

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Civil and Environmental Engineering

1.017 Computing and Data Analysis for Environmental Applications

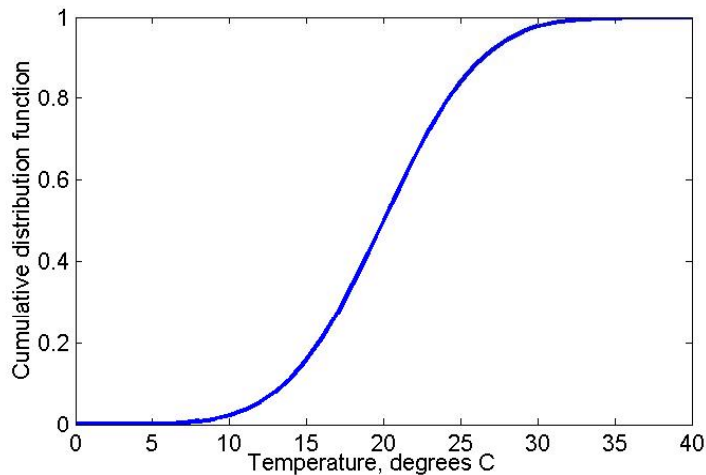
Quiz 1

Thursday, October 8, 2002

Please answer all questions on a separate piece(s) of paper with your name clearly identified:

Problem 1 (15 points)

Suppose that the annual maximum temperature at a particular location has a normal distribution with a population mean of 20 degrees C and a population standard deviation of 5 degrees C. The corresponding CDF is plotted below. Compute (approximately) the probability that the maximum temperature in a given year will be in the range between 15 and 25 degrees C. Would this probability increase or decrease if the population standard deviation decreased to 3 degrees C?



Problem 2 (30 points)

Suppose that the algae growth rate (g) in a lake is related to nitrate concentration (N) by the following equation:

$$g = \frac{2N}{5 + N}$$

Also, suppose that N is exponentially distributed with a distributional parameter $a = 5$. Write a MATLAB program (a script or a function) to compute the probability that g is greater than 1. You should do this by applying the relative frequency approach. That is, generate a set of many replicates of g from many exponentially distributed replicates of N and compute the required probability by dividing the number of replicates yielding g values greater than 1 by the total

number of replicates. Independent exponentially distributed random variables can be generated with the MATLAB `exprnd` function, which has the following form:

```
array_of_random_numbers = exprnd(a,m,n)
```

where `a` = distributional parameter (use `a = 5`), `m` = number of rows in `array_of_random_numbers` and `n` = number of columns in `array_of_random_numbers`.

Be sure to use element-wise MATLAB operations when you want to multiply or divide corresponding individual elements of two arrays. Please write your code neatly and precisely so we can enter it into MATLAB to see if it works.

Problem 3 (20 points)

Characterize the set of 12 monthly total stream flow values given below by plotting or computing the following descriptive statistics:

Histogram
Sample cumulative distribution function (CDF)
Sample mean
Sample standard deviation
Sample median

4.2	0.4	6.0	21.0
1.0	9.8	17.	0.3
0.4	12.4	16.	15.

Use the sample CDF to estimate the probability of obtaining a monthly flow less than or equal to 10. **Do this problem by hand, do not write a MATLAB code.**

Problem 4 (10 points)

Explain the difference between the mean of a set of data (the sample mean) and the mean of a random variable described by a probability distribution (the population mean).

Problem 5 (25 points)

Write a MATLAB program (script or function) that predicts the concentration of a solute in a well-mixed holding basin described by the following equation:

$$VC(t+1) = 0.8 VC(t) + QC_{in}(t) - QC(t) \quad ; \quad t_0 \leq t \leq T$$

where V is the (constant) volume of the basin (m^3), $C(t)$ is the concentration (kg/m^3) in the basin at time t , Q is the (constant) inflow rate (m^3/sec) through the basin, $C_{in}(t)$ is the (time-dependent) inflow concentration (kg/m^3), and t_0 and T are the beginning and end times of the simulation period, respectively. Assume that V , Q , $C_{in}(t)$, $C(t_0)$, t_0 , and T are specified inputs. Your program should provide a plot of $C(t)$ vs t . Please write your code neatly and precisely so we can enter it into MATLAB to see if it works.