MATH2019 – Statistics for Civil and Environmental Engineering

Worksheet 4 – Probability models

1. The data in the file `quake.dat` represent the time in days between successive serious earthquakes worldwide, between 16th December 1902 and 4th March 1977.

   An exponential distribution is proposed as a model for the time between earthquakes.

   Using the mean of the sample to estimate the mean of the exponential distribution, calculate directly (without using MINITAB) the probability that the gap between successive earthquakes is

   (i) less than a week

   (ii) more than a year

   Check your answers using MINITAB.

   Use a probability plot to check whether the model seems appropriate.

2. The data in the file `stress2.dat` are the measurements of stress resulting from wave action on the joints of an off-shore structure.

   Investigate whether any of the normal, lognormal, Weibull or exponential distributions are plausible models for this variable.

   Use your preferred model to estimate the probability that the stress is between 50 and 100.

3. The file `susque.dat` has one column of data, which contains the annual peak flows of the River Susquehanna in Pennsylvania over a period of 24 years.

   Verify that the Extreme Value Distribution for Greatest Values (EVG1) is a good fit to these data, and estimate the parameters of the distribution.

   Suppose that you are designing a structure on this river which needs to be able to withstand a specified flow. If the required design life is 40 years, calculate the flow which the structure must be able to withstand if the risk of failure is to be

   (i) 5%

   (ii) 10%

   [First calculate the annual probability of exceedance, and then use the distribution function of the EVG1 distribution to calculate the threshold flow value.]

   Compare your answers with those you get using the Fundamental Formula for Flood Control.