CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

1. PMLS2, problem 3.3

3.3 *Random walk displacement distribution*

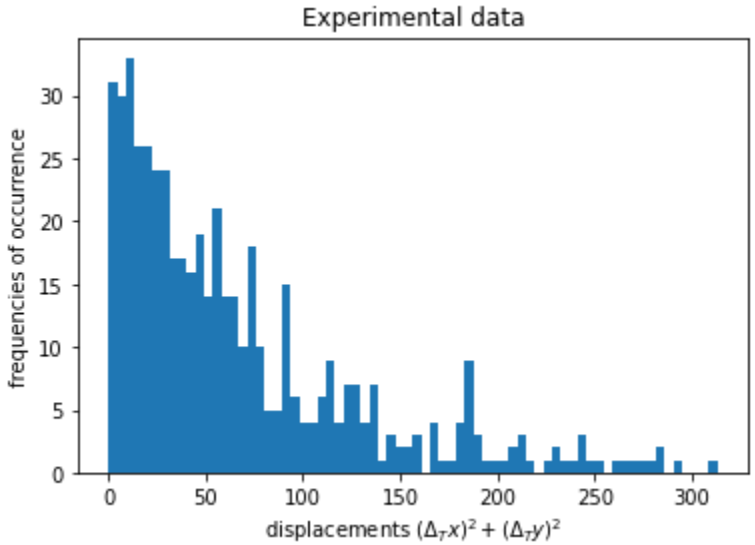
This problem introduces the "random walk," a physical model for Brownian motion. Get Dataset 4, which contains experimental data.

A micrometer-scale particle undergoing Brownian motion was observed at successive times separated by $T = 30\,\mathrm{s}$ intervals. In either of the data files, the two columns contain $\Delta_T x$ and $\Delta_T y$, the relative displacements of a particle from one observation to the next. There were many molecular collisions during each such interval.

Figures 3.3a,b show the cumulative displacements, that is, the actual positions (relative to the start), throughout two such trajectories. You can see that some of the displacements are longer than others. So first, investigate the length-squared of the displacements:

a. Tabulate the values of $(\Delta_T x)^2 + (\Delta_T y)^2$ in each of the two experimental data sets. Find the range of values that occur, subdivide it into a suitable set of bins, classify each sample into the corresponding bin, and display the frequencies of occurrence as a histogram. [*Hint:* What is a "suitable" set of bins? To keep it simple, make all bins have equal width. They should be wide enough that at least some of them have lots of counts (say, more than 30)—enough to estimate accurately their probability. Bins should also be narrow enough to see the general trend of their probabilities.]

Range: (Min: 1.2320099999999998e-29, Max: 312.9451718552014)
Subdivided into 70 bins



See code and values in the .ipynb.

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

Point **5a** (page 36) proposed that this motion arises from a huge number of molecular collisions in between each observation. To model those invisible individual collisions, imagine that a tiny chesspiece is placed at the center of a tiny chessboard, at a point we'll call $(0, 0)$. Once per second, the chesspiece is moved along x.

> Each displacement is by a distance $L = 1\ \mu$m along either the $+x$ or $-x$ direction. The choice is random, each direction is equally probable, and each step is statistically independent of all the others.

first physical model for Brownian motion

(3.30)

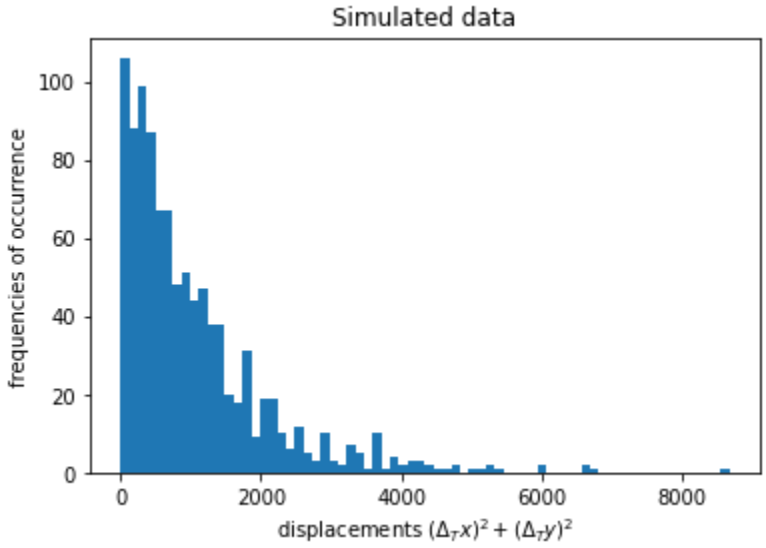
At the same time, the chesspiece is also moved randomly by the same distance L along the $\pm y$ direction; again, the choice is random, with equal probability for each choice, and independent of the x step.

In short, the elementary steps have $\Delta x = \pm L$ and $\Delta y = \pm L$. Imagine making many trajectories, all starting at the same point $(0, 0)$.

b. Simulate 1000 such walks. To do this, at each step draw *two* independent random variables, both equal to $\pm 1\ \mu$m, and let these individual displacements accumulate. Repeat the steps in (a), but with your simulated total displacements after 500 steps instead of the experimental data. Does your answer qualitatively resemble the experimental result in (a)?

Range: (Min: 0, Max: 8676)

Yes, the distribution is similar to the result in (a), but the range is wider.



See code and values in the .ipynb.

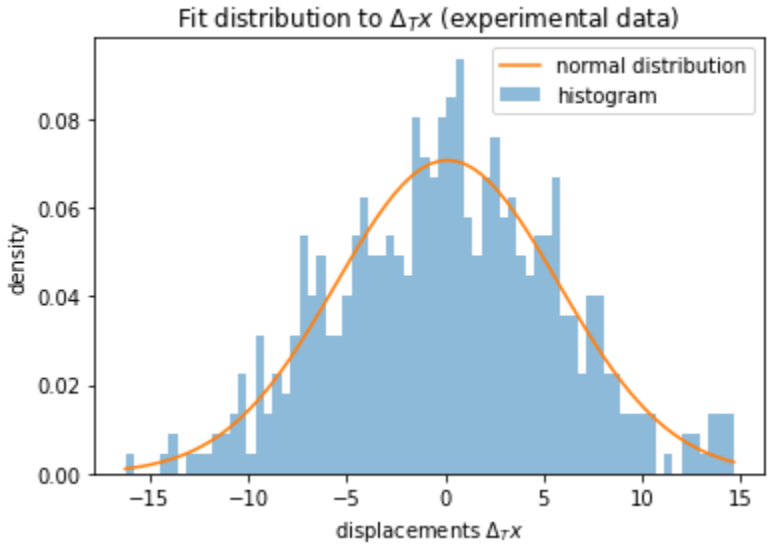
CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

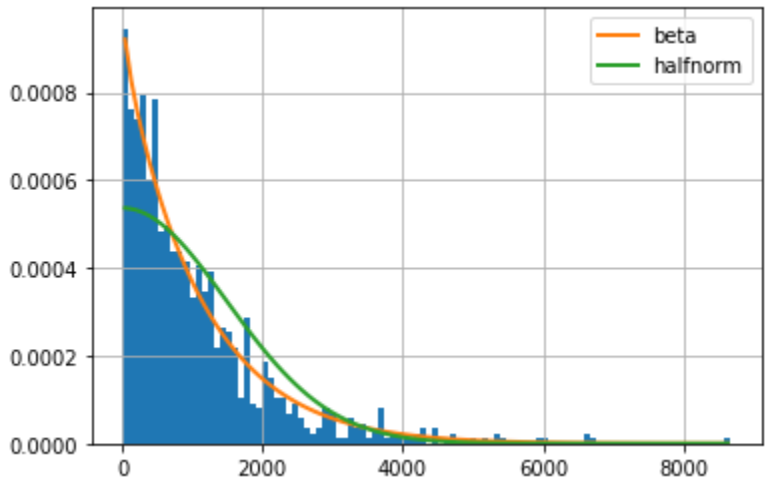
c. Do your results in (a,b) suggest a possible mathematical form for the distribution of this quantity? How could you replot both the experimental and simulated data to check this idea? Try it. What could you change in your simulation to increase your confidence?

Yes, $\Delta_T x$ and $\Delta_T y$ values can be fit to normal distribution.
The length-squared of the displacements $(\Delta_T x)^2 + (\Delta_T y)^2$ can be fit to beta distribution.
Example:

1. $\Delta_T x$ of experimental data:



2. $(\Delta_T x)^2 + (\Delta_T y)^2$ of simulated data:



We can increase the data point in simulation to increase confidence.

See code and values in the .ipynb.

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

2. PMLS2, problem 3.5

3.5 *Negation rule*

a. Suppose that you are looking for a special type of cell, perhaps those tagged by expressing a fluorescent protein. You spread a drop of blood on a slide marked with a grid containing N boxes, and examine each box for the cell type of interest. Suppose that a particular sample has a total of M tagged cells. What is the probability that at least one box on the grid contains more than one of these M cells? [*Hint:* Each tagged cell independently "chooses" a box, so there are N^M equally probable outcomes. Find how many of these outcomes have *no* box with more than one tagged cell, and then use the negation rule.]

$$N\ boxes,\ M\ cells,\ N^M probable\ outcomes$$

$$\#(at\ least\ 1\ box\ contains\ >\ 1\ cell)\ =\ N^M\ -\ \#(no\ box\ with\ >\ 1\ cell)$$

$$no\ box\ with\ >\ 1\ cell$$

$$\Rightarrow a\ cell\ cannot\ choose\ a\ box\ which\ is\ already\ occupied\ by\ another\ cell$$

$$\Rightarrow permutation\ question:\ P_M^N = \frac{N!}{(N-M)!}$$

The probability that at least one box on the grid contains more than one of these M cells

$$= 1 - \frac{P_M^N}{N^M} = 1 - \frac{\frac{N!}{(N-M)!}}{N^M}$$

b. Evaluate your answer for $N = 400, M = 20$.

$$1 - \frac{\frac{N!}{(N-M)!}}{N^M} = 1 - \frac{\frac{400!}{380!}}{400^{20}} = 1 - 0.6169838377 = 0.3830161623$$

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

3. PMLS2, problem 3.6

3.6 *Vaccine efficacy*

In 2021, a real country reported that over a certain period it had 214 severe cases of SARS-CoV-2 infection among unvaccinated persons, and 301 severe cases among fully vaccinated persons. On the face of it, this looks bad—many commentators concluded that the vaccine in question, initially so promising, had lost effectiveness. Let's look more carefully.

Here is are some more detailed, real, data:

	unvaccinated		vaccinated	
	age < 50	age ≥ 50	age < 50	age ≥ 50
entire population	1 116 833	186 077	3 501 117	2 133 515
severe cases	43	171	11	290

Persons not included in the second row either did not get ill at all, or got a mild case (not requiring hospitalization).

Clear thinking begins with the question, "What do I really want to know?" In this situation, one may really want to know

How much will vaccination change ***my*** *chances of getting severely ill?* (3.31)

Conditional probability gives a more sure-footed approach to addressing this question than just comparing 214 to 301.

a. Combine the two age groups to estimate the overall $\mathcal{P}(\text{severe} \mid \text{vax})$ and $\mathcal{P}(\text{severe} \mid \text{unvax})$.

$$P(severe \mid vax) = \frac{301}{3501117+2133515} = \frac{301}{5634632} = 5.34196377 \times 10^{-5}$$

$$P(severe \mid unvax) = \frac{214}{1116833+186077} = \frac{214}{1302910} = 1.64247722 \times 10^{-4}$$

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

b. We can define the vaccine's **efficacy** against severe illness via the ratio of those quantities:

$$1 - \frac{\mathcal{P}(\text{severe} \mid \text{vax})}{\mathcal{P}(\text{severe} \mid \text{unvax})}.$$

Why is this a reasonable name for this quantity? Compute it for the overall population.

The formula represents a proportionate reduction in severe cases between the unvaccinated and vaccinated populations. That is 1 - the relative risk of getting severe illness for vaccinated people versus unvaccinated people.

$$1 - \frac{5.34196377\times10^{-5}}{1.64247722\times10^{-4}} = 0.6747617742$$

c. Merely dividing 214/301 amounts to computing $\mathcal{P}(\text{severe}$ **and** $\text{vax})/\mathcal{P}(\text{severe}$ **and** $\text{unvax})$. Why is this not a very informative response to (3.31)?

Because this formula does not take the population number of vax and unvax into consideration. It is comparing two different groups of people that have two different backgrounds (conditions).

We can do even better. You are an individual, not a randomly chosen member of a population. To answer question (3.31), then we should ask,

> *How much did vaccination change the outcomes for* ***people who are like me in relevant aspects?***

The outcome of infection depends on *age,* and you have a particular age.[32] Crucially, in this population the rates of vaccination were also quite different for different ages.

d. Estimate $\mathcal{P}(\text{severe} \mid \text{vax}, \text{age})$ and $\mathcal{P}(\text{severe} \mid \text{unvax}, \text{age})$ for each of the two age groups given in the table. We call this procedure **stratifying** by age.

$$P(severe \mid vax,\ age < 50) = \frac{11}{3501117} = 3.14185444 \times 10^{-6}$$

$$P(severe \mid vax,\ age \geq 50) = \frac{290}{2133515} = 1.35925925 \times 10^{-4}$$

$$P(severe \mid unvax,\ age < 50) = \frac{43}{1116833} = 3.85017277 \times 10^{-5}$$

$$P(severe \mid unvax,\ age \geq 50) = \frac{171}{186077} = 9.18974403 \times 10^{-4}$$

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

e. Compute the efficacy against severe illness for each age group separately and comment.

In the population of age < 50:

$$1 - \frac{P(severe \mid vax, age<50)}{P(severe \mid unvax, age<50)} = 1 - \frac{3.14185444\times10^{-6}}{3.85017277\times10^{-5}} = 0.9183970531$$

In the population of age ≥ 50:

$$1 - \frac{P(severe \mid vax, age\geq50)}{P(severe \mid unvax, age\geq50)} = 1 - \frac{1.35925925\times10^{-4}}{9.18974403\times10^{-4}} = 0.8520895418$$

The vaccine efficacy is higer for the younger population.

f. Now represent these ideas graphically. First restrict attention to the younger group. There are four subcategories of such people. Draw squares whose areas are an overall constant times the joint probabilities $\mathcal{P}$(vaxed **and** not-severe | under 50), and the other three such probabilities. (Actually, magnify the two squares representing severe cases by equal amounts, to make them big enough to see.) Explain how your picture relates to your answer in (c).

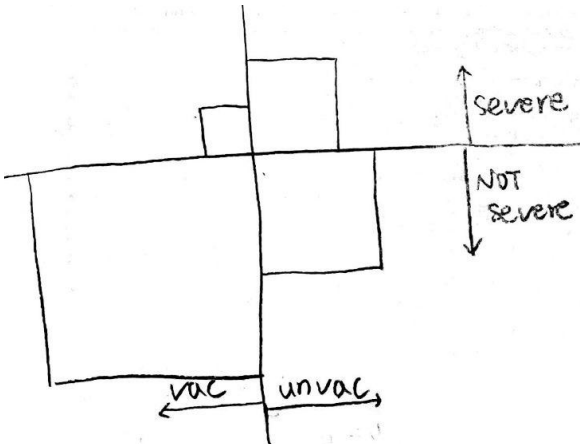
$$age < 50 = 1116833 + 3501117 = 4617950$$

$$P(vax\ \&\ severe \mid age < 50) = \frac{11}{4617950} = 2.38200933 \times 10^{-6}$$

$$P(vax\ \&\ not\ severe \mid age < 50) = \frac{3501117-11}{4617950} = 0.7581515608$$

$$P(unvax\ \&\ severe \mid age < 50) = \frac{43}{4617950} = 9.31149103 \times 10^{-6}$$

$$P(unvax\ \&\ not\ severe \mid age < 50) = \frac{1116833-43}{4617950} = 0.2418367457$$



CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

g. Repeat for the over-fifty crowd.

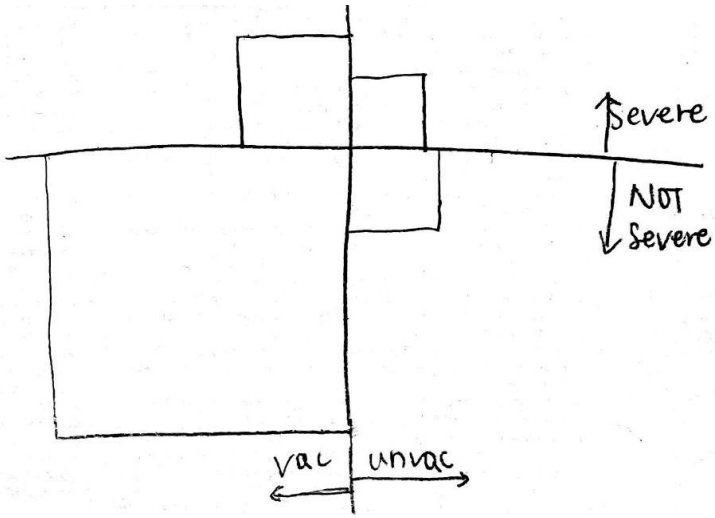
$$age \geq 50 = 186077 + 2133515 = 2319592$$

$$P(vax\ \&\ severe \mid age \geq 50) = \frac{290}{2319592} = 1.25021987 \times 10^{-4}$$

$$P(vax\ \&\ not\ severe \mid age \geq 50) = \frac{2133515-290}{2319592} = 0.919655267$$

$$P(unvax\ \&\ severe \mid age \geq 50) = \frac{171}{2319592} = 7.37198611 \times 10^{-5}$$

$$P(unvax\ \&\ not\ severe \mid age \geq 50) = \frac{186077-171}{2319592} = 0.08014599119$$



If we compare the not-severe cases in the over-fifty crowd, the probability of not-severe for vax is higher than not-severe for unvax. However, if we look at severe cases, the vaccinated population is higher than unvaccinated. Because the rates of vaccination is much higher in over-fifty crowd, we can see the probability of vax & severe is a little higher than unvax & severe.

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

4. PMLS2, problem 3.7

3.7 *Virus evolution*

The genome of the HIV virus, like any genome, is a string of "letters" (base pairs) in an "alphabet" containing only four letters. The message for HIV is rather short, just $n \approx 10^4$ letters in all.

The probability of errors in reverse transcribing the HIV genome is about one error for every $3 \cdot 10^4$ "letters" copied. Suppose that each error replaces a DNA base by one of the three other bases, chosen at random. Each time a virion infects a T cell, the reverse transcription step creates opportunities for such errors, which will then be passed on to the offspring virions. The total population of infected T cells in a patient's blood is roughly 10^7 in the quasi-steady state.[33]

a. Find the probability for a T cell infection event to generate one *particular* error, for example, the one lucky mutation that could confer resistance to a drug. Multiply by the population to estimate the number of T cells already present with a specified mutation (prior to administering any drug). Those cells will later release resistant virions.

$$error\ rate \times one\ specific\ position \times one\ particular\ nucleotide \times \#\ infected\ T\ cells$$

$$\frac{1}{3\times10^4} \times \frac{1}{C_1^{10^4}} \times \frac{1}{3} \times 10^7 = 1.11111111 \times 10^{-2}$$

b. Repeat (a), but this time for the probability of spontaneously finding *two* specified errors, and comment. What about *three*?

[*Note:* You may assume that each infected T cell was infected by a wild-type virion, so that mutations do not accumulate. For example, the wild-type may reproduce faster than any mutant, crowding them out in the quasi-steady state.]

Finding two specified mutations:

$$\left(\frac{1}{3\times10^4}\right)^2 \times \frac{1}{C_2^{10^4}} \times \left(\frac{1}{3}\right)^2 \times 10^7 = 2.46938274 \times 10^{-11}$$

It's 10^9 times less likely to happen.

Finding three specified mutations:

$$\left(\frac{1}{3\times10^4}\right)^3 \times \frac{1}{C_3^{10^4}} \times \left(\frac{1}{3}\right)^3 \times 10^7 = 8.23292239 \times 10^{-20}$$

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

5. PMLS2, problem 4.3 (5 pts)

4.3 *Gene frequency*

Consider a gene with two possible variants (alleles), called A and a.

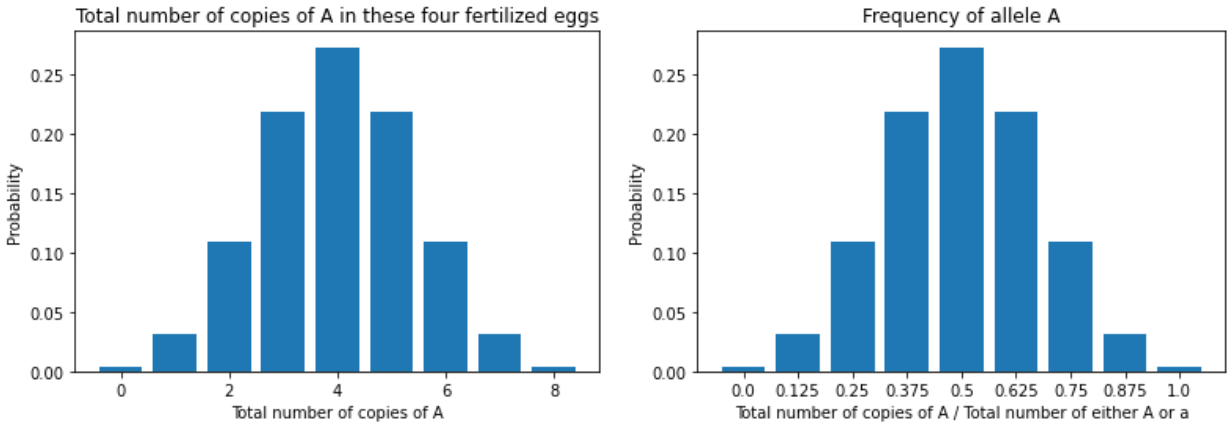
Father Fish has two copies of this gene in every somatic (body) cell; suppose that each cell has one copy of allele A and one copy of a. Father Fish releases a zillion sperm, each with just one copy of the gene. Mother Fish also has genotype Aa. She releases a zillion eggs, again each with just one copy of the gene.

Four sperm and four eggs are drawn at random from these two pools and fuse, giving four fertilized eggs, which grow as usual.

a. What is the total number of copies of A in these four fertilized eggs? Re-express your answer in terms of the "frequency of allele A" in the new generation, which is the total number of copies of A in these four individuals, divided by the total number of either A or a. Your answer should be a symbolic expression.

The probabilities for each "total number of copies of A" in the new generation should be a binomial distribution.

Total number of copies of A in these four individuals: I = [0,1, 2,...,8]
Total number of either A or a: M = 8
$P(an\ egg\ or\ a\ sperm\ with\ A) = 0.5$



CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

b. What is the probability that the frequency of allele A is exactly the same in the new generation as it was in the parent generation (two individuals)? Your answer should be a number. What is the probability that the frequency of allele A is *zero* in the new generation?

Frequency of allele A in the parent generation = 2/4 = 0.5

8 * 0.5 = 4

$$P(4\ fertilized\ eggs\ have\ 4\ A\ in\ total)$$
$$= P_{binom}(4; 0.5, 8)$$
$$= \frac{8!}{4!(8-4)!} 0.5^4 (1 - 0.5)^{8-4}$$
$$= 0.2734375$$

$$P(4\ fertilized\ eggs\ have\ 0\ A\ in\ total)$$
$$= P(4\ fertilized\ eggs\ all\ have\ aa) = 0.25^4$$
$$= 0.00390625$$

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

6. PMLS2, problem 4.6 (5 pts)

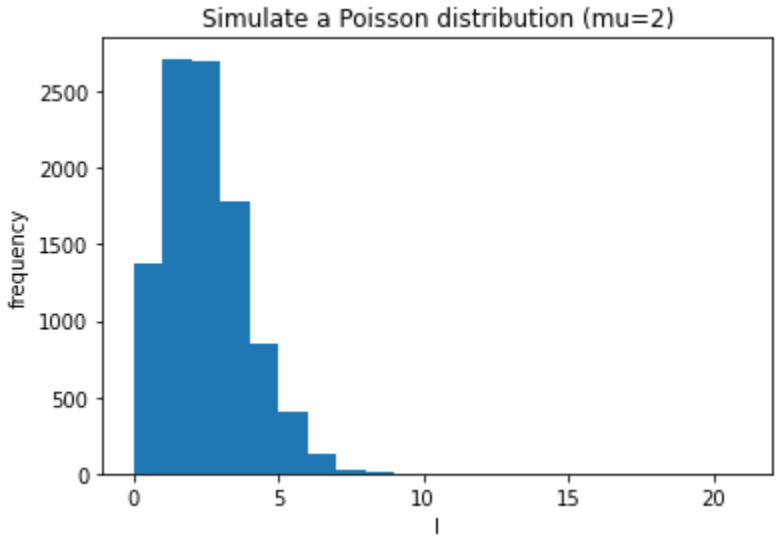
4.6 *Simulate a Poisson distribution*

a. Write a function for your computer called `poissonSetup(mu)` that is similar to the one described in Your Turn 4B, but which prepares a set of bin edges suitable for simulating a Poisson distribution with expectation `mu`. In principle, this distribution has infinitely many bins, but in practice you can cut it off; that is, use either 10 or 10*`mu` bins (rounded to an integer), whichever is larger. (Or you may invent a better way to set a finite cutoff.)

See code in the .ipynb. (4.6.a)

b. Write a little "wrapper" program that calls `poissonSetup(2)`, generates 10 000 numbers from the distribution, finds the sample mean, and makes a histogram. Also estimate the variance by replacing the expectations in Equation 3.20 (page 54) by sample means and evaluating them.

mean = 1.9957
variance = 2.02408151



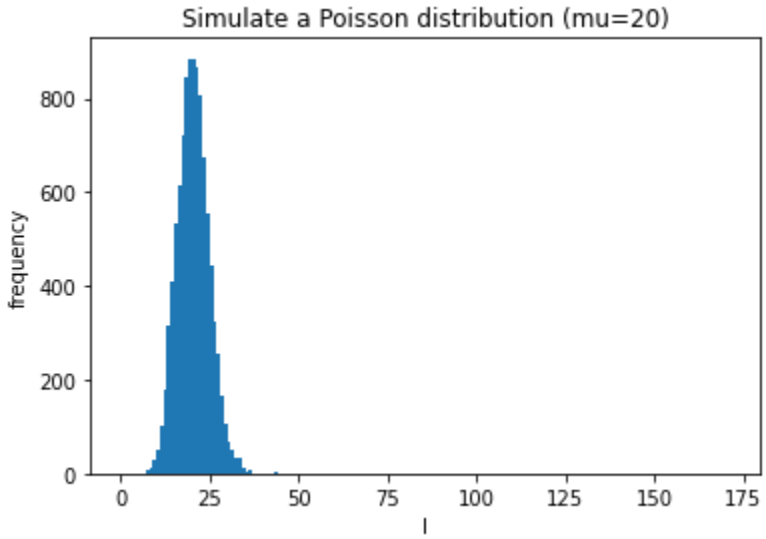
See code and values in the .ipynb. (4.6.b)

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

c. Repeat with `mu` = 20, and comment on the different symmetry of the peak between this case and (b). Confirm that, for a big enough sample, the sample mean and variance you found agree with direct calculation from the definition of the Poisson distribution.

mean = 19.941
variance = 19.828919



(c) is more symmetric than (b).

The direct calculation from the definition of the Poisson distribution:

$\langle \ell \rangle = \mu$ for the Poisson distribution with parameter μ.

$\operatorname{var} \ell = \langle \ell \rangle$ for any Poisson distribution.

The values of mean and variance are both very close to μ (=20)

See code and values in the .ipynb. (4.6.c)

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

7. PMLS2, problem 4.9 (5 pts)

4.9 *Cultures and colonies*

a. Suppose that you add $2 \cdot 10^8$ virions to a culture containing 10^8 cells. Suppose that every virus "chooses" a cell at random and successfully infects it, but some mechanism prevents infected cells from dying. A cell can be infected by more than one virus, but suppose that a prior infection doesn't alter the probability of another one. What fraction of the cells will remain uninfected? How many virions would have been required had you wished for over 99% of the cells in the culture to be infected?

Because the total number of cells (V_*) and the total number of virions (M_*) are both very big, we can assume that the number of virions that will infect each cell is a Poisson distribution with mu=2.

$$\mu = \langle l \rangle = \frac{2\times10^8}{10^8} = 2$$

$P_{pois}(0;2) = 0.1353352832366127$
Thus, about 13.5335% of the total 10^8 cells will remain uninfected.
To let over 99% of the cells be infected, we need a mu so that $P_{pois}(0;\mu) < 0.01$
$e^{-\mu} = 0.01$
$\mu = -\ln 0.01 = 4.605170186$

4.605170186 * 10^8 virions are required for over 99% of the cells to be infected.

b. Suppose that you take a bacterial culture and dilute it by a factor of one million. Then you spread 0.10 mL of this well-mixed, diluted culture on a nutrient plate, incubate, and find 110 well-separated colonies the next day. What was the concentration of live bacteria (colony forming units, or CFU) in the original culture? Express your answer as CFU/mL and also give the standard deviation of your estimate.

Diluted: 110 CFU /0.1mL = 1100 CFU/mL
Original: 1100 * 1,000,000 = $1.1*10^9$ CFU/mL
Assume this is a Binomial distribution, the standard deviation is equal to the square root of [$n * P * (1 - P)$].
Standard deviation: $\sqrt{1.1 \times 10^9 \times 10^{-6} \times (1 - 10^{-6})} = 33.16623132$

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

8. PMLS2, problem 4.16 (10 pts)

4.16 *Luria-Delbrück experiment*

First do Problem 4.6, and be sure that your code is working the way you expect before attempting this problem.

Imagine a set of C cultures (separate flasks) each containing n_0 bacteria initially. Assume that all the cells in a culture divide at the same time, and that every time a cell divides, there is a small probability α_g that one of the daughter cells will mutate to a form resistant to phage attack. Assume that the initial population has no resistant mutants ("pure wild-type"), and that all progeny of resistant cells are resistant ("no reversion"). Also assume that mutant and wild-type bacteria multiply at the same rate (no "fitness penalty"), that at most one of the two daughter cells mutate (usually neither), and that no cell ever dies.

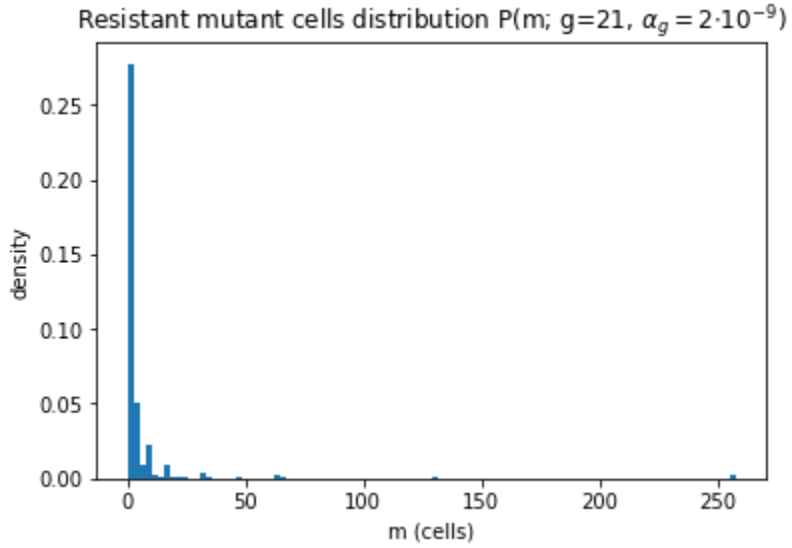
a. Write a computer code to simulate the situation and find the number of resistant mutant cells in a culture after g doublings. The Poisson distribution gives a good approximation to the number of new mutants after each doubling, so use the code you wrote in Problem 4.6. Each simulated culture may end up with a different number m of resistant mutant cells, due to the random character of mutation.

See code and values in the .ipynb. (4.16.a)

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

b. For $C = 500$ cultures with $n_0 = 200$ cells initially, and $\alpha_g = 2 \cdot 10^{-9}$, find the number of cultures with $m = 1, 2, \ldots$ resistant mutant cells after $g = 21$ doublings, for every value of the random variable m. Plot your result as an estimated probability distribution $\mathcal{P}(m; g, \alpha_g)$. Find its sample mean. Estimate its variance by replacing the expectations in Equation 3.20 (page 54) by sample means. Compare the estimated variance of m to the sample mean and comment.



Mean = 5.076
Variance = 474.498224

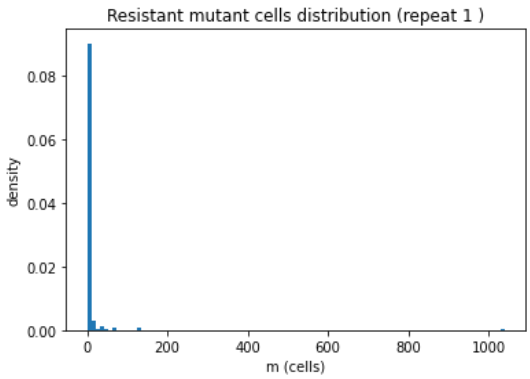
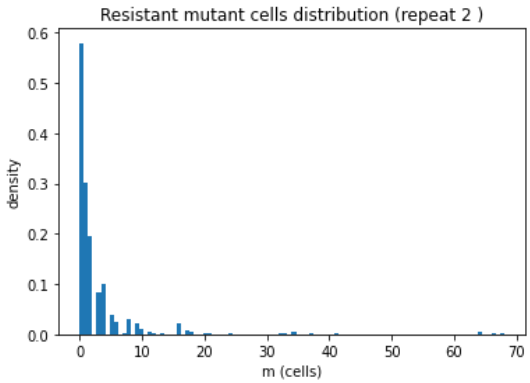
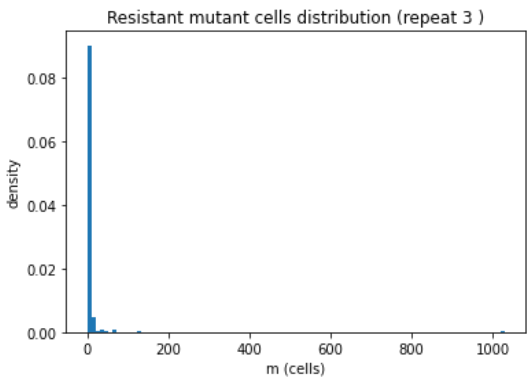
The estimated variance is very big compared to the sample mean, because there are some data points (m>250) far from the main group (m<25).

See code and values in the .ipynb. (4.16.b)

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

c. Repeat the simulation $M = 3$ times (that is, make three graphs each for 500 cultures, and report three estimates of sample mean and of estimated variance). Then comment on how accurately we can expect to find the true expectation and variance of the distribution from such experiments.

	mean: 6.164 var: 2320.8971040000006
	mean: 3.084 var: 57.204944
	mean: 5.558 var: 2187.4266359999997

We cannot expect to find the true expectation and variance of the distribution accurately because they vary a lot among the experiments.

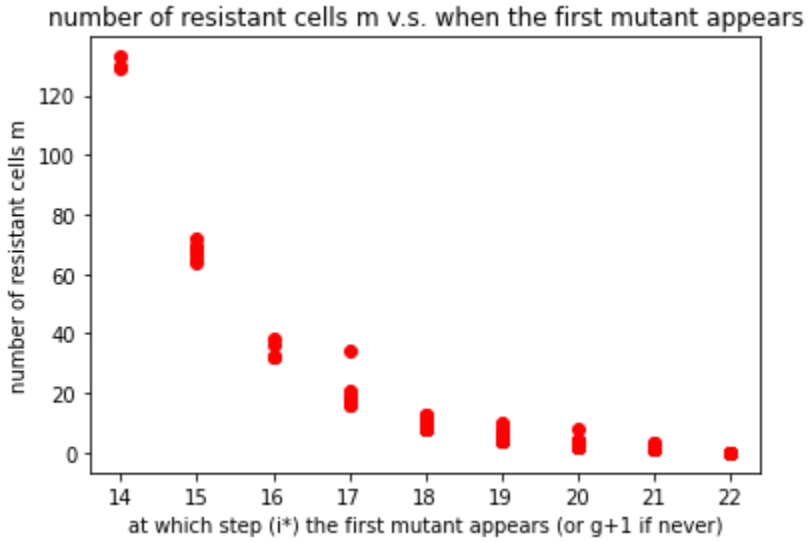
See code and values in the .ipynb. (4.16.c)

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

d. The chapter claimed that the origin of the long tail in the distribution is that, on rare occasions, a resistant mutant occurs earlier than usual, and hence has many offspring (Figure 4.7, page 87). For each simulated culture, let i_* denote at which step (number of doublings) the *first* mutant appears (or $g + 1$ if never). Produce a plot with m on the vertical axis and i_* on the horizontal, and comment.

[*Hints:* (*i*) This project may require a couple dozen lines of code. Outline your algorithm before you start to code. Keep a list of all the variables you plan to define, and give them names that are meaningful to you. (That practice reduces the risk of having two unrelated variables both named n.)
(*ii*) Start with smaller numbers, like $C = 100$, $M = 1$, so that your code runs fast while you're debugging it. When it looks good, then substitute the requested values of those parameters.
(*iii*) Compared to our earlier examples, a key feature here is that the mutant numbers in each generation are random but *not independent* of the preceding generations: The number of remaining wild-type (candidates for mutation) is itself a random variable. One way to proceed is to use three nested loops in your code: The outermost loop repeats the code for each simulated experiment, from 1 to M. The middle loop involves which culture in a particular experiment is being simulated, from 1 to C. The innermost loop steps through the g doublings of a particular experiment, in a particular culture.[22]
(*iv*) Remember that in each doubling step the only candidates for mutation are the remaining unmutated cells. So keep a running tally of that number.]



The m values are higher if the first mutant appears at an earlier step, and can almost be fit with the function: $m = 2^{g-i^*}$
See code and values in the .ipynb. (4.16.d)

CELLULAR AND SYSTEMS MODELING - HW2

Student: Shu-Ting Cho (shc167@pitt.edu)

9. PMLS2, problem 4.19 (5 pts)

4.19 **T2** *Skewed distribution*

Suppose that ℓ is drawn from a Poisson distribution. Find the expectation $\langle(\ell - \langle\ell\rangle)^3\rangle$, which depends on μ. Compare your answer with the case of a symmetric distribution, and suggest an interpretation of this statistic.

The expectation $\langle(l - \langle l\rangle)^3\rangle = \mu$

$$\langle(\ell-\langle\ell\rangle)^3\rangle = \langle \ell^3 - \langle\ell\rangle^3 - 3\ell^2\langle\ell\rangle + 3\ell\langle\ell\rangle^2\rangle \qquad * \langle\ell\rangle = \mu$$

$$= \langle\ell^3\rangle - \langle\ell\rangle^3 - 3\langle\ell^2\rangle\langle\ell\rangle + 3\langle\ell\rangle\langle\ell\rangle^2 \qquad * \text{var}\,\ell = \mu$$

$$= \langle\ell^3\rangle - \mu^3 + 3\mu(\langle\ell\rangle^2 - \langle\ell^2\rangle) \quad \hookrightarrow \mu$$

$$= \langle\ell^3\rangle - \mu^3 - 3\mu^2 \quad — ①$$

From Moment Generating Function of Poisson Distribution:

$M_x(t) = e^{\mu(e^t-1)}$ and $\langle\ell^3\rangle = E(\ell^3) = M'''(0)$

$M'(t) = \mu e^{\mu(e^t-1)+t}$

$M''(t) = \mu(\mu e^t + 1)e^{\mu(e^t-1)+t}$

$M'''(t) = \mu^2 e^{\mu(e^t-1)+2t} + \mu(\mu e^t+1)^2 e^{\mu(e^t-1)+t}$

$\langle\ell^3\rangle = M_x'''(0)$

$= \mu^2 e^{\mu(e^0-1)+2\times0} + \mu(\mu e^0+1)^2 e^{\mu(e^0-1)+0}$

$= \mu^2 + \mu(\mu+1)^2 = \mu^3 + 3\mu^2 + \mu \quad — ②$

put ② in ①

$\langle(\ell-\langle\ell\rangle)^3\rangle = \mu^3 + 3\mu^2 + \mu - \mu^3 - 3\mu^2 = \mu$ ※

This term $\langle(l - \langle l\rangle)^3\rangle$ represents the skewness of a distribution.
In the case of symmetric distribution (e.g., normal distribution), the skewness is 0 (not skewed).