ASTR 5550: HW3

Jasmine Kobayashi

```python
import numpy as np
import matplotlib.pyplot as plt
```

1. Joint Probability and Find the Gap

Suppose the goal is to take a series of images of galaxies which necessitates very long exposures in a narrow field-of-view (FOV). The expected (mean) number of galaxies in this FOV is 19.3.

Part (a)

How many exposures in separate FOVs are needed so that the probability of seeing at least 100 galaxies is greater than 50%? (Think simple!)

$\mu = 19.3$ (galaxies per FOV)

$\langle x \rangle = x\, P(x) \Rightarrow (19.3) = x\, P(x)$

$$P(x) = \frac{19.3}{x} \qquad \Rightarrow P(x) \equiv \text{Probability of } x \text{ galaxies per 1 FOV}$$

$$P_n(x) = \frac{19.3}{x}\, n \qquad \Rightarrow \begin{cases} n \equiv \# \text{ of FOV} \\ P_n(x) \equiv \text{Probability of } x \text{ galaxies per } n \text{ FOVs} \end{cases}$$

$$\Rightarrow 0.5 < P_n(x=100) = \frac{19.3}{100}\, n$$

$$0.5\, \frac{100}{19.3} < n$$

$$\Rightarrow \boxed{n > 2.59} \approx \boxed{\text{at least } \sim 3 \text{ FOV exposures}}$$

Part (b)

Your proposal review insists that you need at least 100 galaxies and requires a 99% chance that you will succeed. How many exposures do you need now?

$$P_n(x) = \frac{19.3}{x}\, n$$

$$0.99 < P_n(x=100) = \frac{19.3}{100}\, n$$

$$\Rightarrow 0.99 < \frac{19.3}{100}\, n$$

$$\boxed{n > 5.130} \implies \boxed{\text{at least } 5 \text{ FOV exposures}}$$

Part (c)

There is a problem, however. A "contaminating" star is one that is bright enough to saturate the image to the point that one cannot determine the properties of galaxies. The expected (mean) number of "contaminating" stars one FOV is 5.1. Your assignment is to find enough FOVs in the sky without contaminating stars to get the needed long-term exposures. Your advisor (who did not take statistics) comes up with a simple plan. Divide the sky in sectors, each with an area of the FOV with no overlap and take a short images; a contaminating star will emerge quickly. How many sectors do you expect to examine to find enough uncontaminated FOVs?

$$\langle S_c \rangle = 5.1 \qquad \equiv \frac{\text{star}}{\text{FOV}}$$

$$P_p(S_c = 0 ; \mu = 5.1) = \frac{\mu^x}{x!} e^{-\mu} = \frac{(5.1)^0}{0!} e^{-5.1} = e^{-5.1} \approx 0.006097 \qquad \equiv \begin{array}{l}\text{Probability of 0 contaminated stars} \\ \text{per } \underline{1} \text{ FOV}\end{array}$$

$$\Rightarrow P_p(S_c = 0 ; \mu = 5.1) * n = 100\% = 1$$

Side note: I'm not sure what qualifies as "enough" uncontaminated FOVs but even 90% probability requires more than 100 FOVs

$$(0.006097) n = 1$$

$$\boxed{n = 164.021}$$

Part (d)

What an effort! But another problem comes up. Due to the long exposure time and scheduling limitations of the telescope, there is a limited region of the sky that can be used. It is 1000 as by 1000 as whereas your FOV is a 100 as diameter circle. At best there are only a little over 100 such sectors. You now remember the "Find the Gap" lecture and you realize that there are many more possible exposures. You take a star map of the entire usable region. Demonstrate you can find enough open regions. Your choice as to how.

Hint: One way is to create a random star map using positions as $x =$ uniform random between 0 and 1000 as and $y =$ uniform random between 0 and 1000 as. Find values on a 1000×1000 grid $[x, y]$ in which no stars are within 50 as.

```
# will come back to this later
```

2. Multi-variant Uncertainties

Stellar parallax measurements are made by measuring the relative motion of a nearby star (S_A) with respect that of a distant object (S_B) due to Earth's orbit about the Sun. All images, however, have finite resolution; even though a star should make a single point on an image the star's intensity has a finite width, often referred to a point-spread function. The probability that a photon is recorded at position x is:

$$P(x) = \frac{1}{\Delta x \sqrt{2\pi}} e^{-\frac{(x-x_0)^2}{2\Delta x^2}} \tag{1}$$

Part (a)

In the above equation, Δx is a property of, for example, the telescope diameter and the CCD resolution. Suppose $\Delta x = 20$ pixels. Star S_A records 10^3 photons and S_B yields 10^4 photons. What is the $1 - \sigma$ variance in the position of each of the stars (in *pixels*)?

$$\Delta x = 20 \text{ pixels}$$
$$\varphi_A = 10^3 \text{ photons}$$
$$\varphi_B = 10^4 \text{ photons}$$

$$\sigma_{\mu}'^2 = \frac{\sigma_i^2}{N}$$

$$\sigma_{\mu_A}'^2 = \frac{\sigma_A^2}{N_A} = \frac{20}{10^3} = 0.02$$

$$\sigma_{\mu_B}'^2 = \frac{20}{10^4} = 0.002$$

Part (b)

Given the above, what is the uncertainty in the distance between the stars?

"uncertainty in distance" = variance of variance (I'm guessing?)

$$\sigma'^2 = \frac{1}{N} \sum_i (x_i - \mu)^2 = \frac{1}{N} \sum_i (x_i - x_0)^2 \quad \Rightarrow x_0 = ? \Rightarrow \sigma'^2 \approx \frac{1}{N-M} \sum_i (x_i - \mu')$$

$$\mu' = \frac{1}{N} \sum_i x_i \qquad ??? \qquad \left(\begin{array}{c} \text{I will come back} \\ \text{to this ...} \end{array} \right)$$

Part (c)

The calibration relating *pixels* to *mas* is $c = 10$ mas/pixel $\pm 0.2\%$. Suppose S_A and S_B are separated by 250 pixels. What is d and the uncertainty in d in *mas*? Which uncertainty (a, b, or c) makes the largest contribution to the uncertainty in d?

```
delta_x = 250
c = 10

d = delta_x*c
print("d = {}".format(d), "mas")
```

d = 2500 mas

$$d = |a-b| * c$$

$$\sigma_d^2 = \sigma_a^2\left(\frac{\partial d}{\partial a}\right)^2 + \sigma_b^2\left(\frac{\partial d}{\partial b}\right)^2 + \sigma_c^2\left(\frac{\partial d}{\partial c}\right)^2 + \cdots$$

$$a-b = 250 \implies b = 250+a$$

$$= (0.02)^2\left(b\cdot c\right)^2 + (0.002)^2\left(a\cdot c\right)^2 + (0.002)^2\left(a-b\right)^2$$

$$= (4\times10^{-4})(10\,b)^2 + (4\times10^{-6})(10\,a)^2 + (4\times10^{-6})(250)^2$$

$$= 4b^2\times10^{-2} + 4a^2\times10^{4} + (4\times10^{-6})(62500)$$

$$= 4b^2\times10^{-2} + 4a^2\times10^{-4} + 0.25$$

$$= 4(250+a)^2\times10^{-2} + 4a^2\times10^{-4} + 0.25$$

$\left(\text{progress so far, but I will come back to try again/finish later}\right)$

Uncertainty in d:

3. Gamma Ray Bursts

Gamma Ray Bursts (GRBs) were first detected in the 1960s by defense satellites that were designed to verify the nuclear test ban treaty. These detectors had little ability to determine the source direction, but enough to determine that GRBs were not from Earth.

To investigate GRBs, the next satellite-borne detectors had all-sky coverage combined with separate detectors with fields-of-view of roughly 0.04 sr (0.3% of the sky; gamma rays are difficult to focus). The GRB energy is so high that it was at thought by many (at the time) that GRBs must be of galactic origin. GRBs are detected at a rate of about one per day.

Part (a)

Start by dividing the sky into two equal parts; one part contains the galactic plane and the other part is the rest of the sky (which has two sub-parts, but not a concern here). Most models predicted that a majority (at least 60%) of GRBs should be detected in a 2π sr region containing the galactic plane. To test this hypothesis, assume that 60 of GRBs are coming from the galactic plane. Estimate (no computer) how long was it expected to take the observations to show that the fraction of events in the galactic plane is 2σ higher than 50%, proving a galactic origin model.

```
# will come back to this later
```

Part (b)

After roughly one year, say that 166 GRBs are detected in the galactic plane out of 320 events. Let's examine the expected value of r_{GP}, the ratio of the number of events from galactic plane to the number of total events. By computer, compute and plot the Gaussian probability distribution as a function of r_{GP} using the mean and σ from the above statistics. Why Gaussian? The number of events may be a bit too high to compute Poisson statistics directly. Mark the $\pm 2\sigma$ area. Compute the CDF. What is the probability that r_{GP} is 60% or higher?

```
# will come back to this later
```

Part (c)

The idea that GRBs are extra-galactic was difficult to accept. However, after 9 years 2512 GRBs were detected of which close to 50% came from the galactic plane, which ruled out all but the most die-hard galactic-origin theories. The next test was to determine if there was any bias in GRB direction. The sky was parsed into 300 equally-sized sectors. Suppose the histogram is:
$$H = [0, 1, 2, 8, 9, 25, 37, 37, 40, 36, 41, 23, 23, 8, 3, 3, 2, 1, 1]$$

Does the GRB histogram match a Poisson distribution?

Hint: Cut and paste H into your code and plot. Overplot the expected distribution of the 300 sectors. What is the most likely number of GRBs per sector (μ')? What is χ_ν^2? How does χ_ν^2 compare with the expected distribution? What do you conclude?

```python
import os,sys
parent_dir = os.path.abspath("..")   # string path to parent directory of file
os.chdir(parent_dir)                 # changes cwd to specified directory path


from hw_helper_func import poisson
# I'm importing code I made separately in a different python script file
# feel free to ask me for it if needed!
# But it's actually the same code that I've written out in HW1 & HW2
```

```python
H = [0,1,2,8,9,25,37,37,40,36,41,23,23,8,3,3,2,1,1]

# getting the mean value based on the values of "H"
ind = []
for i in range(len(H)):
    ind.append(i*H[i])

s = sum(ind)
mean = s/(sum(H))

print("mean (based on histogram results):",mean)
```
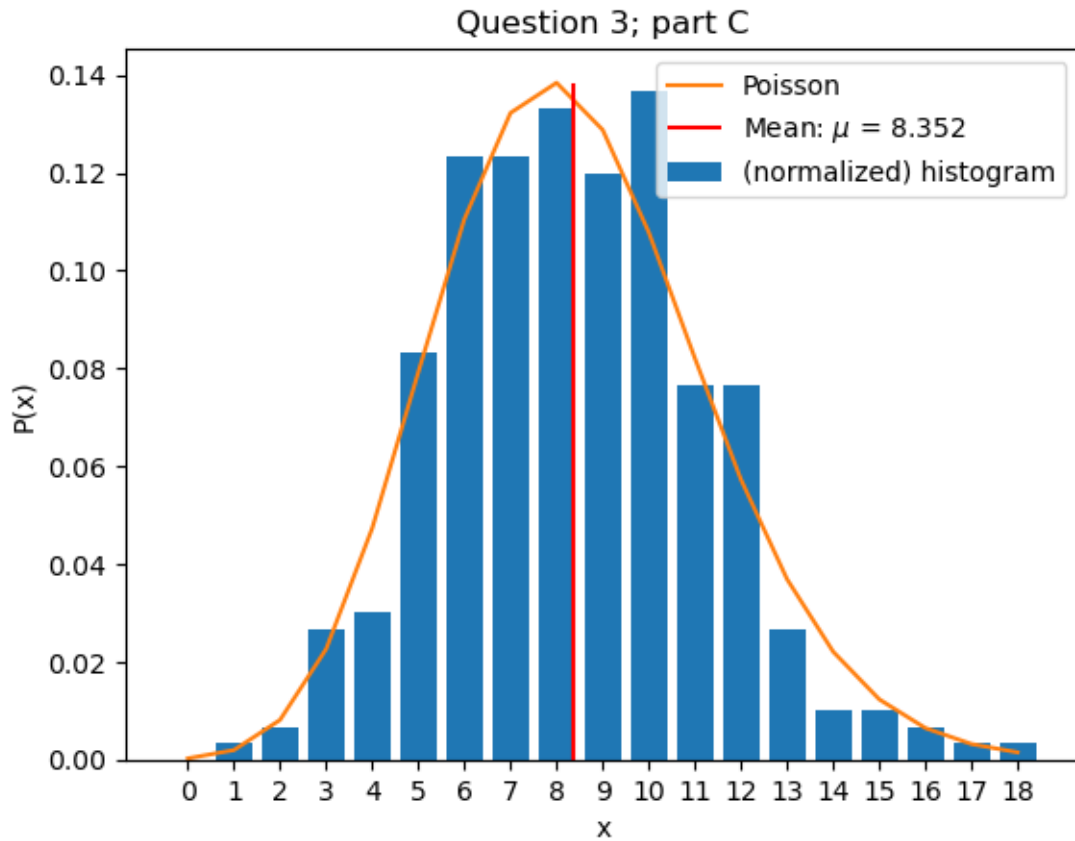
```
mean (based on histogram results): 8.373333333333333
```

```python
q3c = poisson(x_min=0,x_max=len(H)-1,mu=mean)
# print("x:",q3c.x)

plt.bar(q3c.x,np.array(H)/sum(H), label='(normalized) histogram')
q3c.plot_poisson(title="Question 3; part C",label='Poisson')
plt.xticks(q3c.x)

q3c.calculate_mean_mode_sigma(x=q3c.x,Px=q3c.Px)
q3c.plot_mark_mean(Px=q3c.Px,colors='red',linestyles='solid')

plt.legend()
plt.show()
```

Jasmine's note to self

$$\chi_N^2 = \sum_{i=1}^{N} \frac{(x_i - \mu')^2}{\sigma^2}$$

$$\approx \frac{\text{expected variance}}{\text{actual variance}} = \frac{\sigma'^2}{\sigma^2}$$

```
P_Hx = np.array(H)/sum(H)   # "probabilities" based on histograms
sigma2_prime = q3c.Expectation(fx=(q3c.x-mean)**2,Px=P_Hx)    # sigma^2 from histogram results

chi2 = sigma2_prime/(q3c.sigma**2)
print("chi-squared:",chi2)
```

```
chi-squared: 0.9604764044774297
```

Part (d)

To perform another χ^2_ν test, all of the 300 sectors are expected to have, within error, the same number of GRBs. That is, $f(x) = \mu'$, which is now well determined. We also know $\sigma^2 = \mu'$ well. What is χ^2_ν? What do you conclude?

Hint: Explode H into a 300-point array (X): $X = [0, 1, 2, 3, 3, 3, 3, 3, 3, 3, 3, (9\ 4s), (25\ 5s), \text{etc.}]$

Jasmine's note to self

$$\chi^2_N = \sum_{i=1}^{N} \frac{(x_i - \mu')^2}{\sigma^2}$$

$$\approx \sum_{i=1}^{N} \frac{(x_i - \mu')^2}{\mu'}$$

```
X = []

for i,hi in enumerate(H):
    X += [i]*hi

print("Check that X is what I want it to be")
print("Preview of X (first 10 elements):",X[:10])
print("Number of elements in X (should be 300):",len(X))
```

```
Check that X is what I want it to be
Preview of X (first 10 elements): [1, 2, 2, 3, 3, 3, 3, 3, 3, 3]
Number of elements in X (should be 300): 300
```

```
chi_i = []

for i,xi in enumerate(X):
    numerator = (xi - mean)**2
    denominator = mean
    chi_i.append(numerator/denominator)

print("chi-squared:",sum(chi_i)/len(X))
```

```
chi-squared: 0.944341825902335
```