`PHSX815_Project3`: Determining Rate Parameter of Exponential Distribution

Vinayak Jha

April 2023

1 Introduction

While performing an experiment, we often have a parameterized model of the results of the experiments. The goal of the experiment in this case is to find the parameters of the model that best fit the measurements. For example, if the experiment is measuring the time between raindrops falling on a unit surface area during rain, then our parameterized model of the time could be an exponential distribution with some rate parameter λ as given below.

$$p(t) = \lambda e^{-\lambda t}$$

The goal here, and of this project, is to determine the rate parameter λ from a given set of measurements of time T. It is also beneficial to provide a lower and upper bound on the rate parameter estimated since the accuracy in estimating λ will be dependent on the number of measurements. Hence, another central question that this project seeks to answer is how the number of measurements per experiment affects the accuracy in estimating λ.

2 Algorithm Design

2.1 Maximum Likelihood Estimation

Consider that on a rainy day, an observer conducts one experiment where N_{meas} measurements of the time between subsequent raindrops falling are recorded. Let this experiment be denoted by the set T. In this project, the experiment set T is constructed by taking N_{meas} samples out of an exponential distribution with rate parameter $\lambda = 2.6$. The goal is to estimate this true rate parameter. The set T is represented as:

$$T = \{t_1, t_2, t_3 \dots t_{N_{meas}}\}$$

Given the set T, we now want to determine the most probable value of the rate parameter λ. Hence, we are interested in what value of λ maximizes $P(\lambda|T)$, which is the posterior probability of λ given the measurements T. Using Bayes Theorem,

$$P(\lambda|T) = \frac{P(T|\lambda)P(\lambda)}{P(T)}$$

Here, $P(T)$ refers to the "evidence" of the observed set, and $P(\lambda)$ is the prior probability distribution of λ. Since $P(T)$ does not change with λ, if we make an additional assumption that all values of λ are equally likely, then we can observe that

$$P(\lambda|T) \propto P(T|\lambda) = P(\{t_1, t_2, t_3 \dots t_{N_{meas}}\}|\lambda)$$

Assuming that time measurements are independent, we can write the above equation as

$$P(\lambda|T) \propto P(\{t_1, t_2, t_3 \dots t_{N_{meas}}\}|\lambda) = \prod_{i=1}^{N_{meas}} P(t_i|\lambda)$$

$$\implies P(\lambda|T) \propto \prod_{i=1}^{N_{meas}} \lambda e^{-\lambda t_i}$$

We need to find $\arg\max_\lambda P(\lambda|T)$ which is equivalent to $\arg\min_\lambda -log(P(\lambda|T))$. Now, we can simplify the product of probabilities as the sum of log probabilities as shown below

$$\arg\min_\lambda -log(P(\lambda|T)) = \arg\min_\lambda \left(-\sum_{i=1}^{N_{meas}} log(\lambda e^{-\lambda t_i}) \right)$$

$$\implies \arg\min_\lambda \left[-N_{meas} log(\lambda) + \lambda \left(\sum_{i=1}^{N_{meas}} t_i \right) \right]$$

Let $f(\lambda)$ denote the functional part of the above equation. Thus,

$$f(\lambda) = -N_{meas} log(\lambda) + \lambda \left(\sum_{i=1}^{N_{meas}} t_i \right)$$

Hence, we need to find λ which minimizes $f(\lambda)$. This is done numerically in this project. This method is called maximum likelihood estimation, and the value of λ obtained is called the maximum likelihood estimate and will be denoted by $\hat{\lambda}$.

2.1.1 Proof of Global Minima of $f(\lambda)$

This section shows that $f(\lambda)$ can have just one global minima. However, to keep it interesting, I don't calculate the point of global minima explicitly, as that is done numerically. Taking the derivative of $f(\lambda)$,

$$f'(\lambda) = -\frac{N_{meas}}{\lambda} + \left(\sum_{i=1}^{N_{meas}} t_i \right)$$

Denoting $\left(\sum_{i=1}^{N_{meas}} t_i \right)$ as S,

$$f'(\lambda) = -\frac{N_{meas}}{\lambda} + S$$

Let's first assume that $S > 0$. Since S is a summation of time intervals, this is a reasonable assumption. The case $S = 0$ will be covered later. Additionally, also assume that $\lambda > 0$. Examining the limits $\lambda \to 0$ and $\lambda \to \infty$,

$$\lim_{\lambda \to \infty} f'(\lambda) = S$$

$$\lim_{\lambda \to 0} f'(\lambda) = -\infty$$

Since $S > 0$, it is guaranteed that $f'(\lambda) = 0$ at some $0 < \lambda < \infty$. Now, we need to show that $f'(\lambda)$ will have just one root (that is the only one λ such that $f'(\lambda) = 0$). For this, it is sufficient to show that $f'(\lambda)$ has no local maxima or local minima. This can be shown by considering $f''(\lambda)$.

$$f''(\lambda) = \frac{N}{\lambda^2}$$

Since $f''(\lambda) > 0$ for $0 < \lambda < \infty$, it is guaranteed that $f'(\lambda)$ does not have a local maxima or minima. This also implies that there is no point of inflection since the second derivative is always positive. Hence, $f'(\lambda) = 0$ will have just one solution. Additionally, $f''(\lambda) > 0$ implies that λ is a local minima and not a local maxima of $f(\lambda)$. Since there is no point of inflection and just one local minima, it is additionally guaranteed that this local minima will be the global minima (provided this local minima doesn't occur at the bounds).

In this proof, it was assumed that $\lambda > 0$ and $S > 0$.

The first assumption (λ > 0) is easy to justify since if $\lambda = 0$, then the probability distribution is no longer meaningful.

For the second assumption, note that $S = 0$ as $N_{meas} \to \infty$ implies that $t_i = 0$ for $1 \leq i \leq N_{meas}$. Hence, if the observer notices that $t_i = 0$ for infinitely many measurements, then $\lambda \to \infty$ almost surely, hence λ also does not have a physical interpretation.

2.2 1σ Bounds

The maximum likelihood estimate of λ from the previous subsection (2.1) provides the most probable value of λ for that experiment. Hence, it does not determine the true rate parameter absolutely. We are now interested in what values of λ are "acceptable" for the true rate parameter which is denoted by 1σ range of λ given the maximum likelihood estimate $\hat{\lambda}$. To find the 1σ range of λ, we find all the values of λ such that:

$$g(\lambda) = log\left(\frac{P(\lambda|T)}{P(\hat{\lambda}|T)}\right) >= -\frac{1}{2}$$

From the previous subsection, it was noted that $f(\lambda) \propto -log(P(\lambda|T))$ has just one global minima. Since $g(\lambda) \propto -f(\lambda)$, it is guaranteed that $g(\lambda)$ will have just one maxima at $\hat{\lambda}$ and $g(\lambda)$ will always decrease as we go in either direction of $\hat{\lambda}$. Hence, we can find the lower and upper bounds of λ as follows:

1. **Lower Bound on $\boldsymbol{\lambda}$**: Set $\lambda_l = \hat{\lambda}$. Update $\lambda_l = \lambda_l - \epsilon$ where ϵ is a small step size. if $g(\lambda_l) < -\frac{1}{2}$, stop. Otherwise, keep updating $\lambda_l = \lambda_l - \epsilon$ until $g(\lambda_l) < -\frac{1}{2}$. The final λ_l is the lower bound of 1σ bound on λ's most likelihood estimate.

2. **Upper Bound on $\boldsymbol{\lambda}$**: Set $\lambda_u = \hat{\lambda}$. Update $\lambda_u = \lambda_u + \epsilon$ where ϵ is a small step size. if $g(\lambda_u) < -\frac{1}{2}$, stop. Otherwise, keep updating $\lambda_u = \lambda_u + \epsilon$ until $g(\lambda_u) < -\frac{1}{2}$. The final λ_u is the upper bound of 1σ bound on λ's most likelihood estimate.

The λ_l and λ_u obtained from this algorithm provide lower and upper bounds on our 1σ estimate of λ

2.3 Multiple Experiments

Note that in subsection 2.1, the maximum likelihood estimate of $\hat{\lambda}$ was for a single experiment of the set T composed of N_{meas} measurements. We can repeat the experiment multiple times and we will get a new estimate of $\hat{\lambda}$ for every single experiment. We can plot a histogram of $\hat{\lambda}$ obtained from each experiment to obtain the probability distribution of $\hat{\lambda}$. From this probability distribution, the expectation of $\hat{\lambda}$ is calculated. This expectation is taken to be the estimated $\hat{\lambda}$ from this set of experiments which is then compared to the true rate parameter. In other words, the expectation of the probability distribution of $\hat{\lambda}$ is calculated and then compared against the true rate parameter to provide the error in our estimation of the true rate parameter. Using the probability distribution of $\hat{\lambda}$, we can also calculate the standard deviation which provides a qualitative indicator of the error in estimating the rate parameter.

2.4 Multiple Estimates of $\hat{\lambda}$ over varying number of measurements

It is expected that as the number of measurements increases per experiment, the standard deviation of the resulting probability distribution of $\hat{\lambda}$ will decrease. This is because as the number of measurements increases, we expect that $\hat{\lambda}$ will more closely resemble the true rate parameter. This implies $\hat{\lambda}$ will lie closer to each other, and to the true rate parameter. Referring to subsection 2.3, it is hypothesized that the expectation of the probability distribution of $\hat{\lambda}$ for a given number of measurements will become closer to the true rate parameter as the number of measurements increases.

3 Analysis

3.1 Measuring rate parameter from single experiment

To simulate an experiment, $N_{meas} = 10$ samples are drawn from an exponential distribution with rate parameter $\lambda = 2.6$. The samples drawn form the set T defined in the previous sections, and consequently, the objective function $f(\lambda)$ (from subsection 2.1) is completely determined. To minimize $f(\lambda)$, I use Scipy's default minimization method Broyden-Fletcher-Goldfarb-Shanno (BFGS). The rate parameter predicted by the algorithm is $\hat{\lambda} = 2.17891$. Using the algorithm described in subsection (2.2) the lower and upper bounds are $\lambda_l = 1.55891$ and $\lambda_u = 2.94891$. Note that this is the estimate of λ from a single experiment.

3.2 Distribution of rate parameter over multiple experiments

As described before, multiple experiments can be conducted, and multiple estimates of $\hat{\lambda}$ can be obtained. Figure 1 shows the histogram of the $\hat{\lambda}$ obtained across 50000 experiments. The number of measurements per experiment (N_{meas}) is 10. The expected value of $\hat{\lambda}$ (represented by the black vertical line) from this distribution is 2.874 which is close to the true rate parameter ($\lambda = 2.6$) and the standard deviation is 0.981.

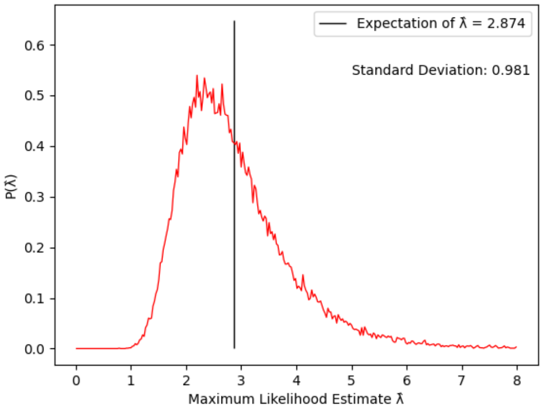

Figure 1: The figure shows the probability distribution of the maximum likelihood estimate of the rate parameter ($\hat{\lambda}$) obtained from 50000 experiments each with 10 measurements. The expectation of $\hat{\lambda}$ obtained is 2.874 (represented by the black vertical line) and the standard deviation is 0.981.

3.3 Multiple estimates of $\hat{\lambda}$ over varying number of measurements

As described in section 2.4, it is expected that as the number of measurements per experiment increases, the resulting distribution of $\hat{\lambda}$ over multiple experiments (for example from section 3.2) will get narrower (standard deviation will decrease) and our estimate of $\hat{\lambda}$ will get more accurate (expectation of $\hat{\lambda}$ will get closer to true rate parameter). Figures 2, 3, 4, and 5 show the histogram of $\hat{\lambda}$ obtained across 50000 experiments with a varying number of measurements - 15, 30, 45, and 60 respectively. As it can be observed, the histograms get narrower as the number of measurements increases. Additionally, it can be seen that the expected values of $\hat{\lambda}$ become closer to the true rate parameter (which is 2.6).

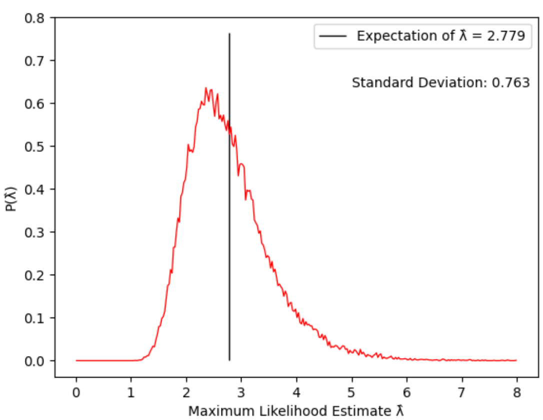

Figure 2: The figure shows the probability distribution of the maximum likelihood estimate of the rate parameter ($\hat{\lambda}$) obtained from 50000 experiments each with 15 measurements. The expectation of $\hat{\lambda}$ obtained is 2.779 (represented by the black vertical line) and the standard deviation is 0.763.

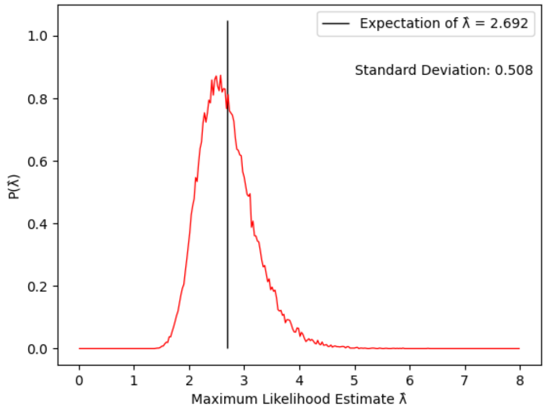

Figure 3: The figure shows the probability distribution of the maximum likelihood estimate of the rate parameter ($\hat{\lambda}$) obtained from 50000 experiments each with 30 measurements. The expectation of $\hat{\lambda}$ obtained is 2.692 (represented by the black vertical line) and the standard deviation is 0.508.

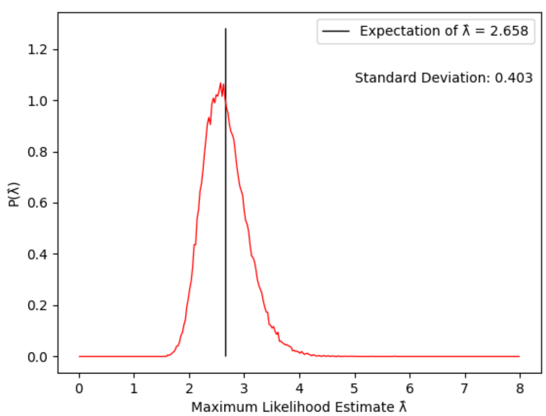

Figure 4: The figure shows the probability distribution of the maximum likelihood estimate of the rate parameter ($\hat{\lambda}$) obtained from 50000 experiments each with 45 measurements. The expectation of $\hat{\lambda}$ obtained is 2.658 (represented by the black vertical line) and the standard deviation is 0.403.

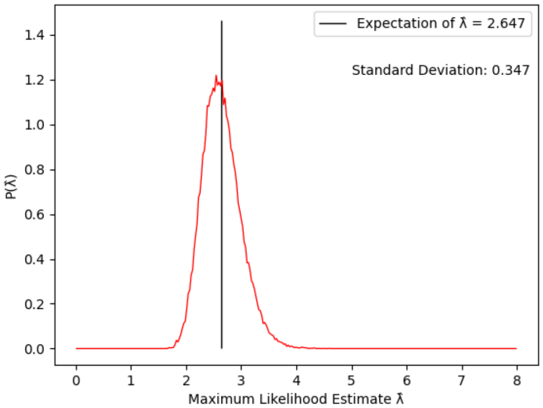

Figure 5: The figure shows the probability distribution of the maximum likelihood estimate of the rate parameter ($\hat{\lambda}$) obtained from 50000 experiments each with 60 measurements. The expectation of $\hat{\lambda}$ obtained is 2.647 (represented by the black vertical line) and the standard deviation is 0.347.

More concretely, Figures 6 and 7 show expected values of $\hat{\lambda}$ and standard deviation of $\hat{\lambda}$ as the number of measurements increases. From Figure 6 it can be seen that the expected value of $\hat{\lambda}$ approaches the true rate parameter (2.6) as the measurements per experiment increase. From Figure 7, the standard deviation decreases as the number of measurements per experiment increases suggesting that the estimates of rate parameters also become more precise with more measurements per experiment.

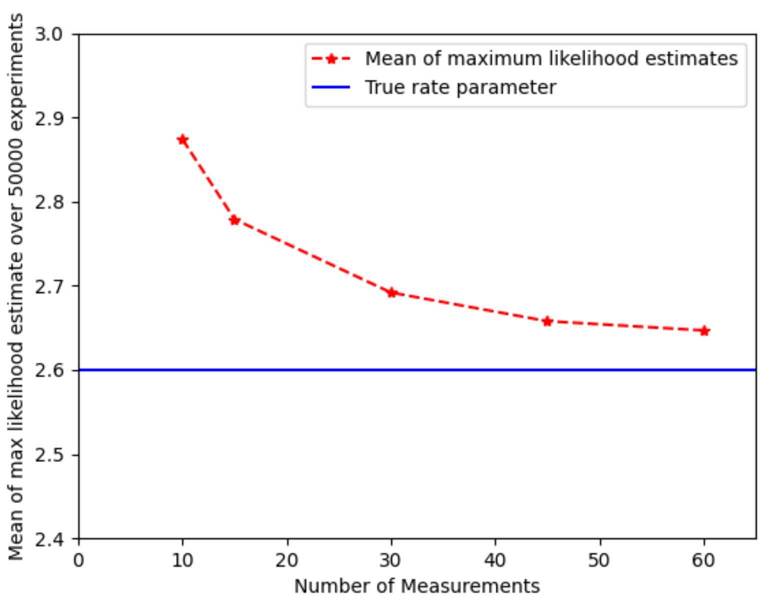

Figure 6: Figure shows the dependence of the expectation (mean) on the number of measurements. As the number of measurements increases, the mean of $\hat{\lambda}$ over 50000 experiments approaches the true rate parameter (2.6)

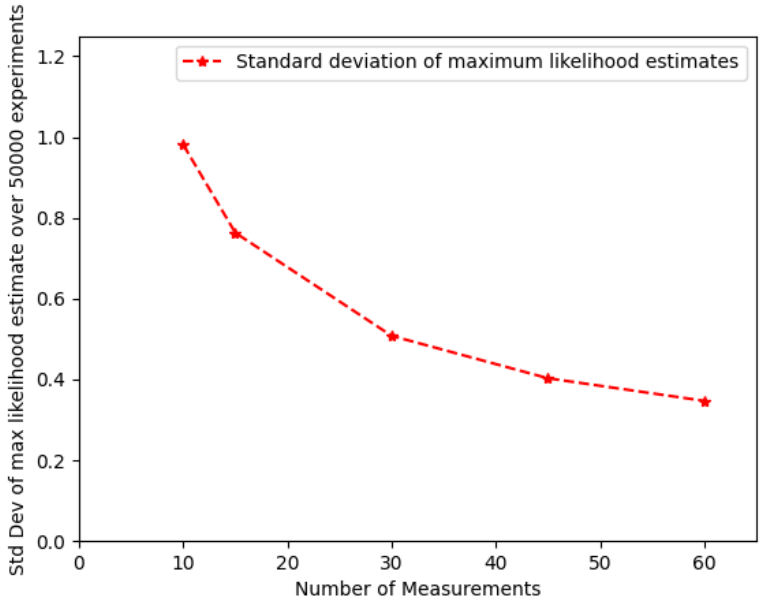

Figure 7: Figure shows the dependence of the standard deviation on the number of measurements. As the number of measurements increases, the standard deviation of $\hat{\lambda}$ over 50000 experiments approaches the 0

4 Conclusion

In this project, the time between raindrops during rain is modeled as an exponential distribution with rate parameter $\lambda = 2.6$. The goal is to best estimate λ from a set of measurements of the time.

Given a single experiment comprised of N_{meas} experiments, the rate parameter of the exponential distribution was estimated using maximum likelihood estimation. Additionally, the lower and upper bounds of this maximum likelihood estimate were also calculated.

Given a fixed number of measurements per experiment, multiple experiments were conducted resulting in a probability distribution of the maximum likelihood estimates of the rate parameter. It was observed that the expectation of this probability distribution approaches the true rate parameter as the number of measurements per experiment increases. To answer the question brought up in Section 1, the error in estimating the true rate parameter decreases with more measurements per experiment. Additionally, the standard deviation of this probability distribution also decreases with more measurements. Hence, the estimates become more precise, and less variant from one another.