Exercise week 14 – Analysis of discharge data for hydrologic design

Goal

Perform the statistical analysis of discharge data using the Log-Pearson Type III distribution (or simply LP3).

1 On Paper

1. **Task:** Write the general equation to estimate the annual maximal discharge $Q_{est}(T)$ that corresponds to a return period T according to the frequency factor approach. Then, write the frequency factor equation for the LP3 approach. (see Lecture 8, slides ≈ 30–43).

2. **Question:** Which terms of the equation depend on the return period? Which depend(s) on the observed discharge time series?

3. **Task:** Estimate the discharge corresponding to a return period T=50 years, knowing that: for the log-transformed discharge (y), the mean and standard-deviation are $\bar{y} = 0.8$ and $S_y = 0.1$ respectively, an that the skewness $G_s(y)$=-0.4. Use tables 10.4.1abcd at slide 40–42 to obtain the frequency factor $K_T(T, Gs)$. The units of the parameters correspond to m^3/s so your estimated discharge will also be in m^3/s.

4. **Task:** Write a pseudo-code on paper to compute the estimated maximal discharges using the LP3 approach, for an arbitrary number of return periods of interest.

2 Numerical application

You will explore a dataset with mean daily discharge data (m^3/s) since 1971 at the BAFU/OFEV station Orbe – Le Chenit, Frontière. You can find all the official information on this station at https://www.hydrodaten.admin.ch/fr/2371.html. The data were provided by BAFU/OFEV through the datafile `2371_Abfluss_Tagesmittel_[...].csv`.

Suppose the municipality of Le Chenit (VD) wants to build a small bridge over the Orbe river, right after the French border. The height of the future bridge should be such that the bridge's road will only be flooded for discharges larger than 13.5 m^3/s. You need to determine the approximate frequency at which the road will be flooded starting from a statistical analysis of available discharge data and the LP3 distribution.

1. **Task:** Check how the datafile `2371_Abfluss_Tagesmittel_[...].csv` is formatted. Then, open the script `compute_yearly_maxima.m` and run it. It will:
   - import the data;
   - plot the full dataset as well as a one selected year;
   - export the yearly maxima to a file: `annual_maxima_LeChenit.csv`.

2. **Task:** Now start a new script and import the newly-generated yearly maxima from file: `annual_maxima_LeChenit.csv`.

3. **Task:** Implement the pseudo-code you have written on paper. It is very convenient that you create a function `computeKTLP3` that takes the variables G_s and T as input, implements equation 1 (see Appendix) and returns the frequency factor K_T for the LP3 distribution. To compute the variable z, you can check the matlab function `icdf` that numerically inverts a cumulative distribution.

4. **Task:** Test your function for a few values of T and G_s and check that they are similar to those reported in the tables at slides 40–41.

5. **Task:** Run your code over a dense range of return periods so you can obtain a smooth function $Q_{est}(T)$, and then plot it. You can also plot the measured discharge Q against the empirical return period as computed from the Weibull plotting position formula (just like you did in assignment 1 for precipitation data).

6. **Question:** What is the probability that the bridge will be flooded during the 5-year mandate of the mayor (*maire* in french)?

7. **Optional Task:** Use the Gumbel distribution, instead of the LP3 distribution, for the estimation of return periods and use it to provide another estimate of the flooding probability of the bridge during 5 years.

Appendix

As the LP3 distribution is not invertible, it is not possible to derive an analytic equation that relates the discharge Q to the return period T. For this reason, we need to use the frequency factor approach.

A good estimate of the frequency factor K_T for the LP3 distribution is given by:

$$K_T(T, Gs) \approx z + (z^2 - 1)\frac{G_s}{6} + \frac{1}{3}(z^3 - 6z)\left(\frac{G_s}{6}\right)^2 - (z^2 - 1)\left(\frac{G_s}{6}\right)^3 + z\left(\frac{G_s}{6}\right)^4 - \frac{1}{3}\left(\frac{G_s}{6}\right)^5 \quad (1)$$

where:

- G_s is the skewness of the log-transformed discharge variable $y = log_{10}(Q)$.
- z is the inverse cumulative standard normal distribution, evaluated in $P = 1 - 1/T$ (see also Lecture 8, slides $\approx$ 10–12).