Problem set 2

Due March 17, 2026

1. The letters A, A, L, M, M, M, S were randomly set in a series. What is the probability that the word "MAMMALS" is obtained?

2. The following is the probability distribution function, $f(R_g)$, of the radius of gyration (R_g; this quantity defines the average dimension of an object) of a 20-residue peptide simulated by Molecular Dynamics.

R_g [Å]	$f(R_g)$
6.6	0.002
6.8	0.012
7.0	0.065
7.2	0.142
7.4	0.205
7.6	0.192
7.8	0.146
8.0	0.101
8.2	0.063
8.4	0.038
8.6	0.017
8.8	0.010
9.0	0.004
9.2	0.002
9.4	0.000
9.6	0.001

   (a) Check if the distribution is normalized to 1 (total probability) and correct the values if necessary (note that the bin size is 0.2 Å).

   (b) Make a bar plot (histogram) of the distribution and a plot of the respective cumulative distribution function ($F(R_g)$). The two plots can be superposed on one graph.

   (c) Determine the most probable value of R_g and the average value of R_g. Comment the result.

   (d) Compute the variance, skewness, and kurtosis and draw conclusions about the character of the distribution.

Hint: When computing the average and central moments keep in mind that the value of the distribution assigned to a given R_g value is an average over the $[R_g, R_g + \Delta R_g]$ interval.

3. Collect raw data of a single quantity of your like, either acquired on your own or from the literature, and convert them into probability density distribution function and cumulative probability density distribution function. Choose a reasonable sample size, a minimum of 10. Examples are temperature readings over 10-14 days at a certain hour of a day. atmospheric CO_2 content, energy or other quantity from an MD run, etc. If your raw-data file is small, you can list the data in the solution, otherwise send it as an ASCII file.