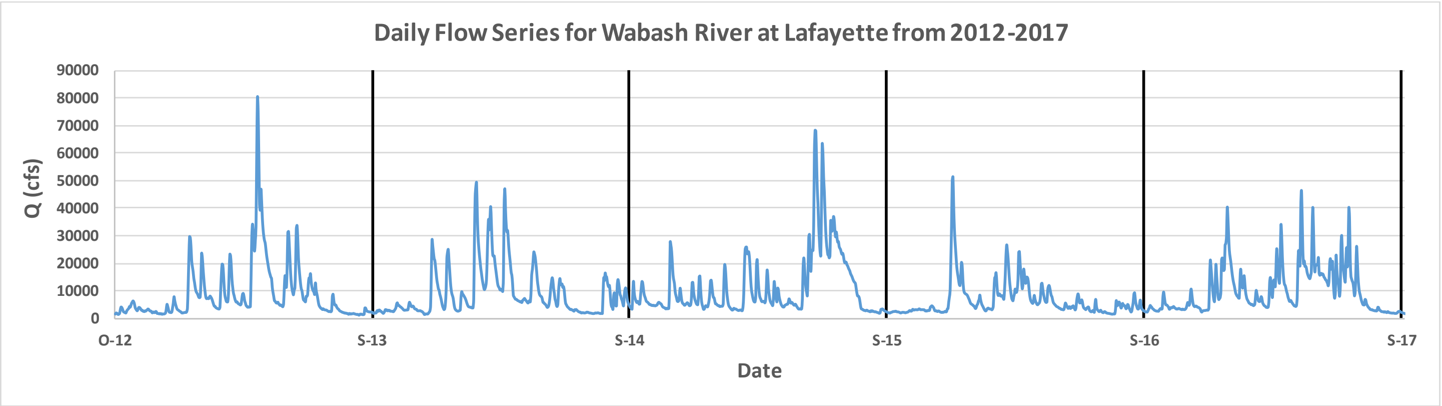
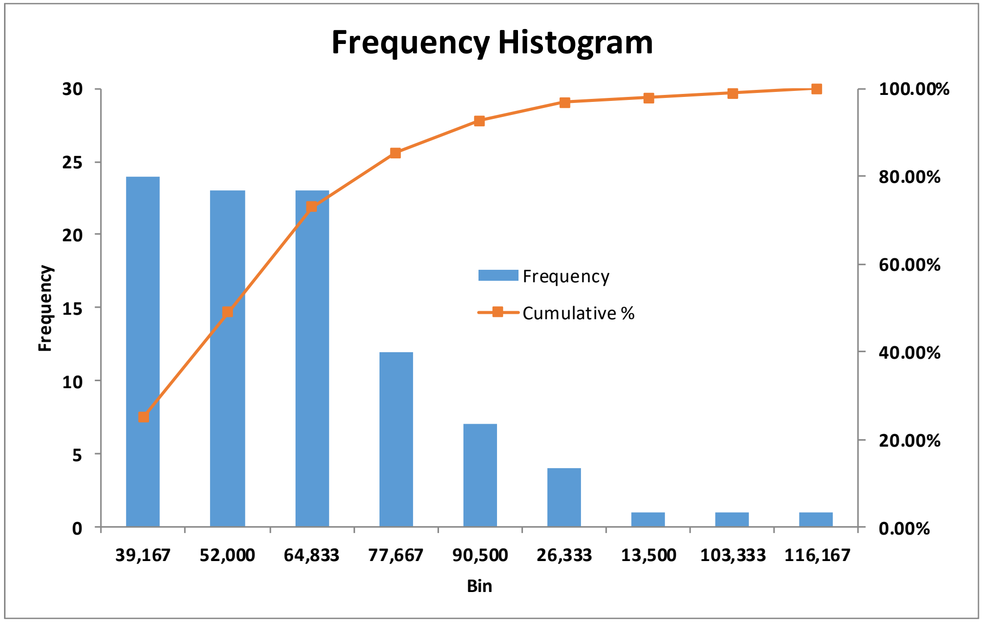
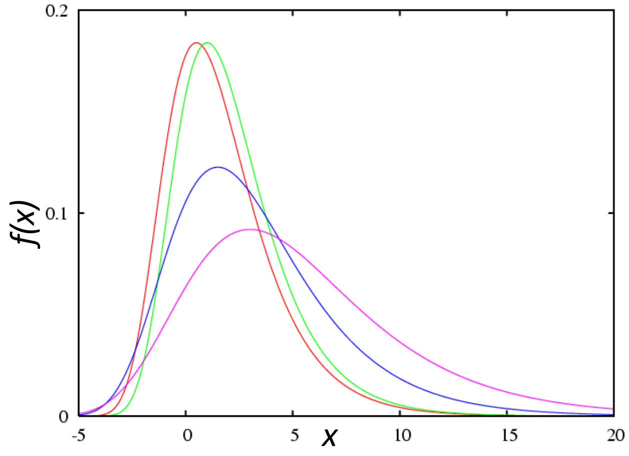


Flood Frequency Analysis - 2





FFA Using Sample Data

- What is Q_2 for this site?
- What is Q_{50} or Q_{100}?

Year	Data	Sorted	Rank	EP	T
1923	38400	129000	1	0.03	31.00
1924	58400	88000	2	0.06	15.50
1925	61900	87200	3	0.10	10.33
1926	57000	73600	4	0.13	7.75
1927	63500	73500	5	0.16	6.20
1928	34100	71500	6	0.19	5.17
1929	38500	67500	7	0.23	4.43
1930	73500	63500	8	0.26	3.88
1931	16500	61900	9	0.29	3.44
1932	40100	59800	10	0.32	3.10
1933	67500	59600	11	0.35	2.82
1934	20400	58400	12	0.39	2.58
1935	36700	58000	13	0.42	2.38
1936	87200	57000	14	0.45	2.21
1937	58000	**50000**	**15**	**0.48**	**2.07**
1938	59800	46000	16	0.52	1.94
1939	73600	43600	17	0.55	1.82
1940	33300	41500	18	0.58	1.72
1941	13500	40100	19	0.61	1.63
1942	43600	40000	20	0.65	1.55
1943	129000	39700	21	0.68	1.48
1944	71500	38500	22	0.71	1.41
1945	46000	38400	23	0.74	1.35
1946	35200	36700	24	0.77	1.29
1947	40000	35200	25	0.81	1.24
1948	39700	34100	26	0.84	1.19
1949	59600	33300	27	0.87	1.15
1950	88000	20400	28	0.90	1.11
1951	50000	16500	29	0.94	1.07
1952	41500	13500	30	0.97	1.03

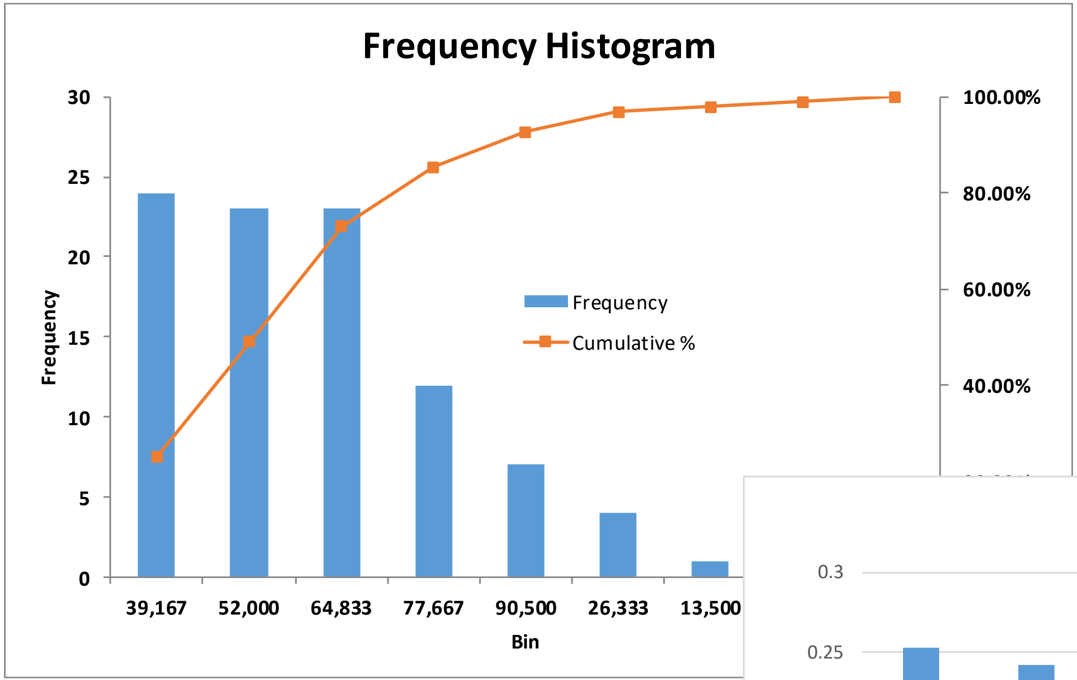
Year	Data	Sorted	Rank	EP	T
1953	34600	97000	1	0.03	31.00
1954	17900	87600	2	0.06	15.50
1955	35000	67100	3	0.10	10.33
1956	29400	65300	4	0.13	7.75
1957	51900	64900	5	0.16	6.20
1958	97000	64100	6	0.19	5.17
1959	87600	60000	7	0.23	4.43
1960	38100	57600	8	0.26	3.88
1961	54700	55300	9	0.29	3.44
1962	45000	54700	10	0.32	3.10
1963	60000	52500	11	0.35	2.82
1964	57600	51900	12	0.39	2.58
1965	36000	49500	13	0.42	2.38
1966	64100	48800	14	0.45	2.21
1967	65300	**46200**	**15**	**0.48**	**2.07**
1968	67100	45000	16	0.52	1.94
1969	64900	43400	17	0.55	1.82
1970	41700	43200	18	0.58	1.72
1971	30000	41700	19	0.61	1.63
1972	38500	39500	20	0.65	1.55
1973	39500	38500	21	0.68	1.48
1974	52500	38100	22	0.71	1.41
1975	34700	36000	23	0.74	1.35
1976	43400	35000	24	0.77	1.29
1977	31400	34700	25	0.81	1.24
1978	49500	34600	26	0.84	1.19
1979	48800	31400	27	0.87	1.15
1980	46200	30000	28	0.90	1.11
1981	43200	29400	29	0.94	1.07
1982	55300	17900	30	0.97	1.03

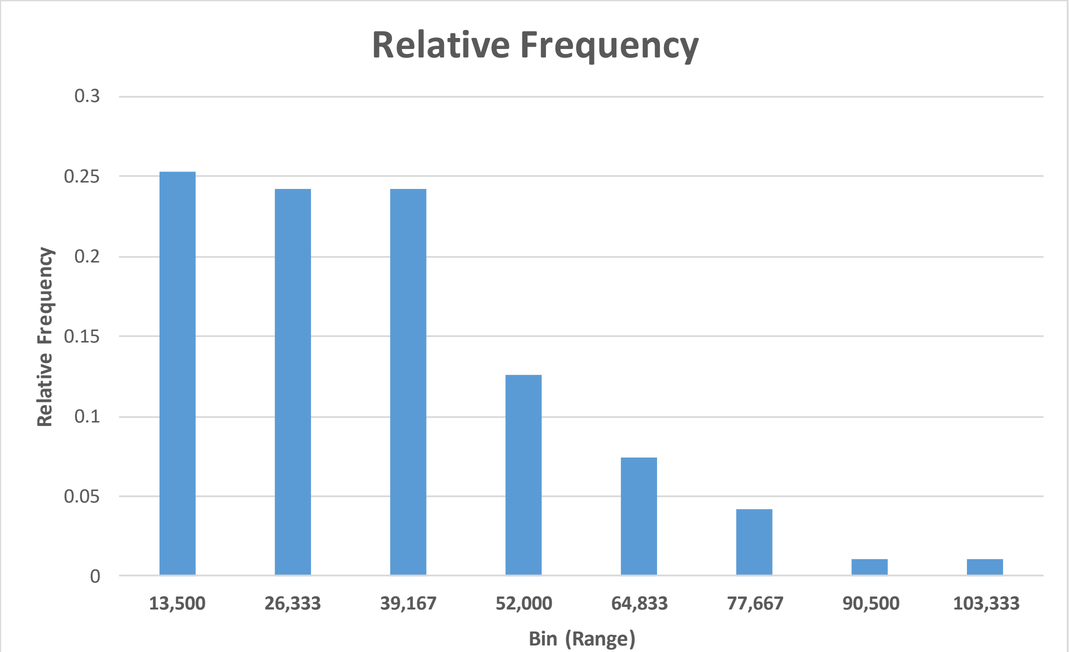
Year	Data	Sorted	Rank	EP	T
1983	59100	79200	1	0.03	31.00
1984	40000	79000	2	0.06	15.50
1985	79200	78300	3	0.10	10.33
1986	30800	75200	4	0.13	7.75
1987	23600	71000	5	0.16	6.20
1988	32600	70300	6	0.19	5.17
1989	39500	63500	7	0.23	4.43
1990	75200	61100	8	0.26	3.88
1991	70300	59100	9	0.29	3.44
1992	53000	58600	10	0.32	3.10
1993	61100	57800	11	0.35	2.82
1994	63500	57100	12	0.39	2.58
1995	29500	56800	13	0.42	2.38
1996	34600	54000	14	0.45	2.21
1997	53000	**53000**	**15**	**0.48**	**2.07**
1998	54000	53000	16	0.52	1.94
1999	58600	49700	17	0.55	1.82
2000	30000	42000	18	0.58	1.72
2001	35100	40000	19	0.61	1.63
2002	39600	39600	20	0.65	1.55
2003	79000	39500	21	0.68	1.48
2004	56800	35100	22	0.71	1.41
2005	78300	34600	23	0.74	1.35
2006	33400	34600	24	0.77	1.29
2007	49700	33400	25	0.81	1.24
2008	71000	32600	26	0.84	1.19
2009	57800	30800	27	0.87	1.15
2010	42000	30000	28	0.90	1.11
2011	57100	29500	29	0.94	1.07
2012	34600	23600	30	0.97	1.03

Sample vs population

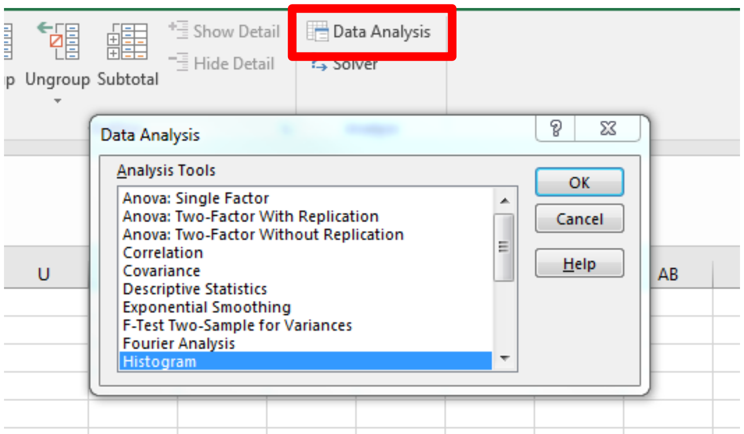
- FFA using sample data requires large sample (>200 years at least) for computing Q_{50} or Q100.
- If we can assume the parent or population distribution for our data, we can extract as many random values as possible
- How do we know the parent probability distribution?

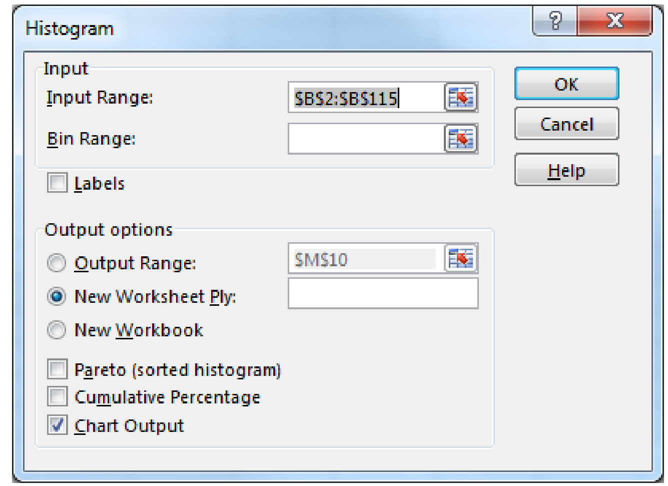
Frequency Histogram and Relative Frequency





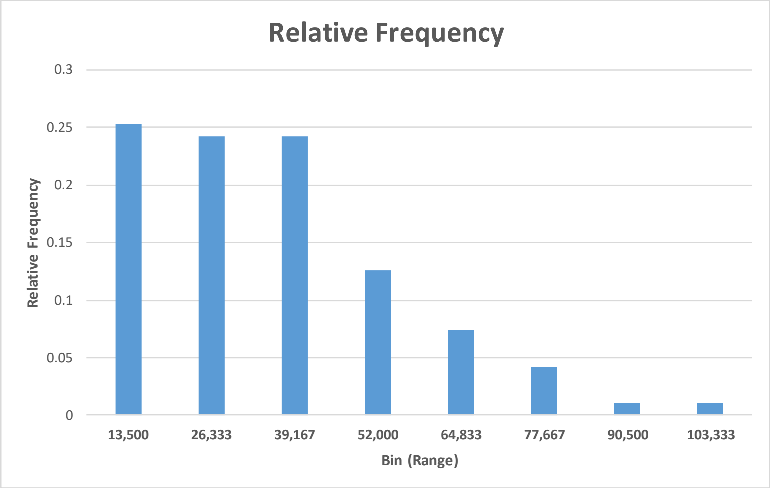
Creating Histogram/Relative Frequency Plot for a sample





- Use the Data Analysis ToolPak in Excel to use the Histogram tool
- Input range is your entire annual maximum series, you can define bins for leave that empty, use new worksheet for displaying the histogram and select chart output.
- You will see a table and the histogram in a new sheet.
- The table gives you the frequency (total number of values within the specified bin). If you divide the frequency by the total number of data points, you will get probability. Plot of these probability against the bin size will give you a probability density plot

How to fit a PDF to Annual Maximum Series?



The properties of sample (mean, standard deviation, skewness, kurtosis) could be used to describe the nature of the parent distribution, which can then provide us flow values for any probability.

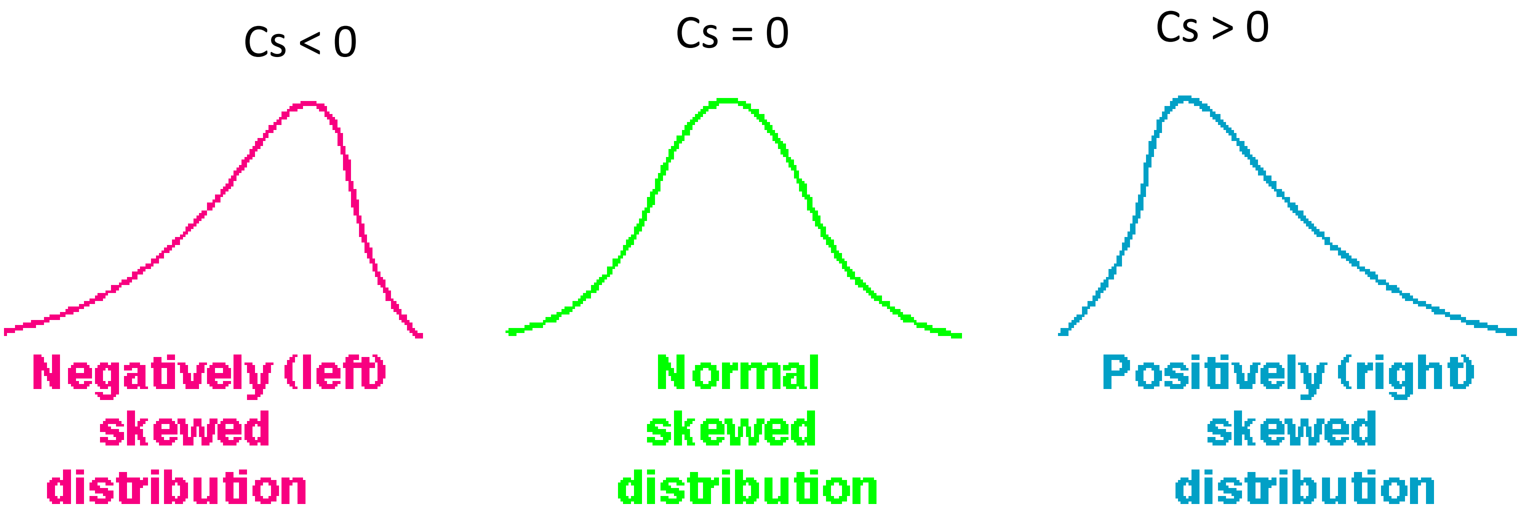
Equations for mean, st. dev. and skewness

Sample mean $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, where n is the sample size.

Sample standard deviation $$S_x = \left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{0.5}$$

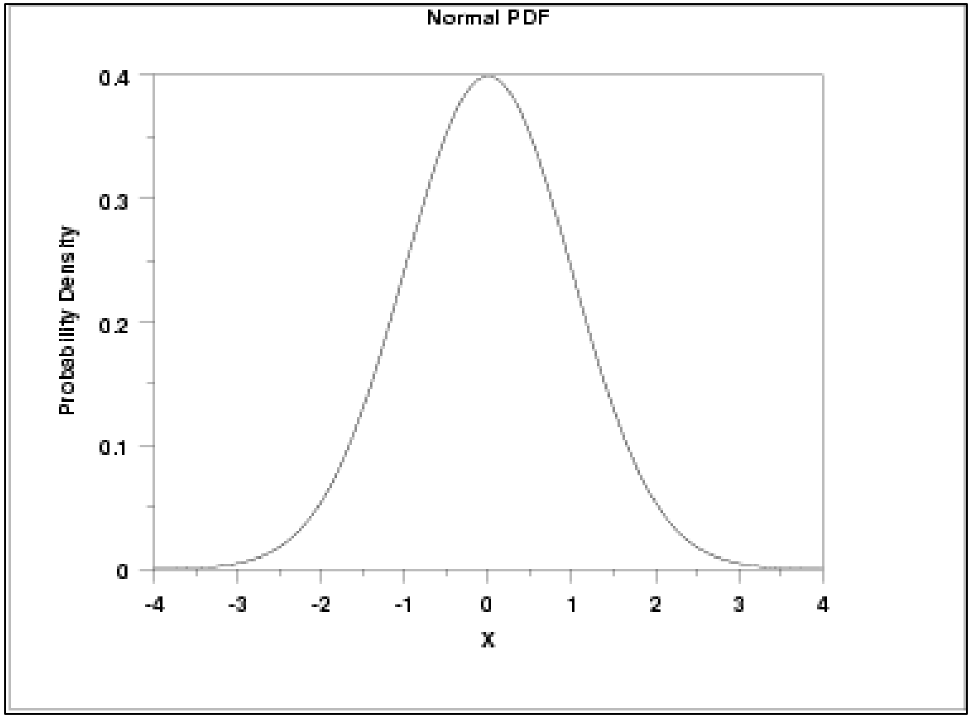
Coefficient of skewness $$C_s = \frac{n\sum_{i=1}^{n}(x_i - \bar{x})^3}{(n-1)(n-2)S^3}$$

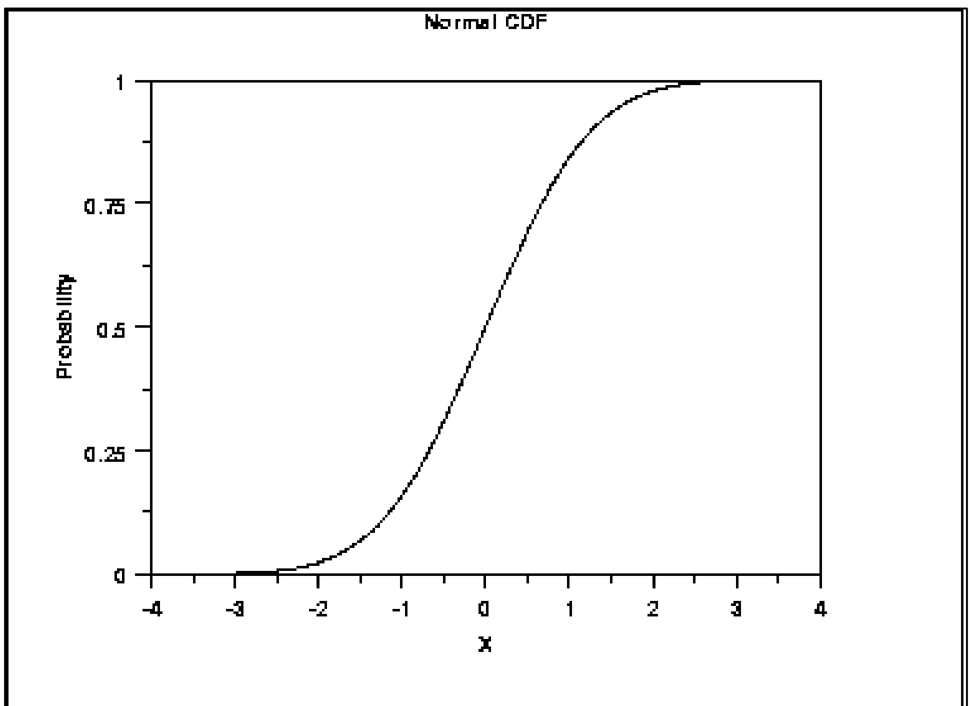
Skewed Distribution



Cs = Coefficeint of Skewness
Annual peak flow series is usually positively skewed

Source: http://researchbasics.education.uconn.edu/normal-distribution/





PDF and CDF

Discrete Probability

$$\mathrm{P}(\mathrm{X} = x) = f(x)$$

Continuous Probability

$$\mathrm{P}(a \leq X \leq b) = \int_a^b f(x)$$

Cumulative Density Function

$$F(x) = \mathrm{P}(\mathrm{X} \leq x)$$

$$1 - F(x) = \mathrm{P}(\mathrm{X} \geq x)$$

$$\mathrm{EP} = 1 - F(x)$$

$$\frac{1}{T} = 1 - F(x)$$

$$F(x) = 1 - \frac{1}{T}$$

$$x_T = F^{-1}\left(1 - \frac{1}{T}\right)$$

Probability Distribution Models

- Normal
- Log Normal
- Exponential
- Gamma
- **Log Pearson III**
- Gumbel
- **Extreme Value Type 1**
- …

Each distribution has a different equation describing the shape of the probability density function and cumulative distribution function.

$$f(x) = \ldots$$

$$\mathrm{F}(x) = \ldots$$

Analysts should select the model that best fits the variable of interest

Annual Peak Flow is commonly modeled using

- **Extreme Value I**
- **Log Pearson III** (recommended by Bulletin 17B)

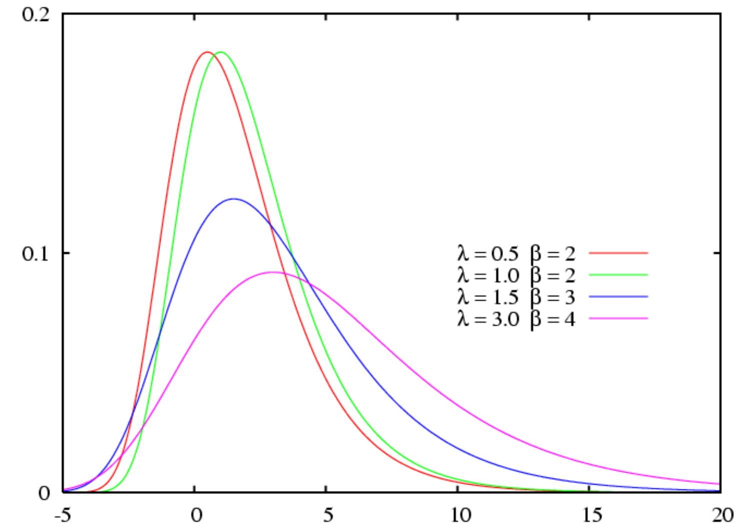
EV type I distribution

- If M_1, M_2..., M_n be a set of daily rainfall or streamflow, and let X = max(Mi) be the maximum for the year. If M_i are independent and identically distributed, then for large n, X has an extreme value type I or Gumbel distribution.

$$f(x) = \frac{1}{\alpha}\exp\left[-\frac{x-u}{\alpha} - \exp\left(-\frac{x-u}{\alpha}\right)\right]$$

$$\alpha = \frac{\sqrt{6}s_x}{\pi} \qquad u = \bar{x} - 0.5772\alpha$$

$$F(x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right]$$



Distribution of annual maximum streamflow follows an EV1 distribution

Expression for x_T using EV1

$$F(x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right] = 1 - EP = 1 - \frac{1}{T}$$

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

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

EV1 is an invertible function – we can simplify F (x) to get value of x for any T.

Pearson Type III

- Named after the statistician Pearson, it is also called three-parameter gamma distribution. A lower bound is introduced through the third parameter (ε)

$$f(x) = \frac{\lambda^{\beta}(x-\varepsilon)^{\beta-1}e^{-\lambda(x-\varepsilon)}}{\Gamma(\beta)} \qquad x \geq \varepsilon; \Gamma = gamma\ \ function$$

It is also a skewed distribution first applied in hydrology for describing the pdf of annual maximum flows.

Log-Pearson Type III

- If y = log X follows a Person Type III distribution, then X is said to have a log-Pearson Type III distribution

$$f(y) = \frac{\lambda^{\beta}(y-\beta)^{\beta-1} e^{-\lambda(y-\varepsilon)}}{\Gamma(\beta)}, y \geq \varepsilon$$

Frequency Factor Method

- Expression for x_T can be derived only if the distribution is invertible, many are not.
- Once a distribution has been selected and its parameters estimated, then how do we use it?
- Chow proposed using: $x_T = \bar{x} + K_T s_x$
- where

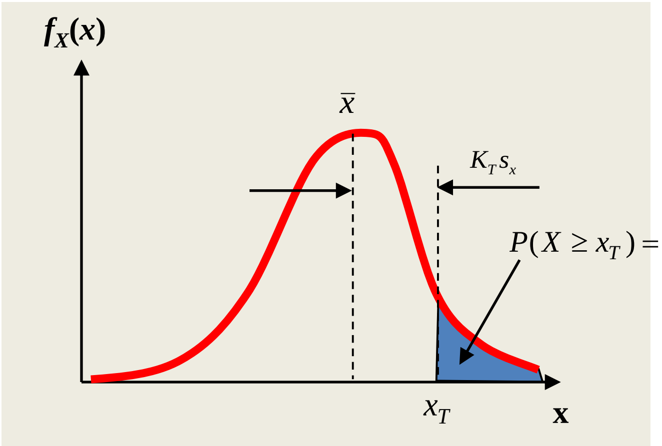
x_T = Estimated event magnitude

K_T = Frequency factor

T = Return period

$\bar{x}$ = Sample mean

s = Sample standard deviation



k_T for EV1 Distribution

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

$$x_T = \bar{x} + K_T s_x$$

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

Using LP III for Flood Frequency Analysis

- Usually x (Q in hydrology) is given
- Develop y = log (x) series
- Find y_{bar} and S_y
- Then find y_T using the frequency factor method, $y_T = \bar{y} + K_T s_y$
- Then $x_T = 10$ ^ y_T

Q: How to find k_T for Pearson Type III distribution?
A: Table 12.3.1 in Chow Maidment Mays gives kT if you know T (return period) and c_s for your data.

TABLE 12.3.1
K_T values for Pearson Type III distribution (positive skew)

Skew coefficient C_s or C_w	Return period in years						
	2	5	10	25	50	100	200
	Exceedence probability						
	0.50	0.20	0.10	0.04	0.02	0.01	0.005
3.0	−0.396	0.420	1.180	2.278	3.152	4.051	4.970
2.9	−0.390	0.440	1.195	2.277	3.134	4.013	4.909
2.8	−0.384	0.460	1.210	2.275	3.114	3.973	4.847
2.7	−0.376	0.479	1.224	2.272	3.093	3.932	4.783
2.6	−0.368	0.499	1.238	2.267	3.071	3.889	4.718
2.5	−0.360	0.518	1.250	2.262	3.048	3.845	4.652
2.4	−0.351	0.537	1.262	2.256	3.023	3.800	4.584
2.3	−0.341	0.555	1.274	2.248	2.997	3.753	4.515
2.2	−0.330	0.574	1.284	2.240	2.970	3.705	4.444
2.1	−0.319	0.592	1.294	2.230	2.942	3.656	4.372
2.0	−0.307	0.609	1.302	2.219	2.912	3.605	4.298
1.9	−0.294	0.627	1.310	2.207	2.881	3.553	4.223
1.8	−0.282	0.643	1.318	2.193	2.848	3.499	4.147
1.7	−0.268	0.660	1.324	2.179	2.815	3.444	4.069
1.6	−0.254	0.675	1.329	2.163	2.780	3.388	3.990
1.5	−0.240	0.690	1.333	2.146	2.743	3.330	3.910
1.4	−0.225	0.705	1.337	2.128	2.706	3.271	3.828
1.3	−0.210	0.719	1.339	2.108	2.666	3.211	3.745
1.2	−0.195	0.732	1.340	2.087	2.626	3.149	3.661
1.1	−0.180	0.745	1.341	2.066	2.585	3.087	3.575
1.0	−0.164	0.758	1.340	2.043	2.542	3.022	3.489
0.9	−0.148	0.769	1.339	2.018	2.498	2.957	3.401
0.8	−0.132	0.780	1.336	1.993	2.453	2.891	3.312
0.7	−0.116	0.790	1.333	1.967	2.407	2.824	3.223
0.6	−0.099	0.800	1.328	1.939	2.359	2.755	3.132
0.5	−0.083	0.808	1.323	1.910	2.311	2.686	3.041
0.4	−0.066	0.816	1.317	1.880	2.261	2.615	2.949
0.3	−0.050	0.824	1.309	1.849	2.211	2.544	2.856
0.2	−0.033	0.830	1.301	1.818	2.159	2.472	2.763

TABLE 12.3.1 (cont.)

K_T values for Pearson Type III distribution (negative skew)

Skew coefficient C_s or C_w	Return period in years						
	2	5	10	25	50	100	200
	Exceedence probability						
	0.50	0.20	0.10	0.04	0.02	0.01	0.005
−0.1	0.017	0.846	1.270	1.716	2.000	2.252	2.482
−0.2	0.033	0.850	1.258	1.680	1.945	2.178	2.388
−0.3	0.050	0.853	1.245	1.643	1.890	2.104	2.294
−0.4	0.066	0.855	1.231	1.606	1.834	2.029	2.201
−0.5	0.083	0.856	1.216	1.567	1.777	1.955	2.108
−0.6	0.099	0.857	1.200	1.528	1.720	1.880	2.016
−0.7	0.116	0.857	1.183	1.488	1.663	1.806	1.926
−0.8	0.132	0.856	1.166	1.448	1.606	1.733	1.837
−0.9	0.148	0.854	1.147	1.407	1.549	1.660	1.749
−1.0	0.164	0.852	1.128	1.366	1.492	1.588	1.664
−1.1	0.180	0.848	1.107	1.324	1.435	1.518	1.581
−1.2	0.195	0.844	1.086	1.282	1.379	1.449	1.501
−1.3	0.210	0.838	1.064	1.240	1.324	1.383	1.424
−1.4	0.225	0.832	1.041	1.198	1.270	1.318	1.351
−1.5	0.240	0.825	1.018	1.157	1.217	1.256	1.282
−1.6	0.254	0.817	0.994	1.116	1.166	1.197	1.216
−1.7	0.268	0.808	0.970	1.075	1.116	1.140	1.155
−1.8	0.282	0.799	0.945	1.035	1.069	1.087	1.097
−1.9	0.294	0.788	0.920	0.996	1.023	1.037	1.044
−2.0	0.307	0.777	0.895	0.959	0.980	0.990	0.995
−2.1	0.319	0.765	0.869	0.923	0.939	0.946	0.949
−2.2	0.330	0.752	0.844	0.888	0.900	0.905	0.907
−2.3	0.341	0.739	0.819	0.855	0.864	0.867	0.869
−2.4	0.351	0.725	0.795	0.823	0.830	0.832	0.833
−2.5	0.360	0.711	0.771	0.793	0.798	0.799	0.800
−2.6	0.368	0.696	0.747	0.764	0.768	0.769	0.769
−2.7	0.376	0.681	0.724	0.738	0.740	0.740	0.741
−2.8	0.384	0.666	0.702	0.712	0.714	0.714	0.714
−2.9	0.390	0.651	0.681	0.683	0.689	0.690	0.690
−3.0	0.396	0.636	0.666	0.666	0.666	0.667	0.667