Ex2 - theoretical Lior Ziv

Poission MLE

1. Since P(X=k) = $\frac{\lambda^k e^{-\lambda}}{k!}$ we get that L(λ :D) = ($\prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$) $\Rightarrow$LogL(λ :D) = log($\prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$) = $\sum_{i=1}^{n}$ $log(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!})$

2. The sufficient statistics is T(x) = $\sum_{i=1}^{n} x_i$ (calculated next)

3. In order to find the MLE I will derive $\sum_{i=1}^{n}$ $log(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!})$ = $\sum_{i=1}^{n}$ $log(\lambda^{x_i})$ $+log(e^{-\lambda})$ - $log(x_i!)$ = $\frac{\sum_{i=1}^{n} x_i}{\lambda}$ -λn , compare to zero $\frac{\sum_{i=1}^{n} x_i}{\lambda}$ -n = 0 $\rightarrow \sum_{i=1}^{n} x_i$ - $\lambda n \rightarrow \lambda = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x}$

Gaussian MLE

1. Since P(x : μ, σ) = $\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$ we get that L(μ, σ :D) = ($\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}$) $\Rightarrow$LogL(μ, σ :D) = log($\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}$) = $\sum_{i=1}^{n}$ $log(\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}})$

2. There's two sufficient statistics :
   (a) With respect to σ^2- $\sum_{i=1}^{n} (x_i)^2$ ($\mu$is known)
   (b) With respect to μ - $\sum_{i=1}^{n} (x_i)$

3. In order to find the MLE I will derive $\sum_{i=1}^{n}$ $log(\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}})$ = $\sum_{i=1}^{n}$ $log(\frac{1}{\sigma\sqrt{2\pi}})$ + $\sum_{i=1}^{n}$ $log(e^{\frac{-(x_i-\mu)^2}{2\sigma^2}})$ = $\sum_{i=1}^{n}$ $log(\frac{1}{\sigma\sqrt{2\pi}})$ + $\sum_{i=1}^{n} \frac{-(x_i-\mu)^2}{2\sigma^2}$

- now we can derive according to both μ, σ :
  1. μ : $\sum_{i=1}^{n} \frac{2(x_i-\mu)}{2\sigma^2}$, compare to zero $\sum_{i=1}^{n} \frac{(x_i-\mu)}{\sigma^2} = 0 \rightarrow \sum_{i=1}^{n} 2(x_i-\mu)2\sigma^2 =$ $\sum_{i=1}^{n} (x_i)\sigma^2$- $\sum_{i=1}^{n} \mu\sigma^2 \rightarrow n\mu = \sum_{i=1}^{n} (x_i) \rightarrow \mu = \frac{\sum_{i=1}^{n} (x_i)}{n}$
  2. σ^2: $\sum_{i=1}^{n}$ $log(\frac{1}{\sqrt{2\pi\sigma^2}})$ + $\sum_{i=1}^{n} \frac{-(x_i-\mu)^2}{2\sigma^2} = \sum_{i=1}^{n} -\frac{1}{\sigma^2} + \sum_{i=1}^{n} \frac{2(x_i-\mu)^2}{4(\sigma^2)^2} \rightarrow -2n\sigma^2 + \sum_{i=1}^{n}$ $2(x_i-\mu)^2 \rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} (x_i-\mu)^2}{n}$

Uniform MLE

1. For uniform distribution $f(x) = \begin{cases} \frac{1}{N} & m \leq N \\ \\ 0 & N < m \end{cases}$, so $L(N : m) = \prod_{i=1}^{n} \frac{1}{N}$ =N^{-n} $\rightarrow$ $logL$(N:m) = log(N^{-n})

2. In order to find the MLE I will derive log(N^{-n}) $\rightarrow$ $\frac{-n}{N}$ now it is clear that this function is decreasing, so the N that will best predict our sampels will be at the maximal x$_i$$\rightarrow$ m

3. For a known N the expected value of m will be m= $\frac{N+1}{2}$(since m is distributed uniformly between (1,N))

Constrained MLE

We have f(x) = $\sum_{i=1}^{n} x_i$, s.t $x_i \in \{1...k\}$ each x_i represents the i'th roll result , θ_i is the probability for each i$\in$ $\{1..k\}$ to appear in the x$_i$$roll$

1. k = 3 g_1 $\rightarrow$$\theta_1$ =$\theta_2 + \theta_3$ and in order to have a probability sum which adds up to 1 we will have g$_2$ $\rightarrow$$\theta_1$+$\theta_2$ + θ_3= 1

   Now we will take the LLR of f(x) $\rightarrow \sum_{i=1}^{n}$ $N_i log(\theta_i)$ where N_i is the number of times the i'th side came out together with the constraints we will compose J = $\sum_{i=1}^{n}$ $N_i log(\theta_i)$ - $\lambda_1(\theta_1$$-\theta_2 - \theta_3)$ - $\lambda_2(\theta_1$+$\theta_2 + \theta_3$- 1) we will derive it by θ_i i$\in$ $\{1, 2, 3\}$

- $\frac{\nabla J}{\nabla \theta_1}$ =$\frac{N_1}{\theta_1}$ -$\lambda_1 - \lambda_2$= 0 $\rightarrow$$\theta_i$ =$\frac{N_i}{+\lambda_1 - 3\lambda_2}$
- $\frac{\nabla J}{\nabla \theta_2}$ =$\frac{N_2}{\theta_2}$ $+\lambda_1 - \lambda_2$= 0
- $\frac{\nabla J}{\nabla \theta_3}$ =$\frac{N_3}{\theta_3}$ $+\lambda_1 - \lambda_2$= 0
  - From θ_1,θ_2 equations we get that $\frac{N_2\theta_3}{N_3\theta_2}$ =1$\rightarrow$ θ_3 =$\frac{N_3\theta_2}{N_2}$
- Now placing back $\theta_3 in$ our constraints equations we get :
  - $\theta_1$$-\theta_2$$-\frac{N_3\theta_2}{N_2}$ = 0 $\rightarrow$ θ_1 = θ_2+$\frac{N_3\theta_2}{N_2}$
  - θ_1+θ_2+$\frac{N_3\theta_2}{N_2}$- 1 = 0 if we place $\theta_1$from above here we get θ_2 =$\frac{N_2}{2(N_3+N_2)}$, by that we also get θ_3= $\frac{N_3}{2(N_3+N_2)}$
  - Now we now θ_1 =$\frac{N_2}{2(N_3+N_2)}$ + $\frac{N_3}{2(N_3+N_2)}$ = 0.5

  Hence I found all the θ_i - θ_1 =0.5 ,θ_2 =$\frac{N_2}{2(N_3+N_2)}$,θ_3= $\frac{N_3}{2(N_3+N_2)}$

1. k $= 4$ $g_1 \to \theta_1 = \theta_2 + \theta_3$ and in order to have a probability sum which adds up to 1 we will have g$_2 \to \theta_1+\theta_2+\theta_3+\theta_4 = 1$ by which we get $2\theta_2+2\theta_3+\theta_4 = 1$

   I will do the same as the previews task :

   J $= \sum_{i=1}^{n} N_i log(\theta_i)$ - $\lambda_1(\theta_1 - \theta_2 - \theta_3)$ - $\lambda_2(\theta_1 + \theta_2 + \theta_3 + \theta_4$- 1) we will derive it by θ_i i$\in \{1, 2, 3, 4\}$

   - $\frac{\nabla J}{\nabla \theta_1} = \frac{N_1}{\theta_1}$ -$\lambda_1 - \lambda_2 = 0$
   - $\frac{\nabla J}{\nabla \theta_2} = \frac{N_2}{\theta_2} + \lambda_1 - \lambda_2 = 0$
   - $\frac{\nabla J}{\nabla \theta_3} = \frac{N_3}{\theta_3} + \lambda_1 - \lambda_2 = 0$
   - $\frac{\nabla J}{\nabla \theta_4} = \frac{N_4}{\theta_4}$ -$\lambda_2 = 0$
     - from θ_3,$\theta_2$equations we get that $\frac{N_1\theta_3}{\theta_2 N_3} = 1 \to \theta_3 = \frac{N_3\theta_2}{N_1}$
     - from θ_1,θ_3 equations we get that $= 0 \to \frac{N_2\theta_3}{\theta_2 N_3} = \frac{\lambda_1+\lambda_2}{\lambda_1-\lambda_2} \to \lambda_1 = \frac{\lambda_2(\theta_2 N_3+\theta_3 N_2)}{\theta_3 N_2-\theta_2 N_3} = \frac{N_1-N_3-N_2}{N_1+N_3+N_2}\lambda_2$
     - from θ_4,$\theta_2$equations we get that $\frac{N_4\theta_2}{\theta_4 N_2} = \frac{\lambda_2}{\lambda_1-\lambda_2} \to \theta_4 = \frac{N_4\theta_2(\lambda_1+\lambda_2)}{N_2\lambda_2}$
   - Now i can place θ_3, θ_4 $2\theta_2 + 2\theta_3 + \theta_4 = 1$
   - Now we can place θ_1,θ_2,$\theta_4$in $2\theta_2 + 2\theta_3+\theta_4 = 1 \to 2\theta_2 + 2\frac{N_3\theta_2}{N_1} + \frac{N_4\theta_2(\lambda_1+\lambda_2)}{N_2\lambda_2} \to \theta_2 = \frac{N_2(N_1+N_3+N_2)}{2(N_1+N_2+N_3+N_4)(N_3+N_2)}$
     - Now we can find the rest
     - by $\theta_2 \to \theta_3 = \frac{N_3\theta_2}{N_1} = \frac{N_2(N_1+N_3+N_2)}{2(N_1+N_2+N_3+N_4)(N_3+N_2)}(\frac{N_3}{N_1})$
     - $\theta_1 = \frac{N_1+N_3+N_2}{2(N_1+N_2+N_3+N_4)}$
     - by g$_1, g_2$we get $2\theta_1 + \theta_4 = 1$ so $\theta_4 =$-$\frac{N_1+N_3+N_2}{(N_1+N_2+N_3+N_4)}$+1

2. k $= 5$ $g_1 \to \theta_1 = \theta_2 + \theta_3 + \theta_4 + \theta_5$-and in order to have a probability sum which adds up to 1 we will have g$_2 \to \theta_1+\theta_2+\theta_3+\theta_4+\theta_5 = 1$

   I will do the same as the previews task :

   J $= \sum_{i=1}^{n} N_i log(\theta_i)$ - $\lambda_1(\theta_1 - \theta_2 - \theta_3$-$\theta_4 - \theta_5)$ - $\lambda_2(\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5$- 1) we will derive it by θ_i i$\in \{1, 2, 3, 4, 5\}$

   - $\frac{\nabla J}{\nabla \theta_1} = \frac{N_1}{\theta_1}$ -$\lambda_1 - \lambda_2 = 0 \to \theta_1 = \frac{N_1}{\lambda_1+\lambda_2}$
     - $\frac{\nabla J}{\nabla \theta_2} = \frac{N_2}{\theta_2} + \lambda_1 - \lambda_2 = 0 \to \theta_2 = \frac{N_2}{-\lambda_1+\lambda_2}$
     - $\frac{\nabla J}{\nabla \theta_3} = \frac{N_3}{\theta_3} + \lambda_1 - \lambda_2 = 0 \to \theta_3 = \frac{N_3}{-\lambda_1+\lambda_2}$
     - $\frac{\nabla J}{\nabla \theta_4} = \frac{N_4}{\theta_4} + \lambda_1 - \lambda_2 = 0 \to \theta_4 = \frac{N_4}{-\lambda_1+\lambda_2}$
     - $\frac{\nabla J}{\nabla \theta_4} = \frac{N_5}{\theta_5} + \lambda_1$-$\lambda_2 = 0 \to \theta_4 = \frac{N_5}{-\lambda_1+\lambda_2}$
   - Now lets look back at out constraints
     - $\theta_1 - \theta_2 - \theta_3$-$\theta_4 - \theta_5 = 0 \to \frac{N_1}{\lambda_1+\lambda_2} = \frac{N_2}{-\lambda_1+\lambda_2} + \frac{N_3}{-\lambda_1+\lambda_2} + \frac{N_4}{-\lambda_1+\lambda_2} + \frac{N_5}{-\lambda_1+\lambda_2}$
       $\Rightarrow N_1(-\lambda_1 + \lambda_2) = (\lambda_1 + \lambda_2)(N_2 + N_3 + N_4 + N_5)$

- θ_1+θ_2+θ_3+θ_4+θ_5- 1 $\rightarrow \frac{N_1}{\lambda_1+\lambda_2}+\frac{N_2}{-\lambda_1+\lambda_2}+\frac{N_3}{-\lambda_1+\lambda_2}+\frac{N_4}{+\lambda_2}+\frac{N_5}{-\lambda_1+\lambda_2}$ $=$ 1$\rightarrow$from the previews be can deduce $\rightarrow \frac{2N_1}{\lambda_1+\lambda_2}$=1 $\rightarrow 2N_1 = \lambda_1+\lambda_2$ so we get λ_1 =$2N_1-\lambda_2$
- Back to the first equation $\Rightarrow N_1(-\lambda_1+\lambda_2)$= $(\lambda_1+\lambda_2)(N_2+N_3+N_4+N_5)$ we can place $\lambda_1 \Rightarrow N_1(-(2N_1-\lambda_2)+\lambda_2)$= $(2N_1-\lambda_2+\lambda_2)(N_2+N_3+N_4+N_5) \rightarrow\Rightarrow N_1(-2N_1+2\lambda_2) = (2N_1)(N_2+N_3+N_4+N_5) \rightarrow \lambda_2 = 2N_1$
- Now if we go back to $\rightarrow 2N_1 = \lambda_1+\lambda_2$ we will get λ_2 =0
- We can go back to the above equations and get θ_i
  * $\theta_1 = \frac{N_1}{2N_1}$ =0.5
  * $\theta_2 = \frac{N_2}{2N_1}$, $\theta_3 = \frac{N_3}{2N_1}$,$\theta_4 = \frac{N_4}{2N_1}, \theta_5 = \frac{N_5}{2N_1}$