200B HW#2 solution

7.5 Maximum Likelihood Estimators

8. The likelihood function is

$$f_n(\mathbf{x}|\theta) = \begin{cases} \exp(n\theta - \sum_{i=1}^n x_i) & \text{for } \min(x_1, \ldots, x_n) > \theta \\ 0 & \text{otherwise.} \end{cases}$$

(a) For each value of $\mathbf{x}$, $f_n(\mathbf{x}|\theta)$ will be a maximum when θ is made as large as possible subject to the strict inequality $\theta < \min(x_1, \ldots, x_n)$. Therefore, the value $\theta = \min(x_1, \ldots, x_n)$ cannot be used and there is no MLE.

(b) Change the pdf to the equivalent (only the inequality sign is changed)

$$f_n(\mathbf{x}|\theta) = \begin{cases} \exp(n\theta - \sum_{i=1}^n x_i) & \text{for } \min(x_1, \ldots, x_n) \geq \theta \\ 0 & \text{otherwise.} \end{cases}$$

Then the likelihood function $f_n(\mathbf{x}|\theta)$ will be nonzero for $\theta \leq \min(x_1, \ldots, x_n)$ and the MLE will be $\hat{\theta} = \min(x_1, \ldots, x_n)$.

10. The likelihood function is

$$f_n(\mathbf{x}|\theta) = \frac{1}{2^n} \exp\left\{ -\sum_{i=1}^n |x_i - \theta| \right\}.$$

Therefore, the MLE of θ will be the value that minimizes $\sum_{i=1}^n |x_i - \theta|$. $\hat{\theta}$ = the sample median of x_i is one of the solutions (*non-unique*) to this minimization problem. To see this:

$$\frac{\partial |x_i - \theta|}{\partial \theta} = \begin{cases} -1, & \text{if } \theta < x_i \\ 1, & \text{if } \theta > x_i \\ \text{does not exist}, & \text{if } \theta = x_i \end{cases}$$

$$\begin{aligned} \frac{\partial L(\theta)}{\partial \theta} &= -\sum_{i=1}^{n} \frac{\partial |x_i - \theta|}{\partial \theta} \\ &= \sum_{x_i > \theta} 1 + \sum_{x_i < \theta} -1 \\ &= -\#\{x_i : x_i < \theta\} + \#\{x_i : x_i > \theta\}, \quad \text{if } \theta \neq x_i \text{ for all } i, \end{aligned}$$

where $\#\{x_i : x_i < \theta\}$ means the number of x_is that are smaller than θ. $\frac{\partial l(\theta)}{\partial \theta} > 0$ ($l(\theta)$ increasing) when there are more x_i lying right to θ, and is negative ($l(\theta)$ decreasing) when more x_i are lying left to θ, so the maximum of $l(\theta)$ is taken when there are equal number of x_i lying on either sides of θ. We have $\hat{\theta}$ is the middle value among $x_i, \ldots, x_n$ if n is odd; or is any point between the two middle values among $x_1, \ldots, x_n$ if n is even (again, *non-unique*). When θ equals any of the x_is, $l(\theta)$ is still continuous at these x_is, Therefore, when n is even, the mle of θ would be any point between the two middle values among $x_1, \ldots, x_n$, including the end point, i.e., if $n = 2k$, the non-unique mle, $\hat{\theta}$, is any point in $[x_{(k)}, x_{(k+1)}]$, where $x_{(i)}$ is the i^{th} order statistic.

7.6 Properties of Maximum Likelihood Estimators

3. The median of an exponential distribution with parameter β is the number m such that

$$\int_0^m \beta \exp(-\beta x)\, \mathrm{d}x = \frac{1}{2}.$$

By calculus $m = \log(2)/\beta$, and it follows from the invariant property of MLE that the MLE $\hat{m} = \log(2)/\hat{\beta}$, where $\hat{\beta}$ is the MLE of β. The joint pdf of $X_1, \ldots, X_n$ is

$$f(\mathbf{x}|\beta) = \begin{cases} \beta^n \exp(-\beta \sum_{i=1}^n x_i) & x_1, \ldots, x_n > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Set $\frac{\partial f(\mathbf{x}|\beta)}{\partial \beta} = 0$ we find $\hat{\beta} = 1/\bar{X}_n$. $\hat{\beta}$ is a maximizer because $\frac{\partial f(\mathbf{x}|\beta)}{\partial \beta} > 0$ when $\beta < \hat{\beta}$ and $\frac{\partial f(\mathbf{x}|\beta)}{\partial \beta} < 0$ when $\beta > \hat{\beta}$. So $\hat{m} = \log(2)\bar{X}_n$.

8. By the invariant property of MLE, the MLE for $\theta = \Gamma'(\alpha)/\Gamma(\alpha)$ is $\hat{\theta} = \Gamma'(\hat{\alpha})/\Gamma(\hat{\alpha})$. By equation (7.6.5) we know the MLE $\Gamma'(\hat{\alpha})/\Gamma(\hat{\alpha}) = \frac{1}{n}\sum_{i=1}^{n}\log(x_i)$.

23.

(a) The means of X_i and X_i^2 are respectively $\alpha/(\alpha+\beta)$ and $\alpha(\alpha+1)/[(\alpha+\beta)(\alpha+\beta+1)]$. We set these equal to the sample moment $\bar{x}$ and $\overline{x^2}$ and solve for α and β. After some tedious algebra, we have

$$\hat{\alpha} = \frac{\bar{x}(\bar{x} - \overline{x^2})}{\overline{x^2} - \bar{x}^2}$$
$$\hat{\beta} = \frac{(1-\bar{x})(\bar{x} - \overline{x^2})}{\overline{x^2} - \bar{x}^2}$$

7.10 Supplementary Exercises

4. Since the joint p.d.f. of the observations is equal to $1/\theta^n$ provided that $\theta \le x_i \le 2\theta$ for $i = 1, \ldots, n$, the M.L.E. will be the smallest value of θ that satisfies these restrictions. Since we can rewrite the restrictions in the form

$$\frac{1}{2}\max\{x_1, \ldots, x_n\} \le \theta \le \min\{x_1, \ldots, x_n\}$$

it follows that the smallest possible value of θ is

$$\hat{\theta} = \frac{1}{2}\max\{x_1, \ldots, x_n\}.$$

10. $\theta = (1/3)(1+\beta)$ and $0 \le \beta \le 1$ implies θ must lie in the interval $[1/3, 2/3]$. The log-likelihood function is

$$L(\theta) = \log(f(\mathbf{x}|\theta)) = n\bar{x}\log(\theta) + (n - n\bar{x})\log(1-\theta),$$
$$\frac{\partial L(\theta)}{\partial\theta} = \frac{n\bar{x} - n\theta}{\theta(1-\theta)},$$
$$\frac{dL(\theta)^2}{d^2\theta} = -\frac{\theta^2 - 2\theta\bar{x} + \bar{x}}{\theta^2(1-\theta)^2} = -\frac{\theta^2 - 2\theta\bar{x} + \bar{x}^2 + \bar{x} - \bar{x}^2}{\theta^2(1-\theta)^2} < 0.$$

We look at 3 cases $\bar{x} < 1/3$, $1/3 \le \bar{x} \le 2/3$ and $\bar{x} > 2/3$.

- Case 1: $\bar{x} < 1/3$. Since $\theta \ge \frac{1}{3}$, $\frac{dL(\theta)}{d\theta} = \frac{n\bar{x} - n\theta}{\theta(1-\theta)} < 0$. It means $L(\theta)$ is a decreasing function of θ and $\hat{\theta} = \frac{1}{3}$.

- Case 2: $1/3 \le \bar{x} \le 2/3$. Since $L(\theta)$ is a concave function of θ and the solution of $\frac{\partial L(\theta)}{\partial \theta} = 0$, $\bar{x}$ is in $[\frac{1}{3}, \frac{2}{3}]$, $\hat{\theta} = \bar{x}$.

- Case 3: $\bar{x} > 2/3$. In this case $\frac{\partial L(\theta)}{\partial \theta} > 0$, indicating $L(\theta)$ is an increasing function of θ. $\hat{\theta} = \frac{2}{3}$.

The MLE for θ is then

$$\hat{\theta} = \begin{cases} 1/3 & \text{if } \bar{x} < 1/3, \\ \bar{x} & \text{if } 1/3 \le \bar{x} \le 2/3, \\ 2/3 & \text{if } \bar{x} > 2/3, \end{cases}$$

Then by the invariant property of MLE, $\hat{\beta} = 3\hat{\theta} - 1$.

Additional problems

1. Let $X_1, \dots, X_n$ be a random sample (i.i.d.) from an exponential distribution, $\text{Exp}(\lambda)$, with rate λ.

(a) Find a method of moments estimator of λ using only the first moment.

Since $E(X_i) = \lambda^{-1}$, we set the sample mean $\bar{x} = \lambda^{-1}$ and obtain the method of moment estimator $\hat{\lambda}_1 = 1/\bar{x}$.

(b) Find a method of moments estimator using only the second moment.

Since $E(X_i^2) = 2/\lambda^2$, we set $\overline{x^2} = 2/\lambda^2$ and obtain the method of moment estimator $\hat{\lambda}_2 = \sqrt{\frac{2}{\overline{x^2}}}$.

(c) Find a third method of moments estimator using both the first and second moments.

Since $\text{var}(X_i) = E(X_i^2) - E(X_i)^2 = 1/\lambda^2$, we set the sample variance $\overline{x^2} - \bar{x}^2 = 1/\lambda^2$ and obtain $\hat{\lambda}_3 = 1/\sqrt{\overline{x^2} - \bar{x}^2}$.

(d) Find a method of moments estimator for $P(X_1 \ge 1)$.

$P(X_1 \ge 1) = e^{-\lambda}$. A method of moment estimator is $e^{-\hat{\lambda}_1} = \exp\{-1/\bar{x}\}$ (non-unique).

2. A Pareto distribution has c.d.f of the form

$$F(x|\theta_1, \theta_2) = 1 - (\frac{\theta_1}{x})^{\theta_2},\ \theta_1 \le x,\ \theta_1 > 0,\ \theta_2 > 0.$$

Find the MLE for (θ_1, θ_2).

The p.d.f is

$$f(x|\theta_1, \theta_2) = \begin{cases} \theta_2 \frac{\theta_1^{\theta_2}}{x^{\theta_2+1}} & \text{if } x \geq \theta_1 \\ 0 & \text{otherwise.} \end{cases}$$

and when $\min\{x_1, x_2, \ldots, x_n\} \geq \theta_1$ the likelihood function

$$L(\theta_1, \theta_2) = n \log \theta_2 + n\theta_2 \log \theta_1 - (\theta_2 + 1) \sum_{i=1}^{n} \log x_i.$$

It can be easily seen it is a monotonically increasing function of θ_1 and $\hat{\theta}_1 = \min\{x_1, x_2, \ldots, x_n\}$. Then we take derivative over θ_2 to get

$$\frac{\partial L(\theta_1, \theta_2)}{\partial \theta_2} = \frac{n}{\theta_2} - (\sum_{i=1}^{n} \log x_i - n \log \theta_1),$$

$$\frac{\partial L(\theta_1, \theta_2)^2}{\partial^2 \theta_2} = -\frac{n}{\theta_2^2} < 0.$$

So it is a concave function of θ_2 and the MLE is $\hat{\theta}_2 = \frac{n}{\sum_{i=1}^{n} \log x_i - n \log \hat{\theta}_1}$.

3. Show that if $\hat{\theta}$ is a method of moments estimator of θ, then for any one to one function g it holds that $g(\hat{\theta})$ is a method of moments estimator of $g(\theta)$.

By $\hat{\theta}$ is a method of moments estimator of θ we know that

$$\theta = h(\mu_1(\theta), \mu_2(\theta), \ldots, \mu_r(\theta))$$

and

$$\hat{\theta} = h(m_1, m_2, \ldots, m_r),$$

where $\mu_j(\theta) = EX^j$ and $m_j = \frac{1}{n} \sum_{i=1}^{n} X_i^j$.
Then we have

$$g(\theta) = g(h(\mu_1(\theta), \mu_2(\theta), \ldots, \mu_r(\theta)))$$

and

$$g(\hat{\theta}) = g(h(m_1, m_2, \ldots, m_r))$$

is a method of moments estimator of $g(\theta)$.