−2.6	−3.889	0.368	0.696	0.747	0.764	0.768	0.769	0.769
−2.8	−3.973	0.384	0.666	0.702	0.712	0.714	0.714	0.714
−3.0	−4.051	0.396	0.636	0.660	0.666	0.666	0.667	0.667

SOURCE: Wurbs & James, 2012

TABLE 7.1 MAXIMUM ANNUAL DISCHARGE IN THE MISSISSIPPI RIVER AT ST. LOUIS

Year	Flow (m^3/s)	Year	Flow (m^3/s)	Year	Flow (m^3/s)
1933	12,400	1955	8,800	1977	11,000
1934	6,260	1956	5,860	1978	16,200
1935	18,500	1957	9,620	1979	19,500
1936	9,450	1958	14,300	1980	9,930
1937	10,600	1959	10,300	1981	14,400
1938	12,300	1960	19,000	1982	20,700
1939	15,100	1961	16,700	1983	20,300
1940	5,240	1962	16,700	1984	16,400
1941	14,000	1963	8,510	1985	19,500
1942	18,900	1964	8,710	1986	20,500
1943	23,700	1965	15,600	1987	11,900
1944	23,700	1966	10,500	1988	8,850
1945	17,400	1967	15,000	1989	9,280
1946	14,200	1968	9,790	1990	17,000
1947	22,300	1969	17,500	1991	12,400
1948	17,900	1970	15,300	1992	14,600
1949	12,000	1971	11,900	1993	30,600
1950	13,100	1972	11,500	1994	17,000
1951	22,200	1973	24,200	1995	22,500
1952	19,400	1974	16,500	1996	17,400
1953	10,400	1975	13,700	1997	15,400
1954	8,230	1976	12,700	1998	15,500

Example 7.5

Estimate the 10-year and 100-year recurrence interval peak flows on the Mississippi River at St. Louis using the data from Table 7.1. Model the flows alternatively with the normal, log-normal, log-Pearson type III, and Gumbel probability distributions.

Solution Equations 7.15 and 7.16 are applied to compute the mean and standard deviation for the 66 flows for use in the normal and Gumbel distributions.

$$\overline{X} = 14{,}776 \text{ m}^3\text{/s}$$

$$S = 5{,}242 \text{ m}^3\text{/s}$$

The base 10 logarithms of each of the 66 flows are computed and substituted into Eqs. 7.15, 7.16, and 7.17 to obtain values for the parameters of the log-normal and log-Pearson type III distributions.

$$\overline{\log X} = 4.149$$

$$S_{\log X} = 0.1511$$

$$G_{\log X} = -0.427$$

Since this textbook does not provide a normal probability table, Table 7.3 is used with $G = 0.0$ to obtain the frequency factor K for the normal and log-normal distributions. The K for T of 10 and 100 years are 1.282 and 2.326, respectively. Values of K for the log-Pearson type III distribution obtained from linear interpolation of Table 7.3 for G of -0.427 and T of 10 and 100 years are 1.227 and 2.009.

Normal Distribution

$$Q_{10\text{ years}} = \overline{Q} + KS = 14{,}776 + (1.282)(5{,}242) = 21{,}500\,\frac{\text{m}^3}{\text{s}}$$

$$Q_{100\text{ years}} = \overline{Q} + KS = 14{,}776 + (2.326)(5{,}242) = 27{,}000\,\frac{\text{m}^3}{\text{s}}$$

Log-Normal Distribution

$$\log Q_{10\text{ years}} = \overline{\log Q} + KS_{\log Q} = 4.149 + (1.282)(0.1511) = 4.343$$

$$Q_{10\text{ years}} = 10^{4.343} = 22{,}000\,\frac{\text{m}^3}{\text{s}}$$

$$\log Q_{100\text{ years}} = \overline{\log Q} + KS_{\log Q} = 4.149 + (2.326)(0.1511) = 4.5005$$

$$Q_{100\text{ years}} = 10^{4.5005} = 31{,}700\,\frac{\text{m}^3}{\text{s}}$$

Log-Pearson Type III Distribution

$$\log Q_{10\text{ years}} = \overline{\log Q} + KS_{\log Q} = 4.149 + (1.227)(0.1511) = 4.334$$

$$Q_{10\text{ years}} = 21{,}600\,\frac{\text{m}^3}{\text{s}}$$

$$\log Q_{100\text{ years}} = \overline{\log Q} + KS_{\log Q} = 4.149 + (2.009)(0.1511) = 4.453$$

$$Q_{100\text{ years}} = 28{,}300\,\frac{\text{m}^3}{\text{s}}$$

Gumbel Distribution

$$P = 1 - e^{-e^{-b}}$$

$$b = \frac{1}{0.7797S}(X - \overline{X} + 0.45S)$$

Gumbel, T = *10 Years*

$$0.10 = 1 - e^{-e^{-b}}$$

$$e^{-e^{-b}} = 0.9$$

$$\ln(e^{-e^{-b}}) = \ln(0.9)$$

$$-e^{-b} = -0.1054$$

$$\ln(e^{-b}) = \ln(0.1054)$$

$$b = 2.2504$$

$$b = \frac{1}{0.7797S}(X - \overline{X} + 0.45S)$$

$$2.2504 = \frac{1}{0.7797(5242)}(Q_{10\text{yr}} - 14{,}776 + 0.45(5{,}242))$$

$$Q_{10\text{ years}} = 21{,}600\,\frac{\text{m}^3}{\text{s}}$$

Gumbel, T = *100 Years*

$$0.01 = 1 - e^{-e^{-b}}$$

$$b = 4.600$$

$$4.600 = \frac{1}{0.7797(5{,}242)}(Q_{100\text{yr}} - 14{,}776 + 0.45(5{,}242))$$

$$Q_{100\text{ years}} = 31{,}200\,\frac{\text{m}^3}{\text{s}}$$

EXAMPLE 7.5 SOLUTION SUMMARY

Probability distribution	T = 10 years Q, m³/s	T = 100 years Q, m³/s
Normal	21,500	27,000
Log-normal	22,000	31,700
Log-Pearson type III	21,600	28,300
Gumbel	21,600	31,200

Example 7.6

Estimate the annual exceedance probability and recurrence interval for a flow of 25,000 m³/s alternatively using the normal, log-normal, log-Pearson type III, and Gumbel probability distributions.

Normal Distribution

$$X = \overline{X} + KS$$
$$25{,}000 = 14{,}776 + K(5{,}242)$$
$$K = 1.9504$$

Linear interpolation of Table 7.3 with $G = 0.0$ yields: $T = 31$ years and $P = 3.2\%$

Log-Normal

$$\log X = \overline{\log X} + KS_{\log X}$$
$$\log(25{,}000) = 4.149 + K(0.1511)$$
$$K = 1.6475$$

Linear interpolation of Table 7.3 with $G = 0.0$ yields: $T = 22$ years and $P = 4.6\%$

Log-Pearson Type III

$$\log X = \overline{\log X} + KS_{\log X}$$
$$\log(25{,}000) = 4.149 + K(0.1511)$$
$$K = 1.6475$$

Linear interpolation of Table 7.3 with $G = -0.427$ yields: $T = 31$ years and $P = 3.2\%$

Gumbel

$$b = \frac{1}{0.7797S}(X - \overline{X} + 0.45S) = \frac{1}{0.7797(5{,}242)}(25{,}000 - 14{,}776 + 0.45(5{,}242)) = 3.0786$$

$$P = 1 - e^{-e^{-b}} = 1 - e^{-e^{-3.0786}} = 0.0450$$

$$T = \frac{1}{P} = \frac{1}{0.045} = 22 \text{ years}$$

EXAMPLE 7.6 SOLUTION SUMMARY

Probability distribution	25,000 m³/s	
	T	P
Normal	31 years	0.032
Log-normal	22 years	0.046
Log-Pearson type III	31 years	0.032
Gumbel	22 years	0.045

SOURCE: Wurbs & James, Water Resources Engineering, Prentice-Hall, 2002