RETURN PERIOD ESTIMATION USING PROBABILITY PAPERS EXERCICE

1. Plotting position formulas

These formulas allow the calculation of the empirical (i.e estimated from the raw data) exceedence probability of observed data points. This method involves ordering the data from the largest event to the smallest event, assigning a rank of 1 to the largest event and a rank of *n* to the smallest event, and using the rank (i) of the event to obtain a probability plotting position. Numerous plotting-position formulas are available.

$$\text{Weibull,} \qquad P_i = \frac{i}{n+1}; \qquad (5.5b)$$

$$\text{Hazen,} \qquad P_i = \frac{2i-1}{2n} = \frac{i-0.5}{n}; \qquad (5.5c)$$

$$\text{Cunnane,} \qquad P_i = \frac{i-0.4}{n+0.2}. \qquad (5.5d)$$

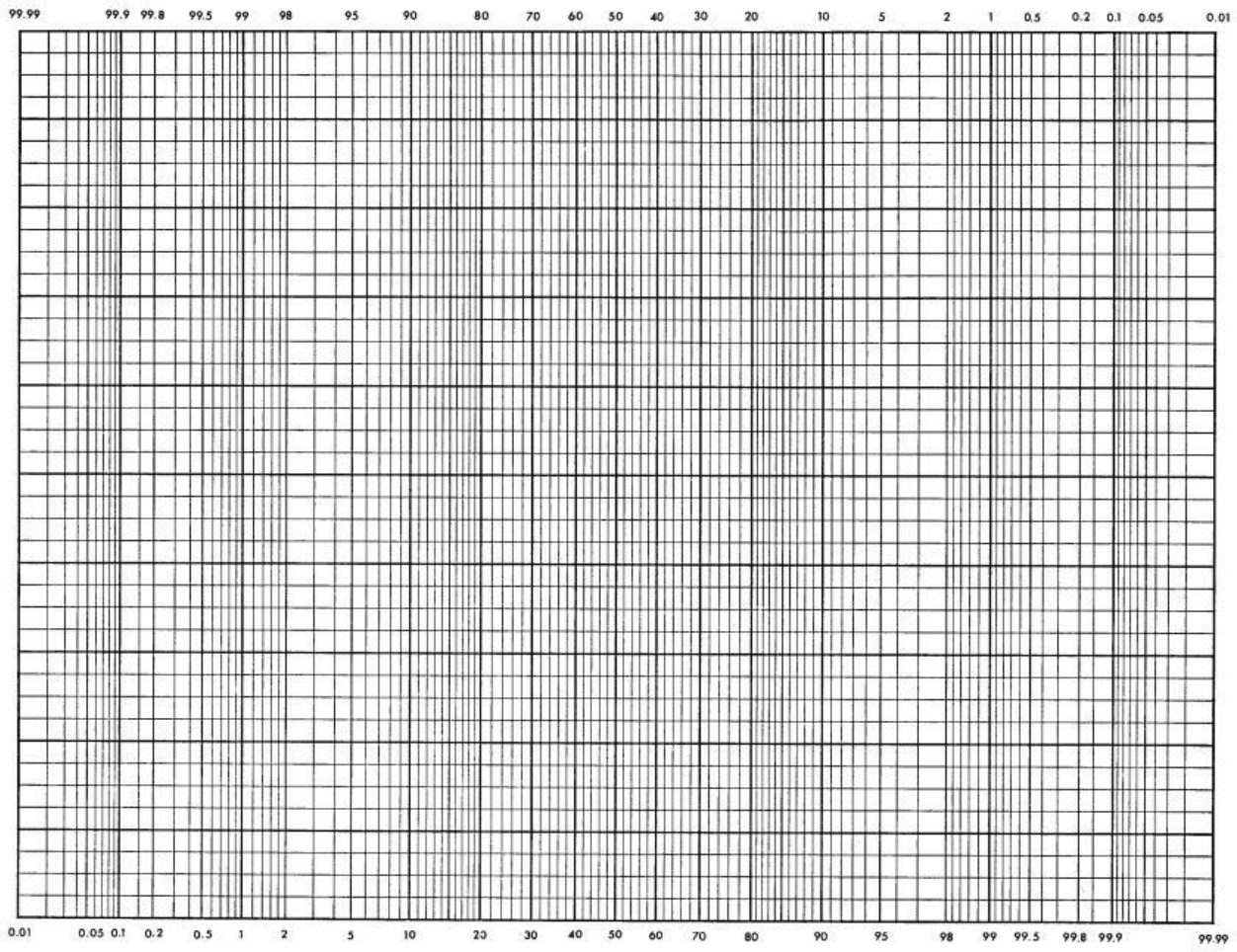
Example: application to the following series of discharges in cfs: { 1000,465,455,290,157,100}

			Empirical exceedance probability		
RANK	**Q(cfs)**	**Ordered Q (cfs)**	**Weibull**	**Hazen**	**Cunnane**
1	455	1000	0.142857143	0.083333333	0.096774194
2	105	465	0.285714286	0.25	0.258064516
3	465	455	0.428571429	0.416666667	0.419354839
4	1000	290	0.571428571	0.583333333	0.580645161
5	290	157	0.714285714	0.75	0.741935484
6	157	100	0.857142857	0.916666667	0.903225806

2. Probability papers

A probability paper is a special type of graph paper in with the ordinate being used to plot the value of the random variable, and the probability of its occurrence is given on the abscissa. The probability scale will vary depending on the probability distribution being used (i.e each distribution will have its own probability paper: **never use excel to plot the graph as excel does not have the right graduations**).

The example below is the normal probability paper.



The scale at the top of the graph is the *exceedence probability*, which is the probability that the random variable will be equaled or exceeded in one time period (it changes from 99.99% to 0.01%). The lower scale is the *nonexceedence probability*, which is the probability that the corresponding value of the random variable will not be exceeded in any one time period. This scale extends from 0.01% to 99.99%. The ordinate of probability paper is used for the random variable, such as the peak discharge.

3. Using probability papers

1. If you think you data set is normal, use a normal probability paper; if you think your data is log-normal, use a lognormal probability paper.
2. Add graduations to the Y axis so that the data ranges covers all your data points and all the values for which you may want to estimate the exceedance probability
3. Calculate the empirical exceedance probability of the data points by
    a. Sort the data in descending order
    b. Using the Weibull plotting position formula, find the empirical exceedance probability

4. Plot the data on the probability paper
5. If the data points are in straight line on the probability paper, then you have selected the right probability paper (hence the right distribution). If the data does not follow a straight line, then you can conclude that the data does not follow the distribution corresponding to the probability paper.

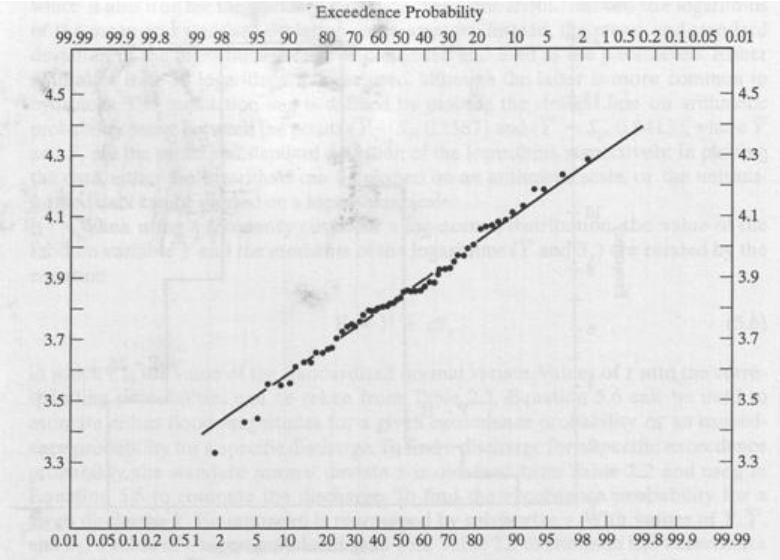


Figure 5.7 Frequency curve for the logarithms of the annual maximum discharge.

6. Plot the best straight line in the data and use that line to relate the discharges (Y –axis) to exceedance probability (X-axis).

APPLICATION

Data sets A and B represent annual floods on two rivers. They are already sorted. Using the probability papers provided in class, find out which data set follows a normal distribution, and which one follows a lognormal distribution. Find the 100-year event for each data set.

DIRECTIONS

At the beginning of the class, you will be provided four probability papers

- A normal probability paper for data set A
- A normal probability paper for data set B
- A lognormal probability paper for data set A
- A lognormal probability paper for data set B

1) Use excel to calculate the weibull plotting position formula for data sets A and B
2) Plot the data sets on the four probability papers
3) Based on whether the data points follow a straight line or not, find which data set is normal and which one is lognormal
4) Find the 100-Year flood for each data set.