CE 374 K – Hydrology

Frequency Analysis

Daene C. McKinney

Extreme Events

- Extreme events: $\text{Magnitude} \propto \dfrac{1}{\text{Frequency of occurence}}$

- Magnitude is related to frequency through a probability distribution
- Assumptions: Events are independent and identically distributed (IID)

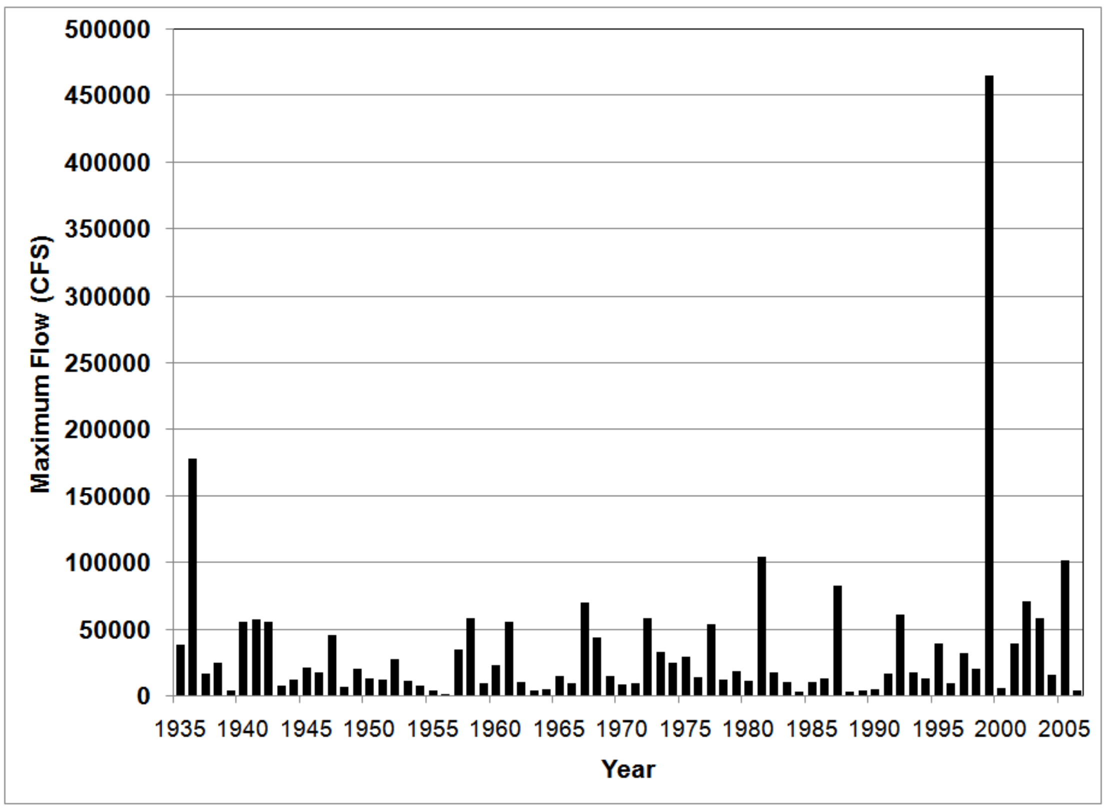
Return Period

- Random variable, X
- Realization, x
- Threshold level x_T
- Extreme event occurs if $X \geq x_T$
- Recurrence interval
  - τ = time between occurrences of $X \geq x_T$
- Return Period, E[τ] = Average recurrence interval

Guadalupe River near Victoria

- **Example: USGS Station 08176500**

  http://nwis.waterdata.usgs.gov/tx/nwis/peak



Exceedence Year	Recurrence Interval
1936	
1940	4
1941	1
1942	1
1958	16
1961	3
1967	6
1972	5
1977	5
1981	4
1987	6
1992	5
1999	7
2002	3
2003	1
2005	2
Number	16
Years (05-36)	69

$$x_T = 50{,}000$$

Exceeded 16 times

16 recurrence intervals in 69 years

$$\bar{\tau} \approx \frac{69}{16} = 4.3\,\text{years}$$

Note: Book estimate (up through 1978 = 5.1 yrs)

Probability of Occurrence

- Probability of an event happening is related to the return period
- Return Period: $E(\tau) = T = \frac{1}{p}$ $\quad p = \Pr(X \geq x_T) = \frac{1}{T}$
- Example:
  - Pr[Max. Discharge in Guad. Riv. > 50K cfs in any year] $\approx \frac{1}{4.3} = 0.23^*$

    $^*1978\ estimate \approx \frac{1}{5.1} = 0.19$

  - Pr[Max. Discharge in Guad. Riv. > 50K cfs at least once in 3 years]

    $= 1 - (1 - 0.2326)^3 = 0.55^{**}$

    $^{**}1978\ estimate \approx 0.48$

Data Series

- Complete duration series:
  - all data available
- Partial duration series:
  - greater than base value
- Annual exceedence series:
  - Partial duration series with # of values = # years
- Extreme value series
  - largest or smallest values in equal intervals
    - Annual series: interval = 1 year
    - Annual maximum series: largest values
    - Annual minimum series : smallest values

Extreme Value Distributions

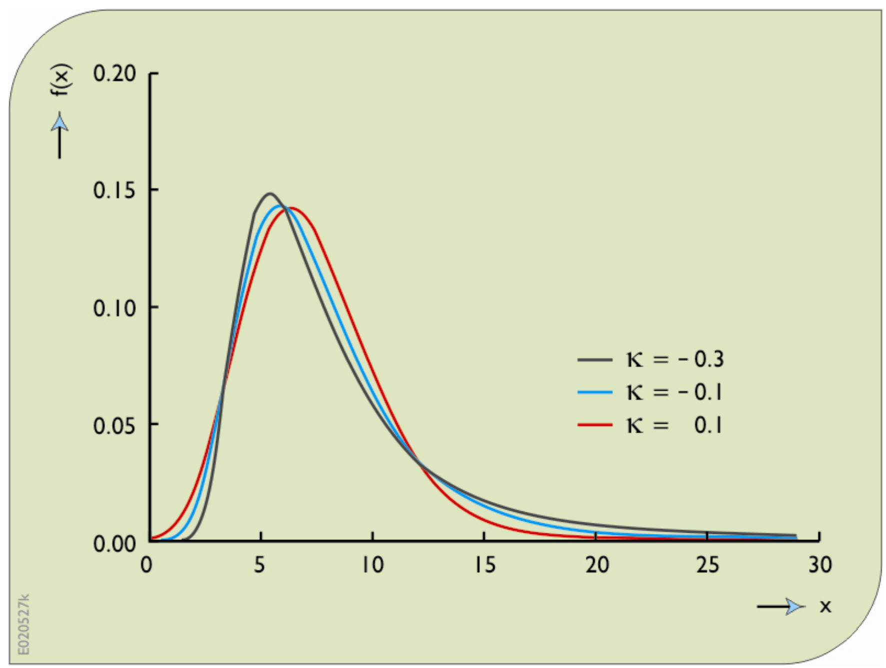
- Consider *N* samples of a random variable
- Put them in order of magnitude
- *Extreme* values: *largest* and *smallest*
- Limiting distributions: EV-I, EV-II, and EV-III

$$F(x) = \exp\left[-\left(1 - k\frac{x-u}{\alpha}\right)^{1/k}\right]$$

- k = 0: EV-I (Gumbel)

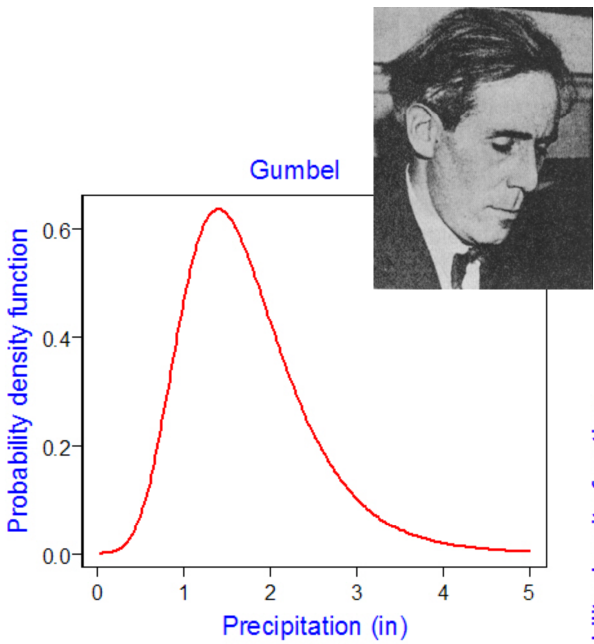
$$F(x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right]$$

- k < 0: EV-II (Frechet)
- k > 0: EV-III (Weibull)

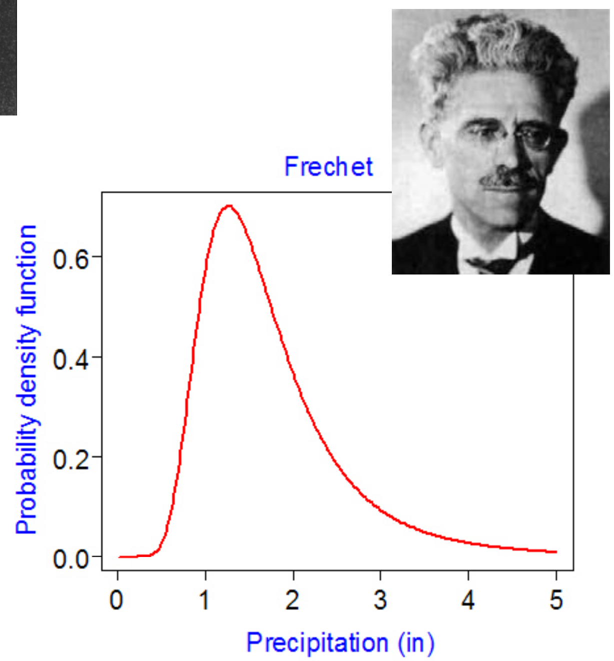


It seems that the rivers know the theory.
It only remains to convince the engineers of the validity of this analysis.
Emil Gumbel

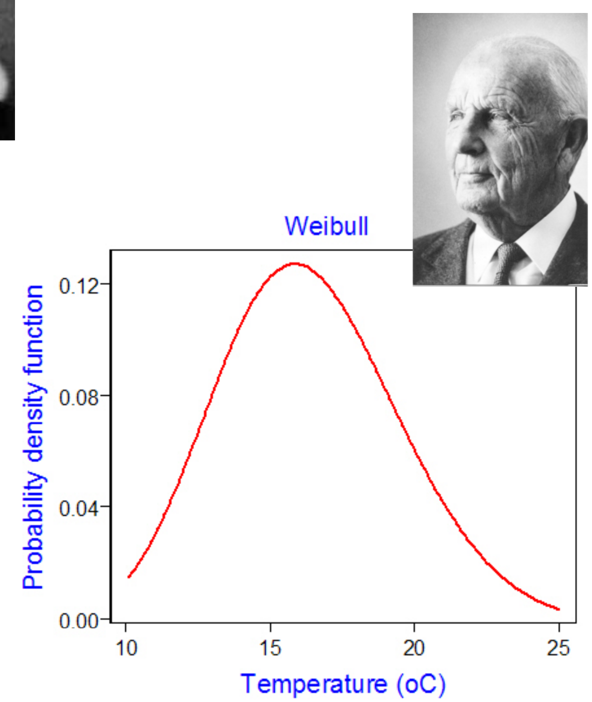
Extreme Value Distributions



EV-I (Gumbel)
Storm Rainfall



EV-II (Frechet)



EV-III (Weibull)
Droughts

EV-I (Gumbell) Distribution

- Often used for maximum type events (floods)

$$F(x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right] \qquad \alpha = \frac{\sqrt{6}s}{\pi} \qquad u = \bar{x} - 0.5772\alpha$$

  - Reduced variate $y = \frac{x-u}{\alpha}$ → $F(x) = \exp[-\exp(-y)]$

$$y = -\ln\left[\ln\left(\frac{1}{F(x)}\right)\right] \rightarrow p = \Pr(X \geq x_T) = \frac{1}{T} \qquad \frac{1}{T} = 1 - F(x_T) \rightarrow F(x_T) = \frac{T-1}{T}$$

$$y_T = -\ln\left[\ln\left(\frac{T}{T-1}\right)\right] \rightarrow x_T = u + \alpha y_T$$

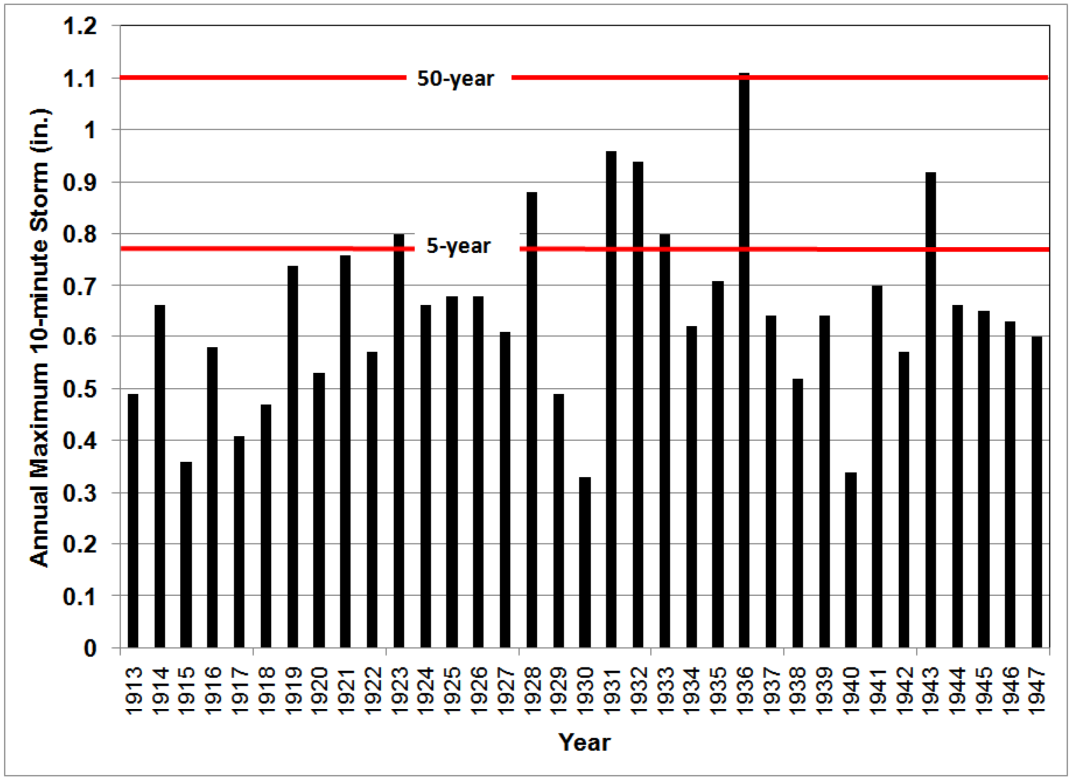
T = Return Period

Example (12.2.1)

- Given annual maxima for 10-minute storms
- Find 5- & 50-year return period 10-minute storms

$$\bar{x} = 0.649\,in \qquad s = 0.177\,in$$

$$y_5 = -\ln\left[\ln\left(\frac{T}{T-1}\right)\right] = -\ln\left[\ln\left(\frac{5}{5-1}\right)\right] = 1.5$$



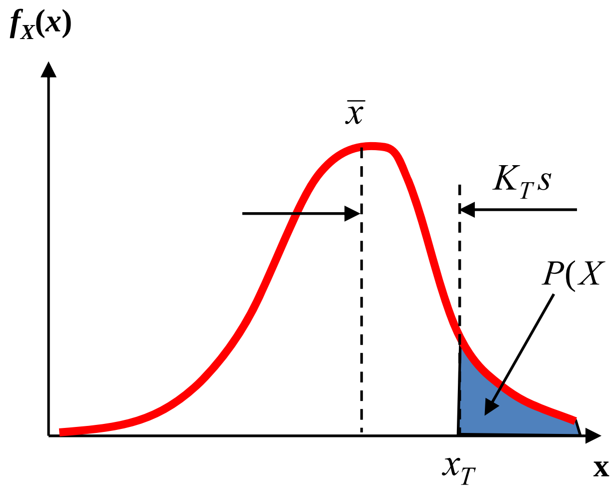
$$\alpha = \frac{\sqrt{6}s}{\pi} = \frac{\sqrt{6} * 0.177}{\pi} = 0.138$$

$$u = \bar{x} - 0.5772\alpha = 0.649 - 0.5772 * 0.138 = 0.569$$

$$x_5 = u + \alpha y_5 = 0.569 + 0.138 * 1.5 = 0.78\,in \qquad x_{50} = 1.11\,in$$

Frequency Factors

- Previous example only works if 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: $$x_T = \bar{x} + K_T s$$
- where x_T = Estimated event magnitude
  K_T = Frequency factor
  T = Return period
  $\bar{x}$ = Sample mean
  s = Sample standard deviation



Normal Distribution

- Normal distribution

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

$$K_T = \frac{x_T - \bar{x}}{s} = z_T$$

- So the frequency factor for the Normal Distribution is the standard normal variate

$$x_T = \bar{x} + K_T s = \bar{x} + z_T s$$

- Example: 50 year return period

$$T = 50;\ p = \frac{1}{50} = 0.02;\ K_{50} = z_{50} = 2.054$$

Look in Table 11.2.1

EV-I (Gumbell) Distribution

$$x_T = u + \alpha y_T$$

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

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

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

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

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

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

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

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

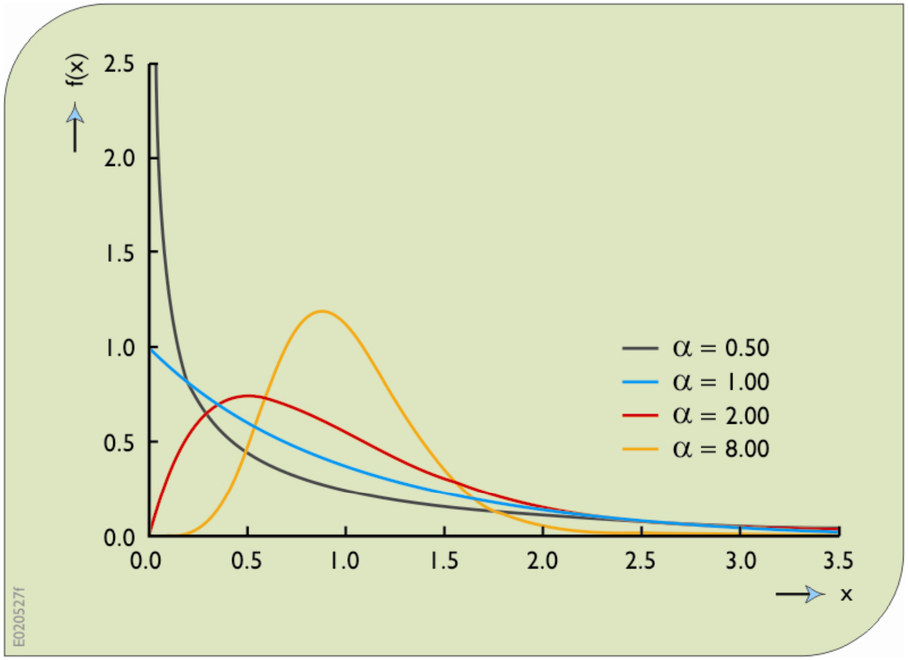
Gamma Distribution

- Gamma Distribution – distribution of sum of n IID exponential variables

$$f_X(x) = \frac{\lambda^\beta x^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)}$$

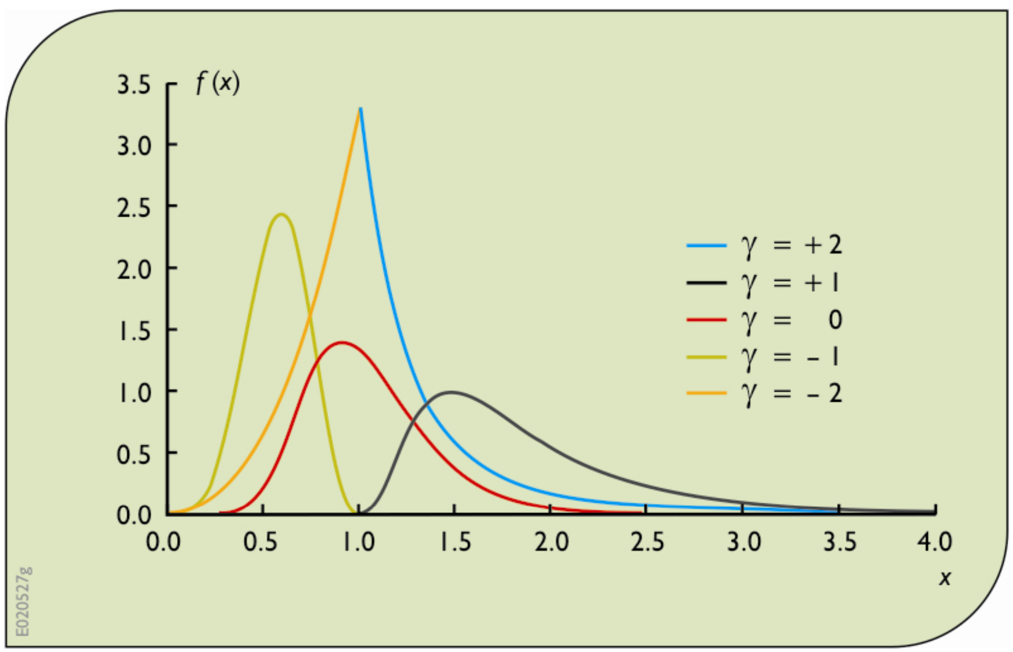
- Used to model many natural phenomena including streamflow
- 2 Special Cases:
  - Exponential $\beta = 1$
  - Chi – squared

$$\lambda = \frac{1}{2};\ 2\beta \text{ is an integer}$$



Generalizations of Gamma Distribution

- Consider a random variable *x*. Subtract a constant e from x
- If $(x - e)$ has a Gamma distribution,
  - then x has a ***Pearson Type III*** Distribution
    - (3 parameter Gamma distribution)
- If $\ln(x - e)$ has a Gamma distribution
  - then x has a ***Log Pearson Type III*** Distribution



Log Pearson Type III Distribution

- 1967 – Bulletin 15 – “A Uniform Technique for Determining Flood Flow Frequencies”
  - US Water Resources Council recommended LP-III as the “standard” flood frequency distribution for all US government agencies
- 1976 – Bulletin 17
  - Extended Bulleting 15 by recommending a “regional skew” parameter for the LP-III distribution.
  - Bulletin 17 describes method for computing flood frequency curves using annual flood series with at least 10 years of data.
- 1892 - Bulletin 17-B
  - Included methods for incorporating regional skewness into the calculations

Log Pearson Type III Distribution

- Event magnitudes are calculated as

$$y_T = \bar{y} + K_T s_y$$

$\bar{y} =$ mean of logs of x (ln *or* $\log_{10}$)

$s_y =$ standard deviation of logs of x

$$x_T = e^{y_T}; \text{ if natural logs are used}$$

$$x_T = 10^{y_T}; \text{ if base}-10 \text{ logs are used}$$

- K_T = frequency factor (quantiles with $p = 1/T$) of an LP-III distribution with skewness coefficient C_s
  - See table 12.3.1

LP-III Example

- Find 50 year return period annual maximum discharges on Guadalupe R. at Victoria, TX using LN and LP-III

Average	4.288369
St. Dev.	0.448573
Skew	0.308895

From the table (12.3.1) with C_s = 0.3

$$K_T = 2.211$$

$$y_T = \bar{y} + K_T s_y$$
$$= 4.28837 + 2.211 * 0.44857$$
$$= 5.28016$$

$$x_T = 10^{y_T}$$
$$= 10^{5.28816}$$
$$= 190{,}616\ cfs$$

Flood of October 1998, logQ= 5.66838592

$$5.6683859 = 4.28837 + K_T * 0.44857$$
$$K_T = \frac{5.6683859 - 4.28837}{0.44857}$$
$$K_T = 3.076$$

Corresponds to a return period of 300 years (see Appendix 3 of Bulletin 17-B)