GRAD C23: Mathematics for Data Science
Hertie School
Fall 2023
Prof. Magazinnik

Problem Set 2

Submitted by: Giulia Maria Petrilli
Group: Benjamin Adoba Ayida, Finn Krueger

1. Poisson distribution to model rare climate events: flooding in Bangladesh

   a. To gain the insight of weather floods in Bangladesh are increasing due to climate change, we begin to write down the likelihood of floods for the first quarter of the century for some unknown λ. Every time span Xt has Poisson distribution with parameter λ. The following is the likelihood of a flooding happening over one time-span.

$$P(X = x^1; \lambda) = \frac{e^{-\lambda}\lambda^{x_i}}{xi!}, xi \in \{1, 2, 3, 4, 5\} \tag{1}$$

Where:
$P(X = x)$ is the probability of observing x floods.
e is the base of the natural logarithm (approximately 2.71828).
λ is the average rate of floods occurrences.
k is a non-negative integer representing the number of floods.

The likelihood of this series of events is the product of all respective marginals. Hence, the likelihood of it happening over all the time spans, assuming that floods are independent from each other, is:

$$\mathcal{L}(x_i; \lambda) = \prod_{xi=1}^{5} P(xi; \lambda) \tag{2}$$

This can be expanded as follows, accounting for the fact that the probability of the events given a unknown sample mean is a Poisson distribution:

$$= \prod_{xi=1}^{5} \frac{e^{-\lambda}\lambda^{x_i}}{xi!} \tag{3}$$

Which is equal to writing it like this:

$$= \frac{\prod_{xi=1}^{5} e^{-\lambda} \prod_{xi=1}^{5} \lambda^{x_i}}{\prod_{xi=1}^{5} xi!} \tag{4}$$

The product of $e^{-\lambda}$ for each time period is equal to $e^{-5\lambda}$, and the product of each sample mean λ can be rewritten as λ at the power of the sum. Writing $\sum_{x_i=1}^{5} x_i$ is the same as writing $\lambda^{x1+x2+x3+x4+x5}$.

---

[1]Writing x_i is equivalent to writing x_1, x_2, x_3, x_4, x_5. Just for clarity purposes, $x_1 = 2000 - 2004$, $x_2 = 2005 - 2009$ and so on.

This is the likelihood of a series of floods over the first quarter of the century for some unknown sample mean λ:

$$\mathcal{L}(x_i; \lambda) = \frac{e^{-5\lambda} \lambda^{\sum_{x_i=1}^{5} x_i}}{\prod_{x_i=1}^{5} x_i!} \tag{5}$$

The simplification process for conducted so that the λ becomes easier to extract.

b. Computing the log likelihood for equation 5 improves numerical stability, simplifies mathematical calculations, and makes it more convenient for optimization. Just as a general note, it is good to remember that the choice to work with the log-likelihood is for computational convenience.

$$\mathcal{L}(x_i; \lambda) = \log \left[\frac{e^{-5\lambda} \lambda^{\sum_{x_i=1}^{5} x_i}}{\prod_{x_i=1}^{5} x_i!} \right] \tag{6}$$

$$= \log \left[e^{-5\lambda} \right] + \log \left[\lambda^{\sum_{x_i=1}^{5} x_i} \right] + \log \left[\frac{1}{\prod_{x_i=1}^{5} x_i!} \right] \tag{7}$$

The last one is a constant in respect to L so we can just drop it.

$$= \log \left[e^{-5\lambda} \right] + \log \left[\lambda^{\sum_{x_i=1}^{5} x_i} \right] \tag{8}$$

This is the final log likelihood:

$$= -5\lambda + \sum_{x_i=1}^{5} \log \lambda x_i \tag{9}$$

c. Now, let's maximize the log-likelihood from part (b) to derive an MLE. Maximizing here entails taking the derivative and setting equal to 0. We set equal to 0 because it is where the derivative, or rate of change, is null, so at the maximum [2]

$$f'(x) = -5 + \frac{1}{\lambda} \sum_{x_i=1}^{5} x_i \tag{10}$$

$$0 = -5 + \frac{1}{\lambda} \sum_{x_i=1}^{5} x_i \tag{11}$$

$$5 = \frac{1}{\lambda} \sum_{x_i=1}^{5} x_i \tag{12}$$

To get λ on the right, we multiply it on both sides

$$5\lambda = \sum_{x_i=1}^{5} x_i \tag{13}$$

Now we divide by 5 to isolate λ

$$\lambda = \frac{\sum_{x_i=1}^{5} x_i}{5} \tag{14}$$

[2] Can also be used to find minimum, but we are not concerned with it right now

d. From the data provided in the problem we see that

$$\sum_{xi=1}^{5} = 7 \tag{15}$$

Subbing it into equation (14), we get this sample mean. It is important to note that the floods could potentially happen under other lambdas, but it is unlikely.

$$\lambda = \frac{7}{5} = 1.4 \tag{16}$$

e. There were 18 major flooding events recorded in Bangladesh over the 20th century.
To understand weather the floods are increasing, let's compare the two sample means. First, we need the proper data points.
If a quarter of a century is 5 time periods of 5 years (that is, 25 years) then a century is 20 time periods of 5 years.
We know that, in the first quarter of the 21st century, floodings are happening under the sample mean of 1.4 mean. How many floods would the 20th century have had if the sample mean were the same? We can easily calculate that by multiplying the time periods of the century, 20, with the current proportion, 1.4, and we derive that 20 $\cdot$ 1.4 = 28.
The number here is much bigger than 18! To be exact, 28 is approximately 55.56% bigger than 18. If the mean λ would have been the same in the 20th century, we would have had way more floods that the ones we observed, 18.
The opposite also holds: calculating the mean probability of the flooding happened in the 20th century, 18/20, yields a mean of 0.9. This is much lower than the current one!

Sample mean for the first quarter of the 21st century:

$$\lambda = \frac{7}{5} \to 1.4 \tag{17}$$

Sample mean for the 20th century:

$$\lambda = \frac{18}{20} \to 0.9 \tag{18}$$

The next comparison clarifies that the sample mean for the 21st century is much higher than the one for the 20th century. Floods are increasing.

$$0.9 < 1.4 \tag{19}$$

2. The second question is on Taylor set approximation.

a. Part a, specifically, requires to find the CDF of the random variable X with the following PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{20}$$

After a few attempts, one quickly realizes that $e^{-\frac{1}{2}x^2}$ has no elementary anti derivative. More specifically, the inability to find an elementary antiderivative for $e^{-\frac{1}{2}x^2}$ does not mean that the integral is impossible to compute, but it typically involves special functions or numerical methods to approximate the integral. This is because, mathematically, $\frac{d}{dx}(e^x) = e^x$. e^x is an exponential function with a base equal to e.

b. Free points!

c. To prepare for a second-order Taylor series approximation to $f(x)$ at the point $x = 1$, I write down here $f'(x)$ and $f''(x)$.
For the derivative of e, we need to consider that $\frac{d}{dx}[e^u] : e^u \cdot u'$.

Let's begin by putting the constant aside for now

$$f'(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} \left[e^{-\frac{1}{2}x^2} \right] \tag{21}$$

Here, we recognize the same pattern mentioned above, where $\frac{d}{dx}[e^u] : e^u \cdot u'$.

$$f'(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} \left[e^{-\frac{1}{2}x^2} \right] = \frac{d}{dx}[e^u] : e^u \cdot u' \tag{22}$$

We compute this and get:

$$f'(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} e^{-\frac{1}{2}x^2} \cdot -\frac{1}{2} \cdot 2x^{2-1} \tag{23}$$

Further simplification steps :

$$f'(x) = e^{-\frac{1}{2}x^2} \cdot -x \tag{24}$$

Multiplication of -x and $e^{-\frac{1}{2}x^2}$:

$$f'(x) = -xe^{-\frac{1}{2}x^2} \tag{25}$$

Putting the costant back.

$$f'(x) = \frac{1}{\sqrt{2\pi}} \cdot -xe^{-\frac{1}{2}x^2} \tag{26}$$

The final answer for $f'(x)$ is;

$$f'(x) = -\frac{1}{\sqrt{2\pi}} \cdot xe^{-\frac{1}{2}x^2} \tag{27}$$

Let's begin the same process for $f''(x)$ as well,

$$f''(x) = -\frac{1}{\sqrt{2\pi}} \cdot xe^{-\frac{1}{2}x^2} \tag{28}$$

We use the product rule.

$$f''(x) = \left(-\frac{1}{\sqrt{2\pi}} x \right) \cdot \frac{\mathrm{d}}{\mathrm{d}x} \left(e^{-\frac{1}{2}x^2} \right) + \frac{\mathrm{d}}{\mathrm{d}x} \left(-\frac{1}{\sqrt{2\pi}} x \right) \cdot \left(e^{-\frac{1}{2}x^2} \right) \tag{29}$$

Differentiating where needed:

$$= \left(-\frac{1}{\sqrt{2\pi}} x \right) \cdot \left(e^{-\frac{1}{2}x^2} \cdot \frac{\mathrm{d}}{\mathrm{d}x} \left(-\frac{1}{2} x^2 \right) \right) + \left(-\frac{1}{\sqrt{2\pi}} \right) \cdot \left(e^{-\frac{1}{2}x^2} \right) \tag{30}$$

$$= \left(-\frac{1}{\sqrt{2\pi}} x \right) \cdot \left(-xe^{-\frac{1}{2}x^2} \right) - \left(\frac{1}{\sqrt{2\pi}} \right) \cdot \left(e^{-\frac{1}{2}x^2} \right) \tag{31}$$

$$= \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{1}{2}x^2} - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{32}$$

The final answer for $f''(x)$ is:

$$f''(x) = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \left(x^2 - 1 \right) \tag{33}$$

d. Now I simply sub x = 1 in both

$$f'(1) = -\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}} \tag{34}$$

$$f''(1) = \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} \left(1^2 - 1\right) \tag{35}$$

$$= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} (0) \tag{36}$$

$$f''(1) = 0 \tag{37}$$

e. Here I write down the second-order Taylor polynomial for f(x) at x = 1.
Let us recall that second-order Taylor polynomial for a function $f(x)$ centered at a is given by:

$$T_2(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 \tag{38}$$

Having clarified that, this is the second order polynomial as I progressively set my f(x) at x = 1

$$T_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} - \frac{1}{\sqrt{2\pi}} \cdot xe^{-\frac{1}{2}x^2}(x-1) + \frac{1}{2!} \cdot \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} (0)\,(x-a)^2 \tag{39}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} - \frac{1}{\sqrt{2\pi}} \cdot xe^{-\frac{1}{2}x^2}(x-1) \tag{40}$$

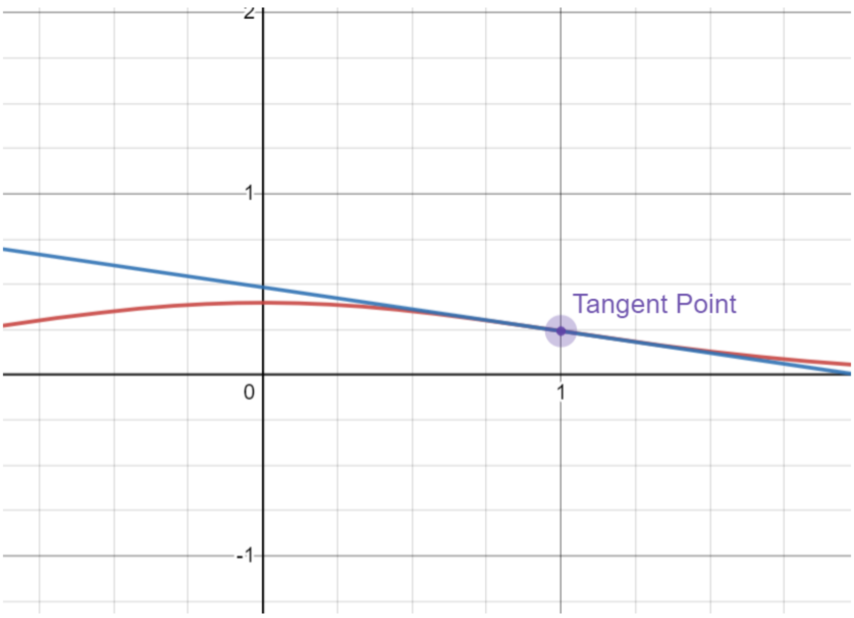
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}1^2} - \frac{1}{\sqrt{2\pi}} \cdot 1e^{-\frac{1}{2}1^2}(x-1) \tag{41}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}} - \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}}(x-1) \tag{42}$$

$$= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} - \frac{e^{-\frac{1}{2}}(x-1)}{\sqrt{2\pi}} \tag{43}$$

$$T_2(1) = \frac{e^{-\frac{1}{2}}(2-x)}{\sqrt{2\pi}} \tag{44}$$

f. The red plot is the PDF while the blue one is the second-order Taylor series approximation found in (e):

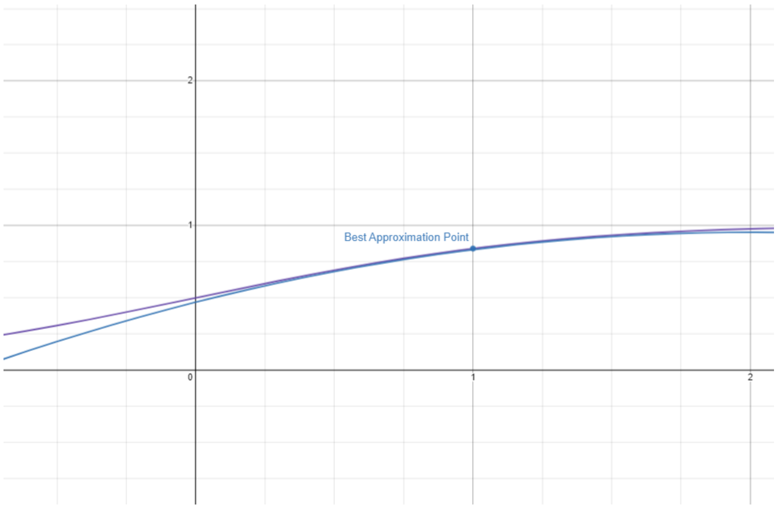


This point represents where the quadratic Taylor polynomial best approximates the behavior of the PDF.

g. Here I compute the integral of the Taylor series approximation to get an approximation to the CDF of X.

$$\int \frac{e^{-1/2}(2-x)}{\sqrt{2\pi}} dx = -\frac{\frac{x^2}{2} - 2x}{\sqrt{2e\pi}} + \text{ C} \tag{45}$$

h. The purple plot is the CDF while the blue one is the integral second-order Taylor series approximation :



C can take any value. Playing around with it, I see that the CDF is the closest to the integral of the second-order Taylor series approximation when C = 0.5

At the labeled point, x = 1 and y = $(1, \text{normaldist}\,(0,1)\,.\,\text{cdf}\,(1))$, the integral of the Taylor polynomial best approximates the CDF. There is no tangent in this case, but there is a close approximation.