WangZhiyi, 922182514
ZhangTianyao, 922182332

Homework 1

Machine Learning
August 11, 2019

Problem 1

As the prisoner cannot remember his former choice,he faces the same situation when he makes choice. Assume expected time to get free is $E(t)$ which is composed by three situations, he chooses the tunnel leads to freedom within 3 hours, others two are lead to freedom within $7 + E(t)$ or $5 + E(t)$. Each of them has a probability of $\frac{1}{3}$. Therefore, the $E(t)$ is calculated by equation1

$$E\left(t\right) = \frac{1}{3} \cdot 3 + \frac{1}{3} \cdot \left[E\left(t\right) + 5\right] + \frac{1}{3} \cdot \left[E\left(t\right) + 7\right] \tag{1}$$

Solve equation 1the expected time it will take the prisoner to obtain freedom is 15 hours.

Problem 2

Question a

Model: $i.d.d$ samples $\{x_i\}_{i=1}^{n}$ are normally distributed, $p\left(x|y\right) \sim \mathcal{N}\left(y, 1\right)$
Parameter: y is the mean of the distribution.
Goal: Get the y when samples are given, which leads the samples have the max probability to happen.
Probability density of X is calculated by equation 2

$$p\left(X = x_i|y = \hat{y}_{MLE}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(x_i - \hat{y}_{MLE}\right)^2}{2\sigma^2}} \tag{2}$$

Likelihood Function is given by equation

$$\begin{aligned} L\left(\hat{y}_{MLE}\right) &= \prod_{i=1}^{n} p\left(X = x_i|y = \hat{y}_{MLE}\right) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(x_i - \hat{y}_{MLE}\right)^2}{2}} \\ &= \frac{n}{\sqrt{2\pi^2}} e^{-\frac{1}{2}\sum_{i=1}^{n}\left(x_i - \hat{y}_{MLE}\right)^2} \end{aligned} \tag{3}$$

Then use ln function to simplify,

$$\ln(L(\hat{y}_{MLE})) = \ln\left(\frac{n}{\sqrt{2\pi}}\right) - \frac{1}{2}\sum_{i=1}^{n}\left(x_i - \hat{y}_{MLE}\right)^2 \tag{4}$$

Derivative,

$$\frac{d\ln\left(L\left(\hat{y}_{MLE}\right)\right)}{d\hat{y}_{MLE}} = \frac{1}{2}\sum_{i=1}^{n}\left(x_i - \hat{y}_{MLE}\right) \tag{5}$$

Set the derivative equal to 0. We can get $\hat{y}_{MLE}$

$$\hat{y}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{6}$$

Question b

MAP function: Use the $\hat{y}_{MLE}$ in equation3 to

$$\begin{aligned} M\left(\hat{y}\right) &= p\left(\hat{y}\right) \cdot L\left(\hat{y}_{MLE}\right) \\ &= \frac{\sqrt{s}}{\sqrt{2\pi}} e^{-\frac{s(\hat{y}-z)^2}{2}} \cdot \frac{n}{\sqrt{2\pi}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\hat{y})^2} \end{aligned} \tag{7}$$

Use $\ln(x)$ function to simplify:

$$L(\hat{y}) = \ln\left(\frac{\sqrt{s}}{\sqrt{2\pi}}\right) - \frac{s\left(\hat{y}-z\right)^2}{2} + \ln\left(\frac{n}{\sqrt{2\pi}}\right) - \frac{1}{2}\sum_{i=1}^{n}\left(x_i - \hat{y}\right)^2 \tag{8}$$

Set $\frac{dL(y)}{dy}$ equals to 0 to get the $arg \max_{\hat{y}\in R} L\left(y\right)$

$$\frac{dL\left(\hat{y}\right)}{d\hat{y}} = -s\left(\hat{y}-z\right) + \sum_{i=1}^{n}\left(x_i - \hat{y}\right) = 0 \tag{9}$$

Solve equation9 then get the $\hat{y}_1$

$$\hat{y}_1 = \frac{n}{n+s}\hat{y}_{MLE} + \frac{sz}{n+s} \tag{10}$$

Question c

The density function of y is

$$p\left(y\right) = \begin{cases} \frac{1}{2}, -1 \leqslant y \leqslant 1 \\ 0, o.w \end{cases} \tag{11}$$

MAP function

$$\begin{aligned} M\left(\hat{y}\right) &= p\left(\hat{y}\right) \cdot L\left(\hat{y}_{MLE}\right) \\ &= \frac{1}{2} \cdot \frac{n}{\sqrt{2\pi}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\hat{y})^2} \end{aligned} \tag{12}$$

Use ln function to simplify:

$$L(\hat{y}) = \ln[M(\hat{y})] = \ln(\frac{1}{2}) + \ln(\frac{n}{\sqrt{2\pi}}) - \frac{1}{2}\sum_{i=1}^{n}(x_i - \hat{y})^2 \tag{13}$$

Set $\frac{dL(y)}{dy}$ equals to 0 to get the $\underset{\hat{y}\in R}{arg\ \max} L(y)$ when $-1 \leqslant y_{MLE} \leqslant 1$, namely

$$\frac{dL(y)}{dy} = \sum_{i=1}^{n}(x_i - \hat{y}) = 0 \tag{14}$$

Solve equation14, we can get $\hat{y}_2 = y_{MLE}$ when $-1 \leqslant y_{MLE} \leqslant 1$, while when $y_{MLE} < -1$, $\hat{y}_2 = -1$, when $y_{MLE} > 1$, $\hat{y}_2 = 1$, therefore

$$\hat{y}_2 = \begin{cases} -1,\ y_{MLE} < -1 \\ y_{MLE}, -1 \leqslant y_{MLE} \leqslant 1 \\ 1, y_{MLE} > 1 \end{cases} \tag{15}$$

Problem 3

Question a

Random variables of Poisson distribution:

$$P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda} \tag{16}$$

a set of iid samples $D = \{x_1, x_2, \cdots, x_N\}$ drawn according to the Poisson probability mass:

$$L(\lambda) = P(x_1, x_2, \cdots, x_N) = P(x_1)P(x_2)\cdots P(x_N) \tag{17}$$

We use ln to simplify

$$\begin{aligned} lnL(\lambda) &= lnP(x_1) + lnP(x_2) + \cdots + lnP(x_N) \\ &= \sum_{i=1}^{N}(x_i ln\lambda - \sum_{j=1}^{x_i} lnj - \lambda) \end{aligned} \tag{18}$$

ML condition tells:

$$\frac{\partial lnL(\lambda)}{\partial \lambda} = \sum_{i=1}^{N}\frac{x_i}{\lambda} - N = 0 \therefore \lambda_{ML} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{19}$$

$$\begin{aligned} E(\hat{\lambda}) &= E(\frac{1}{N}\sum_{i=1}^{N} x_i|\lambda) \\ &= \frac{1}{N}\sum_{i=1}^{N} E(x_i|\lambda) \\ &= \frac{1}{N}\sum_{i=1}^{N} \lambda \\ &= \lambda \end{aligned} \tag{20}$$

note that $E[X] = \lambda$:

$$\begin{aligned} E(\hat{x_k}) &= \sum_{i=1}^{\infty} x_i \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \\ &= \lambda \sum_{i=0}^{\infty} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \end{aligned}$$

So the MLE is not biased.

Question b

Use y and k to reexpress λ:

$$y = P(X = 0)^2 = e^{-2\lambda} \tag{21}$$

So $\lambda = -\frac{1}{2} lny$. Then replace λ into the formula.

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} = \frac{(-\frac{1}{2} lny)^k}{k!} \sqrt{y} \tag{22}$$

to a single k, here is maximum likelihood estimator:

$$\hat{y}_{MLE} = arg \max_{y} P(X = k) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{23}$$

Because $n = 1$, $\hat{y}_{MLE} = e^{-2\lambda}$

$$b(y) = E(y) - \hat{y}_{MLE} = \sum_{k=0}^{\infty} e^{-2k} \frac{\lambda^k}{k!} e^{-lambda} = e^{e^{-2}\lambda-\lambda} - e^{-2\lambda} \neq 0 \tag{24}$$

Obviously, $\hat{y}_{MLE}$ is biased.

Question c

Assume that y_U is biased:

$$E(y_{MLE}) = \sum_{k=0}^{\infty} y_U P(X=k) = \sum_{k=0}^{\infty} y_U \frac{\lambda^k}{k!} e^{-\lambda} = y = e^{-2\lambda} \tag{25}$$

Because $e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$,$y_U = (-1)^k$

Question d

$$\begin{aligned} MSE_{ML} &= E\left[(p_{MLE} - p)^2\right] \\ &= \sum_{k=1}^{\infty} \left(e^{-2k} - e^{-2\lambda}\right)^2 \frac{\lambda^k}{k!} e^{-\lambda} \\ &= e^{\left(e^{-4}-1\right)\lambda} - 2e^{\left(e^{-2}-2\right)\lambda} + e^{-4\lambda} MSE_U \\ &= E\left[(p_U - p)^2\right] \\ &= \sum_{k=1}^{\infty} \left((-1)^k - e^{-2\lambda}\right)^2 \frac{\lambda^k}{k!} e^{-\lambda} \\ &= 1 - e^{-4\lambda} \end{aligned} \tag{26}$$

calculate it with Python, set:

$$\begin{aligned} f(x) &= MSE_{ML} - MSE_U \\ &= [e^{(e^{-4}-1)x} - 2e^{(e^{-2}-2)x} + e^{-4x}] - [1 - e^{-4x}], x > 0 \end{aligned} \tag{27}$$

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

e = math.e

x = np.linspace(0, 20, 500)
y = (e**((e**(-4)-1)*x) - 2*e**((e**(-2)-2)*x) + e**(-4*x)) - (1 - e**(-4*x))

plt.title("compare(ML vs U)")
plt.plot(x, y)
plt.show()
```

and then draw the image of $f(x)$, it shows in figure1

$f(x) < 0, x > 0 \therefore MSE_{ML} < MSE_U$

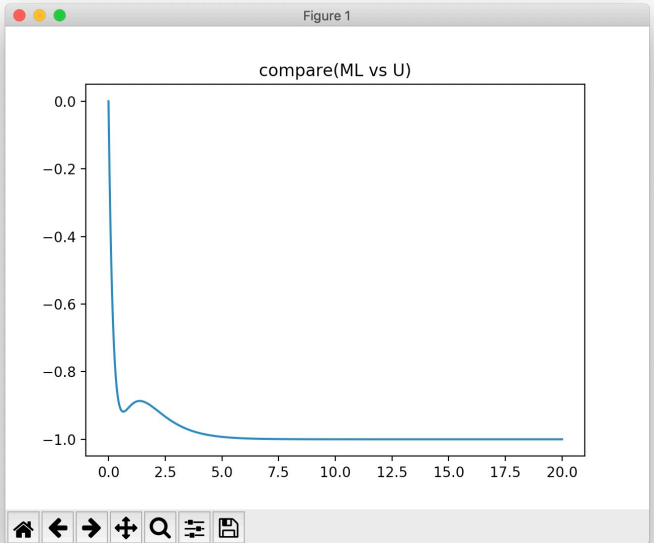

Figure 1: Figure1

Problem 4

Question a

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{28}$$

$$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})(x_i - \hat{\mu})^T \tag{29}$$

$$
\begin{aligned}
E\left[\hat{\Sigma}_{MLE}\right] &= E\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\hat{\mu}_{MLE}\right)\left(x_i-\hat{\mu}_{MLE}\right)^T\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}E\left[x_i\cdot x_i^T-x_i\cdot\hat{\mu}_{MLE}^T-\hat{\mu}_{MLE}\cdot x_i^T+\hat{\mu}_{MLE}\cdot\hat{\mu}_{MLE}^T\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}E\left[x_i\cdot x_i^T-x_i\cdot\frac{1}{n}\sum_{j=1}^{n}x_j^T-\frac{1}{n}\sum_{j=1}^{n}x_j\cdot x_i^T+\frac{1}{n}\sum_{j=1}^{n}x_j\cdot\frac{1}{n}\sum_{j=1}^{n}x_j^T\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(E\left[x_i\cdot x_i^T\right]-\frac{2}{n}\sum_{j=1,j\neq i}^{n}E\left[x_i\cdot x_j^T\right]-\frac{2}{n}E\left[x_i\cdot x_i^T\right]+\frac{1}{n^2}\sum_{j=i}E\left[x_j\cdot x_j^T\right]+\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1,k\neq j}^{n}E\left[x_j\cdot x_k^T\right]\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{n-2}{n}E\left[x_i\cdot x_i^T\right]-\frac{2n-2}{n}\mu_{MLE}\cdot\mu_{MLE}+\frac{1}{n}E\left[x_i\cdot x_i^T\right]+\frac{n^2-n}{n^2}\mu_{MLE}\cdot\mu_{MLE}^T\right) \\
&= \frac{n-1}{n^2}\sum_{i=1}^{n}E\left[x\cdot x^T-x\cdot\mu^T-\mu\cdot x^T+\mu\cdot\mu^T\right] \\
&= \frac{n-1}{n^2}\sum_{i=1}^{n}E\left[\left(x-\mu\right)\left(x-\mu\right)^T\right] \\
&= \frac{n-1}{n^2}\sum_{i=1}^{n}\Sigma \\
&= \frac{n-1}{n}\Sigma
\end{aligned}
\tag{30}
$$

Obviously, $\hat{\Sigma}_{MLE}$ is biased

Question b

$$
\hat{\Sigma}'=\frac{n}{n-1}\hat{\Sigma}_{MLE}=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\hat{\mu})(x_i-\hat{\mu})^T \tag{31}
$$

Problem 5

Question a

We choose k samples from range $[1,2,\cdots,N]$So the **ML Function**is

$$
M(N)=p\left(x_1,x_2,\cdots,x_k|N\right)=\prod_{i=0}^{i=k-1}\frac{1}{N-i} \tag{32}
$$

When we get the $arg\max_{N\in R}M\left(N\right)$ Then define a new function $t(N)$ to simplify, we just need to consider about denominator of $M(x)$

$$
t\left(N\right)=N\cdot\left(N-1\right)\cdots\left(N-k+1\right)=\frac{N!}{k!\left(N-k\right)!}\cdot k!=C_N^k\cdot k! \tag{33}
$$

It is easy to find that $\underset{N \in R}{arg\ \max} M\left(N\right) = \underset{N \in R}{arg\ \min} t\left(N\right)$ Then we find the $\underset{N \in R}{arg\ \min} t\left(N\right)$

$$\begin{aligned} \underset{N \in R}{arg\ \min} t\left(N\right) &= \underset{N \in R}{arg\ \min}\ C_N^k \cdot k! \\ &= \underset{N \in R}{arg\ \min}\ \frac{N!}{k!\left(N-k\right)!} \\ &= \underset{N \in R}{arg\ \min} N \\ &= \underset{x \in \{x_i\}_1^k}{arg\ \max} x_i \end{aligned} \tag{34}$$

Therefore, $\hat{N}_{MLE} = max\{x_i | x_i \in \{x_i\}_1^k\}$

Question b

We should prove $E(\hat{N}_{MLE}) \neq N$

$$
\begin{aligned}
E(\hat{N}_{MLE}) &= \sum_{i=1}^{N} P\left(\hat{N}_{MLE} = i\right) \cdot i \qquad (35)\\
&= \sum_{i=1}^{N} \left[P\left(\hat{N}_{MLE} \leqslant i\right) - P\left(\hat{N}_{MLE} \leqslant i-1\right)\right] \cdot i\\
&= \underbrace{\left(\frac{C_k^k}{C_N^k} - 0\right) \cdot k}_{i \leqslant k} + \underbrace{\sum_{i=k+1}^{N} \left[\frac{C_i^k}{C_N^k} - \frac{C_{i-1}^k}{C_N^k}\right] \cdot i}_{k+1 \leqslant i \leqslant N} + \underbrace{\left(1 - \frac{C_N^k}{C_N^k}\right)}_{i=N+1} + \underbrace{(1-1)}_{i>N+1}\\
&= \frac{k}{\frac{N!}{k!(N-k)!}} + \sum_{i=k+1}^{N} \left[\frac{\frac{i!}{k!(i-k)!}}{\frac{N!}{k!(N-k)!}} - \frac{\frac{(i-1)!}{k!(i-k-1)!}}{\frac{N!}{k!(N-k)!}}\right] \cdot i\\
&= \frac{k \cdot k!\,(N-k)!}{N!} + \sum_{i=k+1}^{N} \left[\frac{i!\,(N-k)!}{(i-k)!N!} - \frac{(i-1)!\,(N-k)!}{(i-k-1)!N!}\right] \cdot i\\
&= \frac{k \cdot k!\,(N-k)!}{N!} + \frac{(N-k)!}{N!} \cdot \sum_{i=k+1}^{N} \left[\frac{i!}{(i-k)!} + \frac{(i-1)!}{(i-k-1)}\right] \cdot i\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + \sum_{i=k+1}^{N} \left[\frac{i!}{(i-k)!} + \frac{i!}{(i-k)} \cdot \frac{i-k}{i}\right] \cdot i\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + \sum_{i=k+1}^{N} \frac{i!}{(i-k)!}\,(2i-k)\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + (k+1)! \sum_{i=k+1}^{N} \frac{i!}{(i-k-1)!\,(k+1)!} \cdot \frac{2i-k}{i-k}\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + (k+1)! \sum_{i=k+1}^{N} \frac{i!}{(i-k-1)!\,(k+1)!} \cdot \left(1 + \frac{i}{i-k}\right)\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + (k+1)! \cdot \left(\sum_{i=k+1}^{N} \frac{i!}{(i-k-1)!\,(k+1)!} + \sum_{i=k+1}^{N} \frac{i!}{(i-k-1)!\,(k+1)!} \cdot \frac{i}{i-k}\right)\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + (k+1)! \cdot \left(\sum_{i=k+1}^{N} \frac{i!}{(i-k-1)!\,(k+1)!} + \sum_{i=k+1}^{N} \frac{i!}{(i-k-1)!\,(k+1)!} \cdot \frac{i}{i-k}\right)\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + (k+1)! \cdot \left(\frac{(N+1)!}{(N-k+1)!\,(k+1+1)!} + \sum_{i=k+1}^{N} \frac{i!}{(i-k-1)!\,(k+1)!} \cdot \left(1 + \frac{k}{i-k}\right)\right)\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + (k+1)! \cdot \left(\frac{2\,(N+1)!}{(N-k+1)!\,(k+1+1)!} + \sum_{i=k+1}^{N} \frac{i!}{(i-k)!\,(k+1)!} \cdot k\right)\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + (k+1)! \cdot \left(\frac{2\,(N+1)!}{(N-k+1)!\,(k+1+1)!} + \frac{k}{(k+1)} \cdot \left(\sum_{i=k}^{N} \frac{i!}{(i-k)!k!} - 1\right)\right)\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + (k+1)! \cdot \left(\frac{2\,(N+1)!}{(N-k+1)!\,(k+1+1)!} + \frac{k}{(k+1)} \cdot \left(\frac{(N+1)!}{(N-k)!\,(k+1)!} - 1\right)\right)\right)\\
&= \frac{(N-k)!}{N!} \left(k \cdot k! + \left(\frac{2\,(N+1)!}{(N-k+1)!\,(k+2)} + \frac{k}{(k+1)} \cdot \left(\frac{(N+1)!}{(N-k)!} - 1\right)\right)\right)\\
&= \frac{k}{k+1}\,(N+1) \neq N
\end{aligned}
$$

Question c

According to question b unbiased estimator is $N_{ub} = \frac{k+1}{k}\hat{N}_{MLE} - 1$

Question d

The simulation result is shown in figure 2

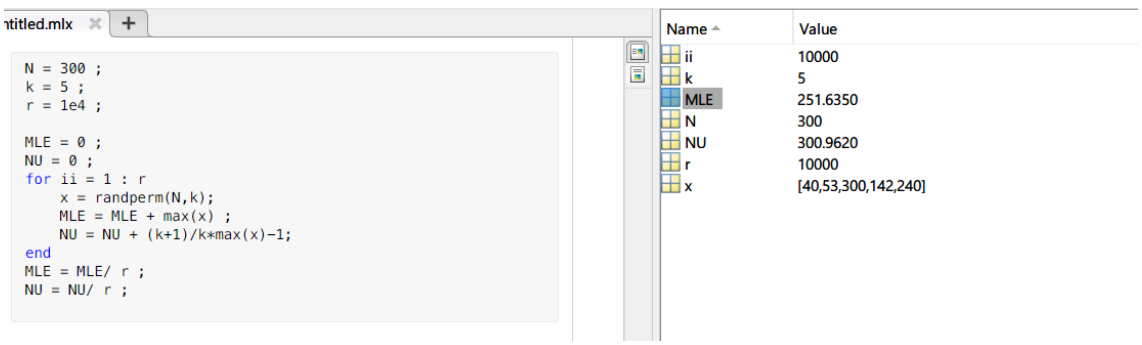

Figure 2: Figure2