ASTR 5550: HW4

Jasmine Kobayashi

```
# Libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import random
import os,sys

# import helper script file
#(change working directory to two folders up)
src_dir = os.path.abspath("")
hw_dir = os.path.dirname(src_dir)
os.chdir(hw_dir)

# import my own code
import hw_helper_func2 as hf    # this is my own code I made (for probability/distribution fu
```

(**JK note:** To view the code with the functions I made myself to (hopefully) help with all assignments click here)

1. Combining Poisson Distributions

Given two Poisson distributions:

$$P(x, \mu_A) = \frac{\mu_A^x}{x!} e^{-\mu_A} \text{ and } P(x, \mu_B) = \frac{\mu_B^x}{x!} e^{-\mu_B}$$

Show that they combine to a Poisson distribution:

$$P(x, \mu_C) \text{ where } \mu_C = \mu_A + \mu_B$$

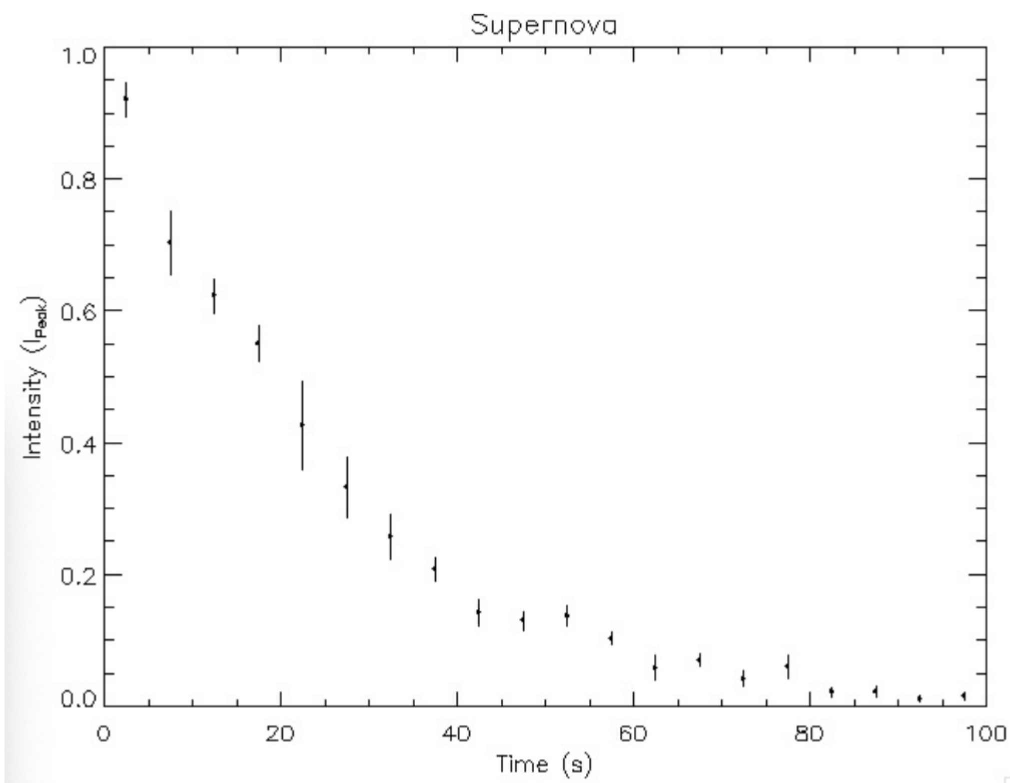
Hint: For any given integer x, the one must sum all possibilities of $P(i, \mu_A)P(x - i, \mu_B)$.

2. Supernova Light Curve

After a supernova reaches its maximum brightness, the light curve exponentially decays as do the radioactive materials. The decay time can tell us its type. Examine the light curve below.

$I = [0.921, 0.704, 0.623, 0.550, 0.426, 0.332, 0.258, 0.208, 0.143, 0.130, 0.137, 0.103, 0.058, 0.070, 0.042, 0.060, 0.022,$

$\sigma = [0.026, 0.048, 0.026, 0.027, 0.068, 0.046, 0.034, 0.017, 0.020, 0.014, 0.015, 0.009, 0.019, 0.010, 0.012, 0.018, 0.007,$




```
I = [0.921, 0.704, 0.623, 0.550, 0.426, 0.332, 0.258, 0.208, 0.143, 0.130,
     0.137, 0.103, 0.058, 0.070, 0.042, 0.060, 0.022, 0.022, 0.011, 0.015]
sigma =  [0.026, 0.048, 0.026, 0.027, 0.068, 0.046, 0.034, 0.017, 0.020, 0.014,
          0.015, 0.009, 0.019, 0.010, 0.012, 0.018, 0.007, 0.008, 0.005, 0.005]
```

Part(a)

Assuming that σ represents a 1-sigma Gaussian uncertainty, find the most likely parameters under the hypothesis that the intensity undergoes an exponential decay:

$$I = I_0 e^{-t/\tau}$$

Here, τ is the decay time. As one can see, I_0 should be nearly unity but, for this problem, do not fix $I_0 = 1$. Calculate the uncertainty in τ. Plot the observations and the fit.

Hint: One way is to perform a linear fit to $ln(I)$. Be careful how you treat the uncertainty σ; Taylor expand $ln(I \pm \sigma)$ to calculate the uncertainties of $ln(I)$.

Part (b)

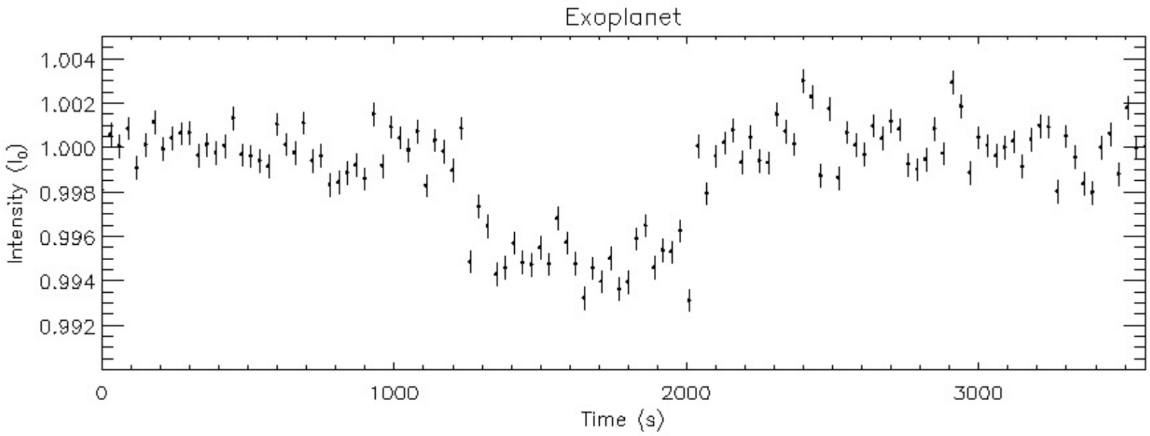
Calculate χ^2_ν and compare it to the expected PDF/CDF of χ^2_ν. Plot your results. Is the hypothesis justified? What is the probability for χ^2_ν to be above the calculated value?

3. Extra-Solar Planet

The Kepler mission used the transit method in which one examines a time series of a star's intensity for a negative excursion. Under this method, the parent distribution of a star's intensity can be well established. In this example, the star's intensity is measured at a 30 sec cadence and found to be $I_0 + 0.001I_0$ (1-sigma) with a Gaussian parent distribution.

Finding a transit often involves several steps. The first step is to identify intervals that may have a transiting planet. One way is to examine one hour (120-point) stretches (sliding every half an hour, 60 points) for a non-constant distribution.

Read in the text file, `HW4_data_A`, from Canvas. It contains 120 points of intensity in units of I_0, one every 30 seconds. Create a corresponding time array going from 0 to 3570 seconds. Assume the uncertainty in time is negligible.




```
df = pd.read_fwf("hw4/HW4_data.txt", sep=" ",header=None,)
df.head()
```

	0
0	0.998088
1	1.000580
2	1.000070
3	1.000850
4	0.999086

Part (a)

Start by eliminating the possibility that the negative excursion is a random fluctuation. Plot the PDF of the expected χ^2_ν under the hypothesis that the intensity is constant. Calculate χ^2_ν and compare to show that this event is **not** consistent with a constant intensity. What is the mean of the intensity (I_μ) and the uncertainty of the mean ($\sigma_{I\mu}$)? Is I_μ less than 1 by more than the $\sigma_{I\mu}$?

Hint: σ of the parent distribution is known ($0.001I_0$)

Part (b)

Now that the interval is identified as significant and negative, let's examine and fit the negative excursion. Keeping it simple, use a three-parameter (I_0, t_{start}, t_{end}) fit:

$$I = \begin{cases} I_0 - \Delta I & t_{start} \le t \le t_{end} \\ I_0 & \text{otherwise} \end{cases} \tag{1}$$

Do a least-squares fit with a method of choose. My method is to guess t_{start} and t_{end} then calculate χ^2_ν along with ΔI. Increment t_{start} and t_{end} and recalculate ΔI until χ^2_ν is minimum. Plot the data (with error bars if you can) and overplot your fit. What are I_0, t_{start}, and t_{end}?

Part (c)

Estimate the uncertainties of I_0, t_{start}, and t_{end}. Explain how you arrive at your values.

Hint: The uncertainty of ΔI is straight-forward. Recall that you can calculate σ_I, but $\partial t / \partial I$ can only be estimated. Can one have an uncertainty in time that is less than δt (30 seconds)?

4. Kolmogorov-Smirnov Test

Using a random number generator, create two distributions:

$$f_1(x) = P(x, \mu_1, n); \mu_1 = 8, n = 100$$

$$f_2(x) = P(x, \mu_2, n); \mu_2 = 5, n = 100$$

Part (a)

Calculated and plot the two CDFs for $n = 100$. Compare the two distributions using the Kolmogorov-Smirnov Test with $\alpha = 0.1$. The more exact formula for the threshold is:

$$D > \sqrt{-\frac{1}{2} \ln\left(\frac{\alpha}{2}\right)} \sqrt{\frac{n+m}{nm}}; n, m \text{ are number of points}$$

Part (b)

Repeat the test several (5 to 10) times recreating the distributions. Do f_1 and f_2 consistently pass or fail the test?

Part (c)

Repeat the test for higher n, say 1000 (for both f_1 and f_2) several times. Does the test at $n = 1000$ reveal that the two distributions are not from the same parent? What does this exercise tell us about the Kolmogorov-Smirnov Test?