PX 4128 – Data Analysis

Continual Assessment Number 3: 20% of final grade
Deadline for submission: 2 pm on Friday 23rd Nov 2017

Please use off-the-shelf random number generators, functions and plotting routines, unless specifically asked to construct your own. **For this piece of assessment, you must write your own fitting algorithms. Do not use pre-existing fitting algorithms.** Show all working, coding, and produce a succinct and tidy report. Full marks can only be awarded where there is proof that the student understands the statistics that they are being asked to explore. Code and spreadsheets count as working, and **must be included** in the report to obtain the relevant marks. For submission you must use the Turn-It-In post via Learning Central. Note that you **can only post one document to Turn-It-In**, so everything will need to be in a single PDF. However, you make multiple submissions of your single document. Your marks here will depend on how well you demonstrate the fidelity of your answers. Please also take note of the University's policy on plagiarism, which is outlined in your student handbook.

1) Fit a straight line to the data in the Learning Central assessment folder (PX4128_CA3_Q1.dat) quoting all the fit parameters and their errors. Comment on the quality of the fit and the uncertainties.

[35 marks]

2) When it comes to nonlinear fits, the errors in the fit parameters can be difficult to estimate. As such we often resort to Monte Carlo techniques. In the Learning Central assessment folder you will find the file PX4128_CA3_Q2.dat. Once again, perform a straight line fit, calculating, and reporting all the errors. Next, randomly sample the y values, using the given values and their errors to define the normal from which to sample. Recalculate the slope and intersect from the new (randomly sampled) data set. Repeat this many different times, using the results of each fit to work out the mean and standard deviation for the fit parameters.

[40 marks]

3) Maximum likelihood can also be used in cases where the likelihood function is not a Gaussian. For example, suppose that we are measuring the weak spectrum from a gas in the lab. The emission is so weak that we are only detecting 10s of photons in each recorded frequency bin, and so the bins are well described by a Poisson distribution. We are trying to fit a straight line to the spectrum. Derive the normal equations for the slope and intersect of the line in this case.

[25 marks]