Hydrologic Frequency Analysis

Return Period and Probability

_T-year:_ if time is very large the average time between events is T-years. The expected number of occurrences of a T-year event in N years is \( N/T \). These are only statistical expectations and are not certainties.

_Recurrence Interval:_ time between occurrences. The average value of recurrence interval is T. Average time of a T-year event recurrence is T years : Probability of a T-year \( (P_T) \) in any year is \( 1/T \).

\[
P_T = 1/T
\]

Assumptions:
- Flows are independent from year to year. May not be true since weather pattern is cyclic in most cases!
- Statistical properties of the system are not changing, i.e., no changes in watershed as an operator. This is also not totally true, but is adjusted with new data about the system.

Risk Analysis

Probability of \( k \) occurrences of \( Q^*_T \) \( (Q^*_T > Q_T) \) in a year is governed by the binomial distribution:

\[
f(k; P_T, n) = \frac{n!}{(n-k)!k!} P_T^k (1-P_T)^{n-k}
\]

where \( n! = n(n-1)(n-2)\ldots1 \); and \( 0! = 1 \)

**Example:**
Probability of 2 occurrences of a 20-year event in 30 years =

\[
P_T = \frac{1}{20} = 0.05
\]

\[
f(2,0.05,30) = \frac{30!}{28!2!} 0.05^2 0.95^{28} = 0.26
\]

Note: All references to equations and tables are from the text book "Design Hydrology and Sedimentology"
This probability is independent of the event $Q_T$.
This means that 26% of large 30-year records will contain exactly 2 peaks exceeding $Q_T$.

**Probability of not exceeding + probability at least one exceedence = 1**

\[
f(P_{T,n}) = 1 - (1 - T)^n
\]

where $f(P_{T,n})$ is the probability of $T$-year will be exceeded at least once in an $n$-year, if $n=T,
\text{then } f(P_{T,T})=0.632$

\[\therefore \text{a structure built on a } T\text{-year, probability that flow will be exceeded is } 0.63 \text{ at least once.}
\]

If you want to be 90% sure not to exceed a design capacity of a 25-year period $f(P_{T,25})$
would be $1 - 0.9 = 0.1$

\[\therefore 0.1 = 1 - (1 - T)^{25} \therefore T = 238 \text{ years}
\]

What does failure mean?
Flow exceeded! Danger! Human life or inconvenience? Consequences
Return periods using economic analysis.
Danger: interest rate sensitivity

See Figure 2.2 in the text book!

**Frequency determination**
Approach depends on quantity, quality, and type of hydrologic data.

**Long flow record at the site (the only one considered)**
Any statistics involve these terms:
- Population
- Sample

**Population statistics:**
- Mean $\mu_x$
- Standard deviation $\sigma_x$, $\sigma^2_x = \text{variance}$
- $C_v$ coefficient of variation $\equiv \text{dimensionless}$
- Skewness $\gamma = \text{measure of symmetry, for normal distribution } \gamma = 0$

**Sample statistics:**
- $\bar{x}$
- $S_x$
- $\hat{C}_v$
- $C_s$

\[
\bar{x} = \frac{\sum x_i}{n}
\]

\[
S_x = \sqrt{\frac{\sum x_i^2 - n \bar{x}^2}{n-1}}
\]

Note: All references to equations and tables are from the text book "Design Hydrology and
Sedimentology"
\[ \hat{C}_v = \frac{s_x}{x} \]
\[ \hat{C}_s = \frac{n \sum (x_i - \bar{x})^3}{(n-1)(n-2)s_x^3} \]
n = sample size

If we assume that the set sample is independent and represents the population with no trend \( \therefore \) we can have frequency analysis

**Probability plotting**
Look at data in Table 2.1 in the textbook. 6 events out of 31 exceeded 2120 cfs
\( \therefore \) 20% probability of exceeding 2120
\( \therefore \) 5-year flood is 2120

This intuitive approach has a limit though. Systematic approach is needed:

1. Rank data from largest to smallest
2. Calculate plotting position \( P = \frac{m}{n+1} \), where \( n= \) total number of records and \( m= \) rank
3. Plot the observation on a log normal or probability paper with \( P \) along the probability scale and the magnitude along the variable scale
   \( P \) represents fraction of data greater or equal to the corresponding value. See the curvature with or without probability paper! Fit a line to the data.

**Probability distributions**
Probability density function (pdf) \( P_x(x) \) \hspace{1cm} vs \hspace{1cm} Cumulative density function (cdf) \( P_X(x) \)

Used for probability of random event

\[ P_x(x) = \int_{-\infty}^{x} P_x(x) dx \]

Bell shape normal pdf

\[ \text{Prob (a} \leq x \leq b) = P_X(b) - P_X(a) \]

Mean and \( \sigma x \) on normal distribution

Normal distribution:

\[ P_x(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[ -\frac{(x - \mu)^2}{2(\sigma x)^2} \right] \]

\[ \text{Prob } X < x \]

Note: All references to equations and tables are from the text book "Design Hydrology and Sedimentology"
Log normal distribution:

\[ P_x(x) = \frac{1}{x \sigma \sqrt{2\pi}} \exp \left[ -\frac{(\ln x - \mu_Y)^2}{2(\sigma_Y)^2} \right] \]

\( \mu_Y \) and \( \sigma_Y \) are mean and s.d. of \( \ln X \)

Extreme value type 1:

\[ P_x(x) = \frac{1}{\alpha} \exp \left[ \frac{-(x - \beta)}{\alpha} \exp \left( \frac{-x - \beta}{\alpha} \right) \right] \]

\[ \alpha, \beta \]

\[ \hat{\alpha} = \frac{S_x \sqrt{6}}{\pi}, \hat{\beta} = 0.45 S_x \]

Pearson type III:

\[ P_x(x) = P_0 e^{-(x - \alpha)/\delta} (x/\alpha)^{\alpha/\delta} \]

with the mode at \( X=\alpha \) and lower bound at \( X=0 \), \( \delta = \text{difference between mean and mode} \)

\( P_0 = P_X(\alpha) \).

Log Pearson : take log of events and apply Pearson type III
Mode at \( x=\alpha \)
\( \delta = \text{difference in the mode and mean} \)

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0</td>
</tr>
<tr>
<td>Lognormal</td>
<td>3C_v + C_v^3</td>
</tr>
<tr>
<td>Extreme value 1</td>
<td>1.139</td>
</tr>
<tr>
<td>Log Pearson III</td>
<td>Any value</td>
</tr>
</tbody>
</table>

**Standardized variable:**

\[ X_T = \bar{X} (1 + C_v K_T) \]

\( X_T \) = magnitude of \( X \) at return period \( T \)

\( \bar{X} = \text{mean}, C_v = \text{coef of variation} \)

\( K_T = \text{frequency factor dependent on distribution} \)

**For normal distribution**

\[ \text{Prob} (Q <= 2500) = \text{Prob} \left( Z <= \frac{2500 - 1599}{1006} \right) \]

\[ \cong \text{Prob} (Z <= 0.896) \]

Note: All references to equations and tables are from the text book "Design Hydrology and Sedimentology"
Prob (Q <= x) = Prob \left( Z <= \frac{(x - \mu x)}{\sigma x} \right)

These are in tables to simplify calculations

Example Frequency Analysis:
Table 2.1 data
\hat{\mu} = 1599
\hat{\sigma} = 1006
C_v = 0.629

Log Normal: use
\begin{align*}
X_T &= \bar{X} (1 + C_v K_T) \\
K_T: & \text{ Table 2.5}
\end{align*}

Extreme Value type I
Use X_T
K_T: Table 2.6

Log Pearson III
K_T: Table 2.7
Use procedure for X_T in book

Steps for LP3 (Probability Plotting):
1. Transform observations to logarithm
   \begin{align*}
   Y_i &= \log X_i \\
   \end{align*}
2. Compute mean logarithm, \bar{Y}
3. Compute standard deviation of the logarithm S_Y
4. Compute \hat{C}_v: skewness
5. Compute \begin{align*}
Y_T &= \bar{Y} + S_Y k_T \\
X_T &= \text{antilog} Y_T
\end{align*}

Please refer to chapter 2 of the text for further details on the subject.

Note: All references to equations and tables are from the text book "Design Hydrology and Sedimentology"