MA 413 - Statistics for Data Science

Solutions to Exercise 9

1. (a) Let X_i, $i = 1 \dots n$ be i.i.d. Geometric random variables with $f(x;p) = (1-p)^x p$. The likelihood is :

$$L(p) = \prod_{i=1}^{n} (1-p)^{x_i} p = (1-p)^{\sum_{i=1}^{n} x_i} p^n$$

   (b) The log-likelihood is :

$$l(p) = \sum_{i=1}^{n} x_i \log(1-p) + n \log p$$

   (c) To get the MLE of p, we have to maximize the likelihood. It is equivalent to maximise the log-likelihood by monotonicity which is easier. We have :

$$l'(\hat{p}) = -\frac{1}{1-\hat{p}} \sum_{i=1}^{n} x_i + n\frac{1}{\hat{p}} = 0 \Rightarrow \hat{p} = \frac{n}{\sum_{i=1}^{n} x_i + n}$$

   To check that it is indeed a maximum, notice that :

$$l''(p) = -\frac{\sum_{i=1}^{n} x_i}{(1-p)^2} - \frac{1}{p^2} < 0 \qquad \forall p \in (0,1)$$

2. (a) Let X_i, $i = 1 \dots n$ be i.i.d. Laplace random variables with $f(x;\sigma) = \frac{1}{2\sigma} \exp(-\frac{|x|}{\sigma})$. The likelihood is :

$$L(\sigma) = \prod_{i=1}^{n} \frac{1}{2\sigma} \exp(-\frac{|x|}{\sigma}) = \left(\frac{1}{2\sigma}\right)^n \exp\left(\frac{\sum_{i=1}^{n} |x_i|}{\sigma}\right)$$

   (b) The log-likelihood :

$$l(\sigma) = -n \log 2\sigma - \frac{\sum_{i=1}^{n} |x_i|}{\sigma}$$

   (c) To maximise the log-likelihood, and thus the likelihood, we solve :

$$l'(\hat{\sigma}) = -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n} |x_i|}{\sigma^2} = 0 \Rightarrow \hat{\sigma} = \frac{\sum_{i=1}^{n} |x_i|}{n}$$

   To check that it is indeed a maximum, notice that :

$$l''(\hat{\sigma}) = \frac{n}{\hat{\sigma}^2} - 2\frac{\sum_{i=1}^{n} |x_i|}{\hat{\sigma}^3} = -\frac{n^3}{(\sum_{i=1}^{n} |x_i|)^2} < 0$$

3. Let X_i, $i = 1 \dots n$ be i.i.d. uniform random variables with $f(x;\theta) = \frac{1}{\theta}\mathbb{1}_{\{-\theta<x<0\}}$. Since the likelihood is zero outside the set $\{x \in \mathbb{R}^n : x_i \in (-\theta, 0) \quad i = 1 \dots n\}$, it is good enough to find the maximum value of the likelihood when the sample satisfies this condition. Now since in this case, the likelihood is a decreasing function, the maximiser is given by $\hat{\theta} = -\min x_i$. More formally, we get :

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta}\mathbb{1}_{\{-\theta<x_i<0\}} = \frac{1}{\theta^n}\mathbb{1}_{\{\min x_i > -\theta\}}\mathbb{1}_{\{\max x_i < 0\}}$$

Drawing this function (see Figure 1), it is easy to see that $\hat{\theta} = -\min x_i$

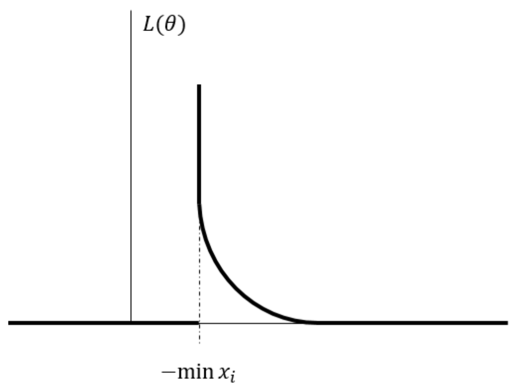

Figure 1: Likelihood for i.i.d. uniform random variables

Remark. In fact, one can say that the MLE is not well defined since it is at the boundary of the likelihood. Nevertheless this discussion is beyond the scope of this course and to simplify, we will assume taking the MLE at the boundary is not a problem.

4. (a) Let X_i, $i = 1 \dots n$ be i.i.d. gamma r.v with $f(x;\alpha) = \frac{x^{\alpha-1}\exp(-x)}{\Gamma(\alpha)}$. The likelihood is :

$$L(\alpha) = \prod_{i=1}^{n} \frac{x^{\alpha-1}\exp(-x)}{\Gamma(\alpha)} = \frac{(\prod_{i=1}^{n} x_i^{\alpha-1})\exp(-\sum_{i=1}^{n}|x_i|)}{\Gamma(\alpha)^n}$$

(b) The log-likelihood is :

$$l(\alpha) = (\alpha-1)\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n} x_i - n\log\Gamma(\alpha)$$

(c) To maximise the log-likelihood and thus the likelihood, we solve :

$$l'(\hat{\alpha}) = \sum_{i=1}^{n} \log x_i - n\Psi(\hat{\alpha}) = 0 \Rightarrow \Psi(\hat{\alpha}) = \frac{\sum_{i=1}^{n} \log x_i}{n}$$

Remark. The above equation does not have analytical solutions. In practice, one can solve it using numerical method such as Newton's method.

5. (a) Let X_i, $i = 1 \dots n$ be i.i.d. Pareto r.v with $f(x;\theta) = \theta \mu^{\theta} x^{-1-\theta}$ with $x \geq \mu, \theta > 1$. Note that the likelihood is zero outside of the set $\{x \in \mathbb{R}^n : x_i \geq \mu, i = 1 \dots n\}$. So it is good enough to consider the maximisation of $L(\alpha)$ when the sample points satisfy this condition. Then :

$$L(\alpha) = \prod_{i=1}^{n} \theta \mu^{\theta} x^{-1-\theta}$$

   (b) The log-likelihood is :

$$l(\alpha) = n \log \theta + n\theta \log \mu - (1+\theta) \log \sum_{i=1}^{n} x_i$$

   (c) Maximising the log-likelihood, we get :

$$l'(\hat{\theta}) = \frac{n}{\hat{\theta}} + n \log \mu - \log \sum_{i=1}^{n} x_i = 0 \Rightarrow \hat{\theta} = \frac{n}{\log \sum_{i=1}^{n} x_i - n \log \mu}$$

   To check that it is indeed a maximum, notice that :

$$l''(\theta) = -\frac{n}{\theta^2} < 0$$

6. (a) Let X_i, $i = 1 \dots n$ be i.i.d. Poisson r.v with $f(x;\lambda) = \frac{\exp(-\lambda)\lambda^{-x}}{x!}$ with $x \in \mathbb{N}_*$. The likelihood is given by :

$$L(\lambda) = \frac{\exp(-n\lambda)\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

   (b) We can now derive from this the log-likelihood :

$$l(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i \log(\lambda) - \sum_{i=1}^{n} \log(x_i!)$$

   (c) Solving $l'(\hat{\lambda}) = 0$, we get :

$$l'(\hat{\lambda}) = -n + \frac{\sum_{i=1}^n x_i}{\lambda} = 0 \Rightarrow \hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n}$$

To check that it is indeed a maximum, notice that :

$$l''(\lambda) = -\frac{\sum_{i=1}^n x_i}{\lambda^2} < 0$$

7. (a) Let X_i, $i = 1 \dots n$ be i.i.d. r.v with $f(x;\theta) = \theta x^{\theta-1} \mathbb{1}_{\{x \in (0,1)\}}$. Note that the likelihood is zero outside of the set $\{x \in \mathbb{R}^n : x_i \in (0,1), i \in \{1 \dots n\}\}$. So it is good enough to consider the maximisation of $L(\theta)$ when the sample points satisfy this condition. Then :

$$L(\theta) = \theta^n \prod_{i=1}^n x_i^{\theta-1}$$

(b) The log-likelihood is:

$$l(\theta) = n \log \theta + (\theta - 1) \sum_{i=1}^n \log x_i$$

(c) Solving $l'(\hat{\theta}) = 0$, we get :

$$l'(\hat{\theta}) = \frac{n}{\theta} + \sum_{i=1}^n \log x_i = 0 \Rightarrow \hat{\theta} = -\frac{n}{\sum \log x_i}$$

To check that it is indeed a maximum, notice that :

$$l''(\theta) = -\frac{n}{\theta^2} < 0$$

8. (a) Let X_i, $i = 1 \dots n$ be i.i.d. Laplace r.v. with $f(x;\theta) = \frac{1}{2} \exp(-|x - \theta|), x \in \mathbb{R}$. The likelihood is :

$$L(\theta) = \frac{1}{2^n} \exp(-\sum_{i=1}^n |x_i - \theta|)$$

(b) The log-likelihood is :

$$l(\theta) = -n \log 2 - \sum_{i=1}^n |x_i - \theta|$$

(c) To solve this equation, notice that $l(\theta) < 0$ and the maximum is reach when $\sum_{i=1}^{n} |x_i - \theta|$ is minimal. Taking the derivative, we have to solve :

$$\sum_{i=1}^{n} \text{sign}(x_i - \theta) = 0$$

If n is even, we solve this equation taking θ such that half of the x_i are lower than θ and half of the x_i are higher than θ. There is an infinite number of solutions, all located in the set $(x_{(\frac{n}{2})}, x_{(\frac{n}{2}+1)})$.

If n is odd, the solution is unique since the only way to solve the equation is to take θ as the median of the sample, that is $\hat{\theta} = x_{(\frac{n+1}{2})}$.