UCF Physics: AST 5765/4762: (Advanced) Astronomical Data Analysis

Fall 2019 Homework 5

Due Tuesday 1 October 2019

Work:
Become sufficiently familiar with probability and error analysis to:

1. Understand all the terms used as headings or italicized in the reading assignments.
2. Calculate the mean, median, mode, range, variance, and standard deviation of a group of measurements.
3. Understand the Gaussian and Poisson distributions, their parameters, and their application in astronomical observation.
4. Memorize the formula for the Gaussian distribution. (On quiz!)

Resources:

1. Chapter 6 of Bevington (**DUE** before class Tuesday, 1 October 2019)
2. Chapters 1 and 2 of Howell (**DUE** before class Tuesday, 8 October 2019)
3. **AST 5765 only:** Read `https://jwst-docs.stsci.edu/display/JTI/NIRISS+Overview` and any one (or more) of the linked pages on its four observing modes (**DUE** before class Tuesday, 8 October 2019). Ambitious students may also read
   `https://jwst-docs.stsci.edu/display/JTI/Near+Infrared+Spectrograph%2C+NIRSpec`,
   `https://jwst-docs.stsci.edu/display/JTI/NIRSpec+Overview`,
   `https://jwst-docs.stsci.edu/display/JTI/NIRSpec+Optics`.
4. **AST 5765 only:** Sections 14.1 – 14.3, 14.8 of Press (**Recommend** you read this before class Tuesday, 8 October 2019)

Hand in:

1. (10 points) Create a 400 element sample with 396 draws from a Poisson distribution for 10,000 photons and 4 draws from the uniform distribution between 0 and 10^6. This sample represents data from a CCD with just 1% bad pixels. Print the mean and median of the sample and also put the result in your homework file in comments. Which is closer to N?
2. (10 points) Calculate the standard deviation, σ, for the sample above. Create a subsample including only points within 5σ of the median (remember your boolean array slicing!). Print the new sample's mean, median, and standard deviation. What has happened?

3. (10 points) Repeat the steps in the previous problem **once** on the subsample, to make a subsubsample. Print the new mean, median, and standard deviation. How different are the final mean and median? How close is the final standard deviation to that expected for the Poisson distribution? Will this method always remove every bad pixel?

4. (10 points) Make a Python routine called `sigrej()` that carries out the "sigma rejection" process given above. It should accept an array containg the data set, a tuple of rejection limits giving the number of standard deviations for each iteration (it would be (5., 5.) for the example above), and an optional Boolean mask, the same shape as the data, indicating which items are good (False=bad, True=good). Return the modified mask. Do not use Numpy's "masked array" feature (numpy.ma). The routine should only use the data-set points indicated as good in the mask, and it should modify the mask by flagging new bad pixels as False. Run it on your data set from question 1. Print the mean of the "cleaned" data set and compare to your calculation from the prior step (they should be the same).

5. (10 points) Use the error propagation equation to derive an expression for σ_f if $f = \sin(x)\mathrm{e}^{-y/1000}$.

   The following arithmetic problems must be handed in electronically, not on paper, as with all other homework in this class. You may typeset them or do them longhand and include a photo or scan. Please express the final answers as decimal numbers.

6. (2 points) What are
$$\begin{array}{r} 9 \pm 3 \\ +20 \pm 4 \\ \hline \end{array}$$

7. (2 points) What are
$$\begin{array}{r} 9 \pm 3 \\ -20 \pm 4 \\ \hline \end{array}$$

8. (2 points) What are
$$\begin{array}{r} 9 \pm 3 \\ \times 20 \pm 4 \\ \hline \end{array}$$

9. (2 points) What are
$$\begin{array}{r} 9 \pm 3 \\ \div\ 20 \pm 4 \\ \hline \end{array}$$

10. (10 points) Include a copy of your class log file in your handin. Print the Git log for your main homework file.

11. **AST 5765 only:** (10 points) Write a Python function to calculate f. Have your function accept an array containing two identically-shaped subarrays, like the output of the `numpy.indices()` function. The first array contains y values and the second contains the corresponding x values. Back in your main homework file, for the closed ranges $x = [0, 30]$

and $y = [0, 3000]$, make a 2D, linearly gray-scaled, PNG image of the value of this function over this space. Give it more than sufficient resolution to see the variation, but keep it a convenient size for viewing on-screen. Be sure the origin is in the lower-left corner.