Homework 2 Responses

Chris Crabtree

Oct. 14, 2021

1. (a) **Prompt**: *Show that* $\mathbf{s}(y_{1:n}) = (\bar{y}, s^2)$ *is a sufficient statistic for* $y_{1:n} = y_1, y_2, ..., y_n \overset{iid}{\sim} Normal(\mu, \sigma^2)$.

I will use the Factorization theorem to solve this problem. From the lecture notes, the theorem can be stated as:

> A statistic $s(x_{1:n})$ is sufficient *iff* the joint conditional pdf, $f(x_{1:n}|\theta)$ can be factored in the following manner:
> $$f(x_{1:n}|\theta) = g(\mathbf{s}|\theta)h(x_{1:n})$$

First we must state the joint conditional distribution $f(y_{1:n})$:

$$f(y_{1:n}|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(\frac{-1}{2}\Big(\frac{x_i - \mu}{\sigma}\Big)^2\Big)$$

From here we can add and subtract $\bar{y}$:

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(\frac{-1}{2}\Big(\frac{y_i - \mu}{\sigma}\Big)^2\Big) = \Big(\frac{1}{\sqrt{2\pi}\sigma}\Big) \exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i + \bar{y} - \bar{y} - \mu}{\sigma}\Big)^2\Big)$$
$$= \Big(\frac{1}{\sqrt{2\pi}\sigma}\Big) \exp\Big(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}((y_i + \bar{y})^2 - 2(y_i - \bar{y})(\bar{y} - \mu) + (\bar{y} - \mu)^2)\Big)$$

Note that $\sum_{i=1}^{n} 2(y_i - \bar{y})(\bar{y} - \mu) = 2(\bar{y} - \mu)\sum_{i=1}^{n}(y_i - \bar{y}) = 0$
Therfore we have:

$$f(y_{1:n}|\theta) = \Big(\frac{1}{\sqrt{2\pi}\sigma}\Big) \exp\Big(\frac{1}{2\sigma^2}\sum_{i=1}^{n}((y_i + \bar{y})^2 - 0 + (\bar{y} - \mu)^2)\Big)$$
$$= \Big(\frac{1}{\sqrt{2\pi}\sigma}\Big) \exp\Big(\frac{(n-1)\cdot s^2}{2\sigma^2}\sum_{i=1}^{n}(\bar{y} - \mu)^2\Big)$$

With that we no longer have $y_{1:n}$ in the expression. Because of this we can trivially let $h(y_{1:n})$ trivially equal 1, and $g(\mathbf{s}|\theta)$ equal the above expression. Hence $\mathbf{s}$ is a sufficient statistic for μ and σ^2.

(b) **Prompt**: *Find the observed information matrix for* $y_{1:n} \overset{iid}{\sim} Normal(\mu, \sigma^2)$.
From the lecture notes, the observed information is defined as:

> $$J(\theta) = -\frac{\partial^2 \mathcal{L}(\theta)}{\partial\theta\partial\theta^\intercal} \text{ with } \mathcal{L} = \log(L(\theta)).$$

For this problem, $\log(L(\theta))$ can be written as:

$$\begin{aligned}\log(L(\theta)) &= \log(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(\frac{-1}{2}\Big(\frac{y_i-\mu}{2\sigma}\Big)^2\Big)) \\ &= \log(\Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n \exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)) \end{aligned} \tag{1}$$

where $\theta^\intercal$ is $[\mu, \sigma^2]$.
First we will find the first derivative w.r.t. θ.

$$\frac{\partial \log(L(\theta))}{\partial \theta} = \begin{bmatrix} \frac{\partial \log(L(\theta))}{\partial \mu} \\ \frac{\partial \log(L(\theta))}{\partial \sigma^2} \end{bmatrix}$$

Starting with $\frac{\partial}{\partial \mu}$ we have:

$$\begin{aligned}\frac{\partial \log(L(\theta))}{\partial \mu} &= \frac{1}{L(\theta)}\frac{\partial}{\partial \mu}\Big[\Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n \exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\Big] \\ &= \frac{1}{L(\theta)}\Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n \exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\frac{\partial}{\partial \mu}\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2 \\ &= \frac{1}{L(\theta)}L(\theta)\frac{\partial}{\partial \mu}\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2 \\ &= \frac{\partial}{\partial \mu}\frac{-1}{2\sigma^2}\sum_{i=1}^{n}\big((y_i-\bar{y})^2+(\bar{y}-\mu)^2\big) \\ &= \frac{-1}{2\sigma^2}\frac{\partial}{\partial \mu}n(\bar{y}-\mu)^2 \\ &= \frac{n(\bar{y}-\mu)}{\sigma^2}\end{aligned}$$

For $\frac{\partial}{\partial \sigma^2}$ we will use the the product rule for derivatives:

$$\begin{aligned}\frac{\partial \log(L(\theta))}{\partial \sigma^2} &= \frac{1}{L(\theta)}\frac{\partial}{\partial \sigma^2}\Big[\Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n \exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\Big] \\ &= \frac{1}{L(\theta)}\Big(\frac{\partial}{\partial \sigma^2}\Big[\Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n\Big]\exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big) \\ &\quad + \Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n \frac{\partial}{\partial \sigma^2}\Big[\exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\Big]\Big)\end{aligned} \tag{2}$$

For space efficiency, I will solve $\frac{\partial}{\partial \sigma^2}\Big[\Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n\Big]$ and $\frac{\partial}{\partial \sigma^2}\Big[\exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\Big]\Big)$ separately.
Solving $\frac{\partial}{\partial \sigma^2}\Big[\Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n\Big]$, we have:

$$
\begin{aligned}
\frac{\partial}{\partial\sigma^2}\Big[\Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^n\Big] &= \frac{\partial}{\partial\sigma^2}[(\sqrt{2\pi}\sigma)^{-n}] \\
&= -n(\sqrt{2\pi}\sigma)^{-n-1}\frac{\partial}{\partial\sigma^2}[\sqrt{2\pi}\sigma] \\
&= -n(\sqrt{2\pi}\sigma)^{-n-1}\sqrt{2\pi}\frac{\partial}{\partial\sigma^2}[\sqrt{\sigma^2}] \\
&= -n(\sqrt{2\pi}\sigma)^{-n-1}\sqrt{2\pi}\frac{1}{2}(\sigma^2)^{-1/2} \\
&= -n(\sqrt{2\pi}\sigma)^{-n-1}\sqrt{2\pi}\frac{1}{2\sigma} \\
&= -n(\sqrt{2\pi}\sigma)^{-n}\frac{1}{\sqrt{2\pi}\sigma}\sqrt{2\pi}\frac{1}{2\sigma} \\
&= \frac{-n}{2\sigma^2}\Big(\frac{1}{\sqrt{2\pi}\sigma^2}\Big)^n \qquad (3)
\end{aligned}
$$

Solving $\frac{\partial}{\partial\sigma^2}\Big[\exp\Big(\frac{1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\Big]\Big)$ we have:

$$
\begin{aligned}
\frac{\partial}{\partial\sigma^2}\Big[\exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\Big] &= \exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\frac{\partial}{\partial\sigma^2}\Big[\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big] \\
&= \exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\frac{-1}{2}\frac{\partial}{\partial\sigma^2}\Big[\frac{1}{\sigma^2}\Big]\sum_{i=1}^{n}(y_i-\mu)^2 \\
&= \exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\Big(\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2(\sigma^2)^2}\Big) \qquad (4)
\end{aligned}
$$

Plugging equations 3 and 4 into equation 2 we have:

$$
\begin{aligned}
\frac{\partial\log(L(\theta))}{\partial\sigma^2} &= \frac{1}{L(\theta)}\Big(\Big[\frac{-n}{2\sigma^2}\Big(\frac{1}{\sqrt{2\pi}\sigma^2}\Big)^n\Big]\exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big) \\
&\quad + \Big(\frac{1}{\sqrt{2\pi}\sigma^2}\Big)^n\Big[\exp\Big(\frac{-1}{2}\sum_{i=1}^{n}\Big(\frac{y_i-\mu}{\sigma}\Big)^2\Big)\Big(\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2(\sigma^2)^2}\Big)\Big]\Big) \\
&= \frac{1}{L(\theta)}L(\theta)\Big(\frac{-n}{2\sigma^2} - \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2(\sigma^2)^2}\Big) \\
&= \frac{-n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2(\sigma^2)^2} \qquad (5)
\end{aligned}
$$

(Finally!) This gives

$$
\frac{\partial\log(L(\theta))}{\partial\theta} = \begin{bmatrix} \frac{\partial\log(L(\theta))}{\partial\mu} \\ \frac{\partial\log(L(\theta))}{\partial\sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{n(\bar{y}-\mu)}{\sigma^2} \\ \frac{-n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2(\sigma^2)^2} \end{bmatrix}
$$

From here finding the second partial derivatives is fairly trivial. You simply need to take derivatives w.r.t the μ and σ^2 separately for each element of the above vector. Doing so

gives the full solution:

$$J(\theta) = -\frac{\partial^2 \log(L(\theta))}{\partial\theta\partial\theta^\intercal} = -\begin{bmatrix} \frac{-n\bar{y}}{\sigma^2} & \frac{-n(\bar{y}-\mu)}{(\sigma^2)^2} \\ \frac{-n(\bar{y}-\mu)}{(\sigma^2)^2} & \frac{n}{2(\sigma^2)^2} - \frac{\sum_{i=1}^n (y_i-\mu)^2}{(\sigma^2)^3}\Big) \end{bmatrix}$$
$$= \begin{bmatrix} \frac{n\bar{y}}{\sigma^2} & \frac{n(\bar{y}-\mu)}{(\sigma^2)^2} \\ \frac{n(\bar{y}-\mu)}{(\sigma^2)^2} & \frac{-n}{2(\sigma^2)^2} + \frac{\sum_{i=1}^n (y_i-\mu)^2}{(\sigma^2)^3}\Big) \end{bmatrix}$$

(c) **Prompt**: *Find the expected information matrix for Normal*(μ, σ^2).
This was explained in lecture. The only term from the above matrix that needs to be calculated is the lower right one.

$$\begin{aligned}
\mathbb{E}[\frac{-n}{2\sigma^2} + \frac{\sum_{i=1}^n (y_i-\mu)^2}{(\sigma^2)^3}] &= \frac{-n}{2(\sigma^2)^2} + \frac{1}{(\sigma^2)^3}\mathbb{E}[\sum_{i=1}^n (y_i-\mu)^2] \\
&= \frac{-n}{2(\sigma^2)^2} + \frac{1}{(\sigma^2)^3}\mathbb{E}[\sum_{i=1}^n ((y_i-\bar{y})^2 + (\bar{y}-\mu)^2)] \\
&= \frac{-n}{2(\sigma^2)^2} + \frac{1}{(\sigma^2)^3}\Big(\mathbb{E}[\sum_{i=1}^n (y_i-\bar{y})^2] + n\mathbb{E}[(\bar{y}-\mu)^2]\Big) \\
&= \frac{-n}{2(\sigma^2)^2} + \frac{1}{(\sigma^2)^3}\Big((n-1)\mathbb{E}[\frac{\sum_{i=1}^n (y_i-\bar{y})^2}{(n-1)}] + n\mathbb{E}[(\bar{y}-\mu)^2]\Big) \\
&= \frac{-n}{2(\sigma^2)^2} + \frac{1}{(\sigma^2)^3}\Big((n-1)\sigma^2 + n\Big(\frac{\sigma^2}{n}\Big)\Big) \\
&= \frac{-n}{2(\sigma^2)^2} + \frac{1}{(\sigma^2)^3} n\sigma^2 \\
&= \frac{n}{2(\sigma^2)^2}
\end{aligned}$$

2. (a) **Prompt**: *Show that the minimal sufficient statistic (MSS) for* $Y_{1:n} = Y_1, Y_2, ..., Y_n \overset{iid}{\sim} \mathcal{N}(\mu, c\mu^2)$ *is the same MSS as for* $\mathcal{N}(\mu, \sigma^2)$.
I will use the common strategy of examining the ratio of the likelihoods of two separate statistics to find the MSS.
Let $Z_{1:m}$ be sampled from the same distribution as $Y_{1:n}$. Then we note that in the ratio $Y_{1:n}/Z_{1:m}$ the normalizing constancts will cancel, leaving the ratio of exponentials. This gives

$$\exp\Big(\frac{\sum_{i=1}^n (Y_i-\mu)^2 - \sum_{k=1}^m (Z_k-\mu)^2}{2c\mu^2}\Big)$$

If we expand the squared binomials, clearly the μ^2s will cancel, but we can also combine the cross terms giving:

$$\exp\Big(\frac{\sum_{i=1}^n Y_i^2 - \sum_{k=1}^m Z_k^2 + 2\mu(\sum_{k=1}^m Z_k + \sum_{i=1}^n Y_i)}{2c\mu^2}\Big) \tag{6}$$

Note that equation 6 **this would be the same as for** $\mathcal{N}(\mu, \sigma^2)$, but swapping the second parameter.
From equation 6 it is clear that dependence on the μ and $c\mu^2$ is eliminated *iff* $\sum_{i=1}^n Y_i^2 = \sum_{k=1}^m Z_k^2$ and $\sum_{i=1}^n Y_i = \sum_{k=1}^m Z_k$. Furthermore, since the manipulations to get to equation 6 would be the same as for $\mathcal{N}(\mu, \sigma^2)$, the MSSs are the same.

(b) **Prompt**: *Find the MLE for μ.*
To obtain the MLE we will first write $\mathcal{L}(\theta)$:

$$\begin{aligned}
\mathcal{L}(\theta) &= \log(L(\theta)) \\
&= \log\Big(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi c}\mu} \exp\Big(\frac{(x-m)^2}{2c\mu^2}\Big)\Big) \\
&= -\log(\sqrt{2\pi c}\mu) + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2c\mu^2} \\
&= -\log(\sqrt{2\pi c}\mu) + \sum_{i=1}^{n} \frac{x_i^2 - x_i\mu + \mu^2}{2c\mu^2} \\
&= -\log(\sqrt{2\pi c}\mu) + \sum_{i=1}^{n} \frac{x_i^2}{2c\mu^2} - \frac{\sum_{i=1}^{n} x_i\mu}{2c\mu^2} + \frac{\sum_{i=1}^{n}\mu^2}{2c\mu^2} \\
&= -\log(\sqrt{2\pi c}\mu) + \frac{1}{\mu^2}\sum_{i=1}^{n} \frac{x_i^2}{2c} - \frac{1}{\mu}\frac{\sum_{i=1}^{n} x_i}{2c} + \frac{n}{2c}
\end{aligned}$$

Taking the derivative of the above expression gives:

$$\frac{d}{d\mu}\mathcal{L}(\theta) = -\frac{1}{\mu} - \frac{1}{\mu^3}\frac{\sum_{i=1}^{n} x_i^2}{c} + \frac{1}{\mu^2}\frac{\sum_{i=1}^{n} x_i}{2c}$$

And if we set that equal to zero we get:

$$\begin{aligned}
&-\frac{1}{\mu} - \frac{1}{\mu^3}\frac{\sum_{i=1}^{n} x_i^2}{c} + \frac{1}{\mu^2}\frac{\sum_{i=1}^{n} x_i}{2c} = 0 \\
\Rightarrow &-\mu^2 + \mu\frac{\sum_{i=1}^{n} x_i}{2c} - \frac{\sum_{i=1}^{n} x_i^2}{c} = 0 \\
\Rightarrow &-\mu^2 + \mu\frac{n\bar{x}}{2c} - \frac{n(\tilde{s}^2 - \bar{x}^2)}{c} = 0
\end{aligned}$$

I tried using the quadratic formula for this, but unfortunately I got an ugly expression that did not seem correct.

3. (a) **Prompt**: *Show a simulated sampling distribution of the mean of 20 samples from a Cauchy distribution. Repeat 1000 times for the simulation.*
As can be seen from Figure 1, the sampling distribution does in fact seem to follow a Cauchy distribution.

(b) **Prompt**: *Fit the spring data to a two-parameter Weibull distribution with shape parameter k and scale parameter λ using three optimization techniques: Step-wise Gradient Descent, Stochastic Gradient Descent, and the Newton-Raphson Method.*
Figure 2 shows the optimization trajectory and fit using Step-wise Gradient Descent. I used an initial learning rate of .5 and used a decay schedule of $.999^{step}$ where *step* was the number of optimization steps taken so far. As you can see from the figure, the algorithm

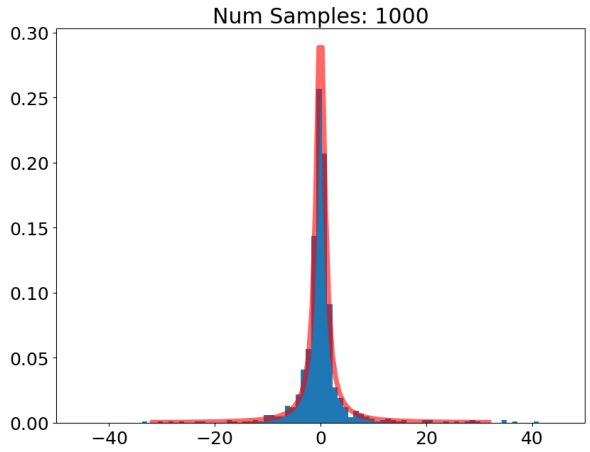

Figure 1: Cauchy Sampling Distribution Simulation

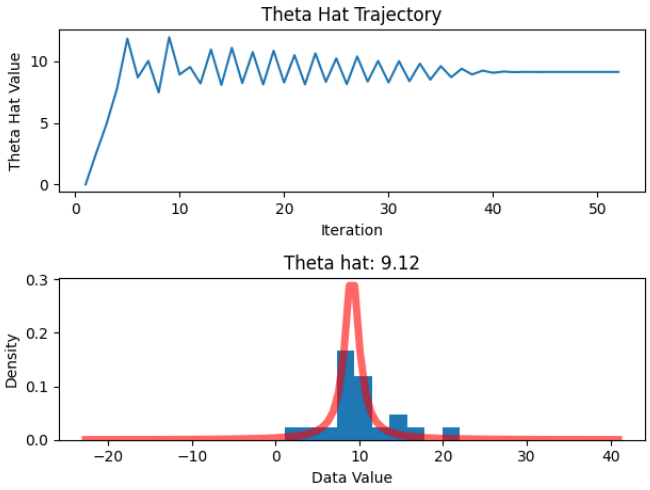

Figure 2: Step-wise Gradient Descent

quickly got very close to the optimal setting, but oscillated around that for about 30 iterations/epochs.

Figure 3 shows the optimization trajectory and fit using the Newton-Raphson Method. This algorithm converged the quickest and did not oscilate which made each update highly targeted.

Figure 4 shows the optimization trajectory and fit using the Stochastic Gradient Descent. This algorithm took longer to converge than the others in terms of update steps (80), but considering that updates were made for each data point instead of the entire dataset (20 data points), the covergence was quite fast in terms of epochs. However, as you can see from the figure it did not obtain a very good fit. This was likely due to the rate of decay. When I flattened the decay rate, the estimate became better, but the algorithm often took well over 1000 iterations to converge.

Also, with the current settings there wer no oscillations. However, by the time the algorithm closed in on the correct solution the learning rate was much smaller (it was around .04 at iteration 70). If the learning rate had been larger at that point I would expect to see the same sort of oscillations present in Step-wise Gradient Descent.

4. **Prompt**: *Use Maximum Likelihood Estimation to fit the data given to a two-parameter Weibull*

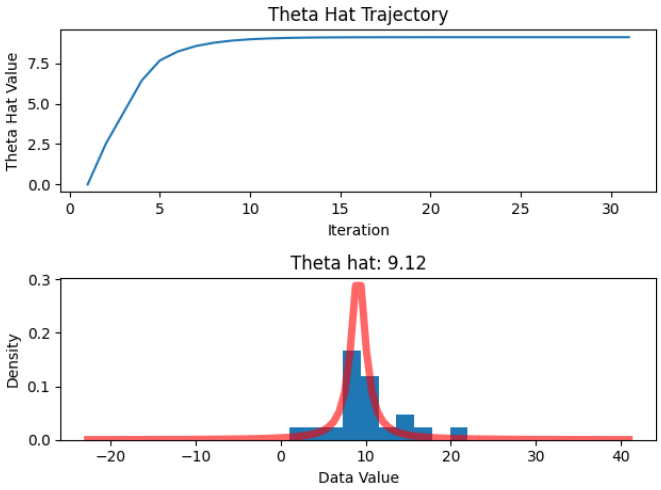

Figure 3: Newton-Raphson

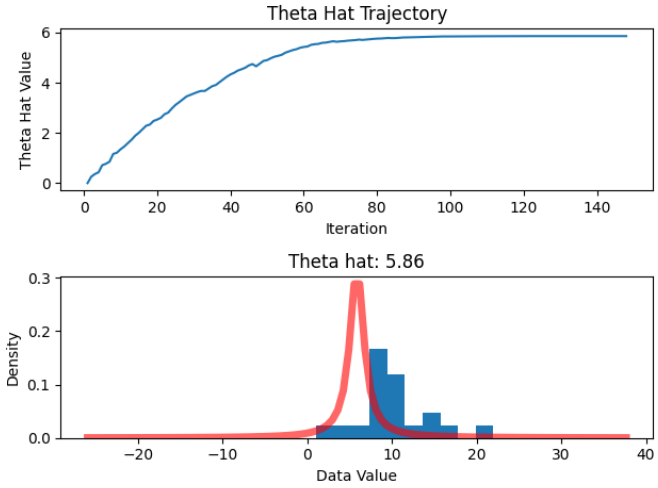

Figure 4: Stochastic Gradient Descent

distribution

The Weibull distribution proved harder to fit than the Cauchy. Not only are there restrictions on the parameter space, but often as the optimization trajectory approached these limits, the gradient would explode and make the estimate unstable. To alleviate this, I used gradient clipping to cap the L2 norm of the gradient to 10. Without this I found the optimization to be unstable and unpredictable.

I used Step-wise Gradient Descent as my algorithm because I did not want to compute the second derivatives. However in retrospect, I could have saved myself a lot of time using the Newton-Raphson Method because the updates would have likely been more stable.

Figure 5 shows the trajectory of both parameter estimates as well as the final fit. It took well over 1000 iterations to achieve a good fit, but it took a lot of experimentation to get this result. I found that the decay rate had to be flattened substantially to $1 - 1e - 10$ when initializing both parameters to 1.0.

5. (a) **Prompt**: *Find the MLE for the parameters of the normal.*
The log likelihood of $\mathcal{N}((,\mu),\sigma^2)$ can trivially be shown to be:

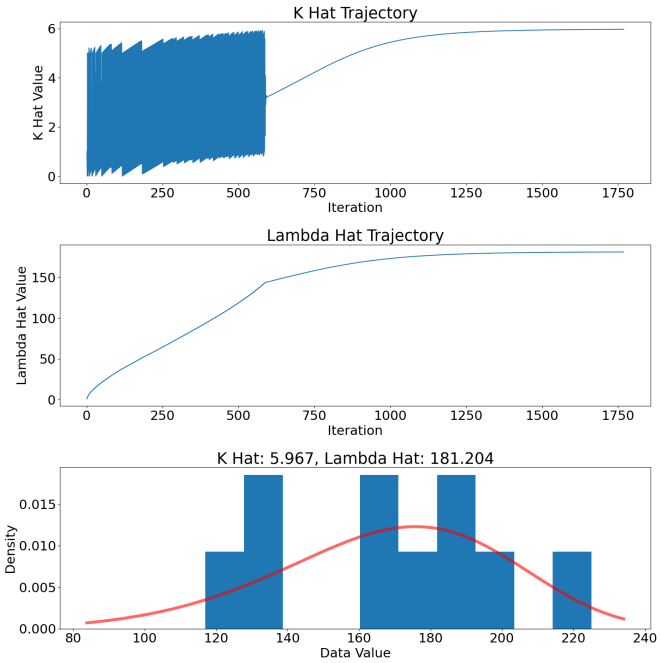

Figure 5: Weibull Parameter Estimation

$$-n/2\log(2\pi) - n\log(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Starting with μ we only need to maximimize $\sum_{i=1}^{n}(x_i-\mu)^2$. This can be obtained easily:

$$\frac{d}{d\mu}\sum_{i=1}^{n}(x_i-\mu)^2 = -2\sum_{i=1}^{n}(x_i-\mu)$$

Setting this equal to zero gives:

$$\begin{aligned} -2\sum_{i=1}^{n}(x_i-\mu) &= 0 \\ \sum_{i=1}^{n} x_i &= n\mu \\ \bar{x} &= \mu \end{aligned}$$

Now for $sigma^2$ we only need to maximize $-n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$. The derivative of this is:

$$-n/\sigma - \frac{1}{2\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2$$

Setting that equal to zero gives:

$$-n/\sigma + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - \mu)^2 = 0$$

$$\sum_{i=1}^{n}(x_i - \mu)^2 = n\sigma^2$$

$$\sigma^2 = \tilde{s}^2$$

(b) **Prompt**: *Show that the MLE for σ^2 is inconsistent and adjust the MLE for σ^2 to make it consistent.*

It was already shown several times in lecture that $\tilde{s}^2$ is biased and that s^2 is not. All we have to do to fix σ_{MLE} then is to multiply it by $n/(n-1)$.

6. **Prompt**: *Show that $\frac{n(\theta-\theta_{MLE})}{\theta} \xrightarrow{dist} E$*

The likelihood for this problem is:

$$L(\theta) = (1/\theta)^n = \theta^{-n}$$

Since we want to make the likelihood large, we must at a minimum make θ larger than the maximum of all the data points.

θ clearly decays in n so the MLE will trivially be the tightest fit on the data (i.e. $\max(Y_{1:n})$ or the n^th order statistic).

The CDF of θ_{MLE} is then $(x/\theta)^n$ which means $\theta_{MLE} \xrightarrow{p} \theta$. I did not have time to finish this problem though.

7. Done

8. Done