STAT 543 Homework 4 Solution

1.

Problem 2.1.2

Consider n systems with failure times $X_1,...,X_n$ assumed to be independent and identically distributed with exponential, $\Sigma(\lambda)$, distributions.

(a) Find the method of moments estimate of λ based on the first moment.
(b) Find the method of moments estimate of λ based on the second moment.
(c) Combine your answers to (a) and (b) to get a method of moment estimate of λ based on the first two moments.
(d) Find the method of moments estimate of the probability $P(X_1 \geq 1)$ that one system will last at least a month.

Solution:

Since $X_1,...,X_n$ $\underset{\sim}{i.i.d}$ $\mathrm{E}(\lambda)$, then

$$f(x_1) = \lambda e^{-\lambda x_1},\ x > 0$$

, and

$$E(X_1) = \int_0^\infty x\lambda e^{-\lambda x}dx = \frac{\Gamma(2)}{\lambda} = \frac{1}{\lambda}$$

, and

$$E(X_1^2) = \int\ x^2\lambda e^{-\lambda x}dx = \frac{\Gamma(3)}{\lambda^2} = \frac{2}{\lambda^2}.$$

(a) The method of moments estimate of λ based on the second moment is

$$\tilde{\lambda} = \frac{1}{\overline{x}} = \frac{n}{\sum_{i=1}^n x_i}$$

(b) The method of moments estimate of λ based on the second moment is

$$\tilde{\lambda} = \sqrt{\frac{2n}{\sum_{i=1}^n x_i^2}}$$

(c) From $\lambda^2 = 2\lambda^2 - \lambda^2 = \mu_2 - \mu_1^2$, the method of moments estimate of λ based on the first two moments is

$$\tilde{\lambda} = \frac{1}{\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2 - \overline{x}^2}} = \frac{1}{\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2 - (\frac{1}{n}\sum_{i=1}^n x_i)^2}}.$$

(d) Since $P(X_1 \geq 1) = \lambda\int_1^\infty e^{-\lambda x}dx = e^{-\lambda}$, then the method of moments estimate of $P(X_1 \geq 1)$ that one system will last at least a month is

$$\tilde{P}(X_1 \geq 1) = e^{-\overline{x}} = e^{-\frac{\sum_i x_i}{n}}.$$

Problem 2.1.3

Suppose that $i.i.d.\ X_1,...,X_n$ have a beta, $\beta(\alpha_1,\alpha_2)$ distribution. Find the method of moments estimates of $\alpha=(\alpha_1,\alpha_2)$ based on the first two moments.

Hint: See Problem B.2.5.

Solution:

Known that for $Beta(\alpha_1,\alpha_2)$ distribution, there exist

$$E(X)=\frac{\alpha_1}{\alpha_1+\alpha_2}$$

, and

$$E(X^2)=\frac{\alpha_1(\alpha_1+1)}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}$$

, then by

$$\begin{cases}\mu_1=\dfrac{\alpha_1}{\alpha_1+\alpha_2}\\ \mu_2=\dfrac{\alpha_1(\alpha_1+1)}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}\end{cases}\Rightarrow\begin{cases}\alpha_1=\dfrac{(\mu_2-\mu_1)\mu_1}{\mu_1^2-\mu_2}\\ \alpha_2=\dfrac{(\mu_1-\mu_2)(\mu_1-1)}{\mu_1^2-\mu_2}\end{cases}$$

Then, plug in $\frac{1}{n}\sum_1^n x_i$, and $\frac{1}{n}\sum_1^n x_i^2$ for μ_1 and μ_2, we get the method of moment estimate for $\alpha=(\alpha_1,\alpha_2)$ .as

$$\begin{cases}\tilde{\alpha}_1=\dfrac{(\frac{1}{n}\sum_i x_i^2-\frac{1}{n}\sum_i x_i)\frac{1}{n}\sum_i x_i}{(\frac{1}{n}\sum_i x_i)^2-\frac{1}{n}\sum_i x_i^2}\\ \tilde{\alpha}_2=\dfrac{(\frac{1}{n}\sum_i x_{i_1}-\frac{1}{n}\sum_i x_i^2)(\frac{1}{n}\sum_i x_i-1)}{(\frac{1}{n}\sum_i x_i)^2-\frac{1}{n}\sum_i x_i^2}\end{cases}.$$

Problem 2.1.11

In Example 2.1.2 with $X\sim\Gamma(\alpha,\lambda)$, find the method of moments estimate based on $\hat{\mu}_1$ and $\hat{\mu}_3$.

Hint: See Problem B.2.4.

Solution:

For $\text{gamma}(\alpha, \lambda)$ distribution, $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, x > 0$

$$\Rightarrow \begin{cases} \text{Ex} = \mu_1 = \frac{\alpha}{\lambda} \\ \text{Ex}^3 = \mu_3 = \frac{\Gamma(\alpha+3)}{\Gamma(\alpha)\lambda^3} = \frac{\alpha(\alpha+1)(\alpha+2)}{\lambda^3} \end{cases}$$

$$\Rightarrow \quad \alpha = \frac{-3\mu_1^3 + \sqrt{\mu_1^6 + 8\mu_1^3\mu_3}}{2(\mu_1^3 - \mu_3)}$$

, and $\lambda = \frac{-3\mu_1^2 + \sqrt{\mu_1^4 + 8\mu_1\mu_3}}{2(\mu_1^3 - \mu_3)}$

By plug in $\frac{1}{n}\sum_1^n x_i$ and $\frac{1}{n}\sum_1^n x_i^3$ for μ_1 and μ_3, we get the MOM estimates for α and λ as

$$\hat{\alpha} = \frac{-3\left(\frac{1}{n}\sum_1^n x_i\right)^3 + \sqrt{\left(\frac{1}{n}\sum_1^n x_i\right)^6 + 8\left(\frac{1}{n}\sum_1^n x_i\right)^3 (\frac{1}{n}\sum_1^n x_i^3)}}{2(\left(\frac{1}{n}\sum_1^n x_i\right)^3 - \frac{1}{n}\sum_1^n x_i^3)}$$

, and

$$\hat{\lambda} = \frac{-3\left(\frac{1}{n}\sum_1^n x_i\right)^2 + \sqrt{\left(\frac{1}{n}\sum_1^n x_i\right)^4 + 8\left(\frac{1}{n}\sum_1^n x_i\right)\left(\frac{1}{n}\sum_1^n x_i^3\right)}}{2(\left(\frac{1}{n}\sum_1^n x_i\right)^3 - \left(\frac{1}{n}\sum_1^n x_i^3\right))}.$$

Problem 2.2.10

Let $X_1, ..., X_n$ denote a sample from a population with one of the following densities or frequency functions. Find the MLE of θ.

(a) $f(x,\theta) = \theta e^{-\theta x}, x \ge 0; \theta > 0$. (Exponential density)

(b) $f(x,\theta) = \theta c^\theta x^{-(\theta+1)}, x \ge c$; c constant>0; $\theta > 0$. (Pareto density)

(c)

<u>Solution:</u>

(a).

$$f(x \mid \theta) = \theta^n e^{-\theta \sum x_i} = L(\theta)$$
$$\ell(\theta) = n\log(\theta) - \theta \sum x_i$$

$$\ell' = \frac{d\ell}{d\theta} = \frac{n}{\theta} - \sum x_i = 0 \Rightarrow \hat{\theta} = \frac{\sum x_i}{n}$$

Check that $\ell'' = -\frac{n}{\theta^2} < 0$, which indicates that $\hat{\theta}$ reaches the maximum of ℓ.

Hence, $\hat{\theta} = \frac{\sum x_i}{n}$ is the MLE of θ.

(b).

$$f(x \mid \theta) = \theta^n c^{n\theta} \prod x_i^{-(\theta+1)}$$

$$\ell(\theta) = n \log \theta + n\theta \log c - (\theta+1)\sum_1^n \log x_i$$

$$\ell' = \frac{d\ell}{d\theta} = \frac{n}{\theta} + n \log c - \sum \log x_i = 0$$

$$\Rightarrow \hat{\theta} = \frac{\sum_1^n \log x_i - n \log c}{n}$$

Check that $\ell'' = -\frac{n}{\theta^2} < 0$, which indicates that $\hat{\theta}$ reaches the maximum of ℓ.

Hence, $\hat{\theta} = \frac{\sum x_i}{n}$ is the MLE of θ.

(c)

$$f(x \mid \theta) = c^n \theta^{nc} \prod x_i^{-(c+1)} \cdot 1\left[\min(x_i) \geq \theta\right], c > 0; \theta > 0$$

$\Rightarrow \hat{\theta} = \min(x_i)$ is the MLE of θ, since θ^{nc} is an increasing function of θ.

(d)

$$f(x \mid \theta) = \theta^{\frac{n}{2}} \prod x_i^{\sqrt{\theta}-1}$$

$$\ell = \frac{n}{2} \log \theta + (\sqrt{\theta} - 1) \sum \log x_i$$

$$\ell' = \frac{n}{2\theta} + \frac{1}{2\sqrt{\theta}} \sum \log x_i = 0$$

$$\Rightarrow \hat{\theta} = \left(\frac{\sum \log x_i}{n} \right)^{-2}$$

Check that $\ell'' = -\frac{n}{2\theta^2} - \frac{1}{4\theta^{\frac{3}{2}}} \sum_i \log x_i < 0$.

So, $\hat{\theta} = \left(\frac{\sum \log x_i}{n} \right)^{-2}$ is the MLE of θ.

(e)

$$f(x \mid \theta) = \theta^{-2n} \left(\prod x_i \right) \cdot \exp\left(-\frac{\sum x_i^2}{2\theta^2} \right)$$

$$\ell = -2n\log\theta - \frac{\sum x_i^2}{2\theta^2} + \sum_i \log x_i$$
$$\ell' = -\frac{2n}{\theta} + \frac{\sum x_i^2}{\theta^3} = 0$$
$$\Rightarrow \hat{\theta} = \sqrt{\frac{\sum x_i^2}{2n}}$$

Check that $\ell'' = \frac{2n}{\theta^2} - \frac{3\sum_i x_i^2}{\theta^4}$ and $\ell''(\hat{\theta}) = \frac{2n}{\hat{\theta}^2} - \frac{3\sum_i x_i^2}{\hat{\theta}^4} = \frac{4n^2}{\sum_i x_i^2} - \frac{12n^2}{\sum_i x_i^2} = -\frac{8n^2}{\sum_i x_i^2} < 0$.

Hence, $\hat{\theta} = \sqrt{\frac{\sum x_i^2}{2n}}$ is the MLE of θ.

(f).

$$f(x \mid \theta) = \theta^n c^n \prod x_i^{-(c-1)} \exp\left(-\theta \sum x_i^c\right)$$
$$l = n\log(\theta) - \theta\sum x_i^c + n\log c - (c-1)\sum_i \log x_i$$
$$\ell' = \frac{n}{\theta} - \sum x_i^c = 0 \Rightarrow \hat{\theta} = \frac{n}{\sum x_i^c}$$

Check that $\ell'' = -\frac{n}{\theta^2} < 0$.

Hence, $\hat{\theta} = \frac{n}{\sum x_i^c}$ is the MLE of θ.

Problem 2.2.14

If $n = 1$ in Example 2.1.5 show that no maximum likelihood estimate of $\theta = (\mu, \sigma^2)$ exists.

Solution:

When $n = 1$, $L(x,\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$,

$$\ell(x,\theta) = -\ln\sqrt{2\pi} - \frac{1}{2}\ln\sigma^2 - \frac{(x-\mu)^2}{2\sigma^2}.$$

Since $\Theta = \{(u,\sigma^2), u \in \mathbb{R}, \sigma^2 > 0\}$ is an open set in $\mathbb{R}^2$, the maximum can not be attained at the bound of Θ.

$$\frac{\partial \ell(x,\theta)}{\partial \sigma^2} = -\frac{1}{2}\frac{1}{\sigma^2} + \frac{(x-\mu)^2}{2\eta^2} = 0 \qquad (1)$$

$$\frac{\partial \ell(x,\theta)}{\partial \mu} = -\frac{(x-\mu)}{\sigma^2} = 0 \qquad (2)$$

From (2), it requires $x = \mu$, however (1) cannot be solved if $x = \mu$. So, there is no maximum in the interior of Θ . Hence, there is not MLE of $\theta = (\mu, \sigma^2)$ if n=1.

Or you can think this way:

Let $\hat{\mu} = x_1$, the $f(x \mid \theta) = \dfrac{1}{\sqrt{2\pi}\sigma} \to \infty$ when $\sigma \to 0$.

So no maximum of *likelihood* function exists, so no MLE of $\theta = (\mu, \sigma^2)$ exists.

Problem 2.2.17 Censored Geometric Waiting Times

If time is measured in discrete periods, a model that is often used for the time X to failure of an item is

$$P_\theta[X = k] = \theta^{k-1}(1-\theta), k = 1, 2, \ldots \quad \text{, where } 0 < \theta < 1 .$$

Suppose that we only record the time of failure, if failure occurs on or before time r an otherwise just note that the item has lived at least $(r+1)$ periods. Thus, we observe $Y_1, \ldots, Y_n$ which are independent, identically distributed, and have common frequency function,

$$f(k, \theta) = \theta^{k-1}(1-\theta), k = 1, \ldots, r$$

$$f(r+1, \theta) = 1 - P_\theta[X \le r] = 1 - \sum_{k=1}^{r} \theta^{k-1}\left(1-\theta\right) = \theta^r .$$

(We denote by “r+1” survival for at least (r+1) periods.) Let M=number of indices I such that $Y_i = r+1$. Show that the maximum likelihood estimate of θ based on $Y_1, \ldots, Y_n$ is

$$\hat{\theta}(Y) = \frac{\sum_{i=1}^{n} Y_i - n}{\sum_{i=1}^{n} Y_i - M} .$$

<u>Solution:</u>

Note that $f(r+1 \mid \theta) = \theta^r = \theta^{y_i - 1}$ if $y_i = r+1$. Then the *likelihood* function

$$L = \theta^{\sum_1^M (y_i - 1)} (1-\theta)^{n-M} \, \theta^{\sum_{M+1}^n (y_i - 1)}$$

$$= \theta^{\sum_1^n (y_i - 1)} (1-\theta)^{n-M}$$

$$\ell = \sum_{i=1}^{n} (y_i - 1) \log \theta + (n - M) \log(1-\theta)$$

$$\ell' = \frac{\sum_{i=1}^{n} (y_i - 1)}{\theta} + \frac{(M-n)}{1-\theta} = 0$$

$$\Rightarrow \hat{\theta} = \frac{\sum_{i=1}^{n} y_i - n}{\sum_{i=1}^{n} y_i - M}$$

Check that $\ell'' = -\frac{\sum_i (y_i - 1)}{\theta^2} + \frac{(M-n)}{(1-\theta)^2}$

, and then $\ell''(\hat{\theta}) = -\frac{\sum_i (y_i - 1)}{\hat{\theta}^2} + \frac{(M-n)}{(1-\hat{\theta})^2} = \frac{(\sum_{i=1}^{n} y_i - M)^3}{(M-n)(\sum_{i=1}^{n} y_i - n)} < 0$

Thus, $\hat{\theta} = \frac{\sum_{i=1}^{n} y_i - n}{\sum_{i=1}^{n} y_i - M}$ is the MLE of θ.

Problem 2.2.23

Let $X_1, ..., X_m$ and $Y_1, ..., Y_n$ be two independent samples from $N\left(\mu_1, \sigma^2\right)$ and $N\left(\mu_2, \sigma^2\right)$ populations, respectively. Show that the MLE of $\theta = \left(\mu_1, \mu_2, \sigma^2\right)$ is $\hat{\theta} = (\bar{X}, \bar{Y}, \tilde{\sigma}^2)$ where

$$\tilde{\sigma}^2 = \frac{[\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{j=1}^{n}(Y_j - \bar{Y})^2]}{m+n}.$$

<u>Solution:</u>

$$L(x, y \mid \theta) = \exp\left(-\frac{\sum (x_i - \mu_1)^2}{2\sigma^2} - \frac{\sum (y_i - \mu_2)^2}{2\sigma^2}\right)\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{m+n}$$

$$l = -\frac{\sum (x_i - \mu_1)^2}{2\sigma^2} - \frac{\sum (y_i - \mu_2)^2}{2\sigma^2} - \frac{m+n}{2}\log(2\pi\sigma^2)$$

$$\frac{\partial l}{\partial \mu_1} = 0 \Rightarrow \hat{\mu}_1 = \bar{x}$$

$$\frac{\partial l}{\partial \mu_2} = 0 \Rightarrow \hat{\mu}_2 = \bar{y}$$

$$\frac{\partial l}{\partial \sigma^2} = 0 \Rightarrow \frac{\sum (x_i - \mu_1)^2}{2(\sigma^2)^2} + \frac{\sum (y_i - \mu_2)^2}{2(\sigma^2)^2} - \frac{\frac{m+n}{2}}{\sigma^2} = 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\left[\sum_1^m (x_i - \bar{x})^2 + \sum_1^n (y_i - \bar{y})^2\right]}{m+n}$$

$| l'' | > 0 \Rightarrow \hat{\theta} = (\bar{x}, \bar{y}, \hat{\sigma}^2)$ is the MLE of $\theta = (\mu_1, \mu_2, \sigma^2)$

2. (Zero-inflated Poisson model)
Condsider a marginal pmf

$$f(x \mid \lambda) = pI[x=0] + (1-p)\frac{\exp(-\lambda)\lambda^x}{x!} \; forx = 0,1,2,...$$

(This is a Poisson distribution with s0me "extra mass" at 0 or alternatively, amixture of a Poisson distribution and a point mass at 0). Suppose parameters are $p \in [0,1]$ and $\lambda \ge 0$. Find $E_{p,\lambda} X$ and $E_{p,\lambda} X^2$. Then, for $X_1, X_2, ..., X_n \, i.i.d.$ with this marginal distribution, find a mothod of moments estimator fro the parameter vector (p, λ) based on the first two sample moments.

<u>Solution:</u>

$$\text{Ex} = \sum_1^\infty (1-p)\frac{e^{-\lambda}\lambda^x x}{x!} = (1-p)\lambda = \mu_1$$

$$\text{Ex}^2 = \sum_1^\infty (1-p)\frac{e^{-\lambda}\lambda^x x^2}{x!} = (1-p)(\lambda + \lambda^2) = \mu_2$$

Then,

$$\lambda = \frac{\mu_2}{\mu_1} - 1, \; p = 1 - \frac{\mu_1^2}{\mu_2 - \mu_1}$$

And by plugging in $\hat{\mu}_1 = \frac{1}{n}\sum_1^n x_i$, and $\hat{\mu}_2 = \frac{1}{n}\sum_1^n x_i^2$, we get the MOM e of $\theta = (\lambda, P)$ as

$$\hat{\lambda} = \frac{\frac{1}{n}\sum_{i=1}^n x_i^2 - \frac{1}{n}\sum_{i=1}^n x_i}{\frac{1}{n}\sum_{i=1}^n x_i}$$

, and

$$\hat{p} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\sum_{i=1}^{n} x_i - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\sum_{i=1}^{n} x_i}.$$

3. (Optional)
Problem 2.1.14
When the data are not i.i.d., it may still be possible to express parameters as functions of moments and then use estimates based on replacing population moments with "sample" moments. Consider the Gaussian $AR(1)$ model of Example 1.1.5.

(a) Use $E(X_i)$ to give a method of moments estimate of μ.

(b) Suppose $\mu = \mu_0$ and $\beta = b$ are fixed. Use $E(U_i^2)$, where

$$U_i = \frac{(X_i - \mu_0)}{\sqrt{\sum_{j=0}^{i-1} b^{2j}}}$$

, to give a method of moments estimate of σ^2.

(c) If μ and σ^2 are fixed, can you five a method of moments estimate of β?

<u>Solution:</u>
(a).

$$EX_i = E\mu + Ee_i = \mu + Ee_i$$

$$Ee_i = \beta Ee_{i-1}E\varepsilon_i = \beta Ee_{i-1} = \cdots = \beta^i Ee_0 = 0$$

So,

$$Ee_i = \mu, i = 1, ..., n$$

Plug in $\frac{1}{n}\sum_{i=1}^{n} x_i$ for μ, and then get the MOM for μ is

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

(b).

$$\varepsilon_i = x_i - \mu(1-\beta) - \beta(x_i - 1)$$

$$= x_i - \mu(1-\beta) - \beta\left(\mu(1-\beta) + \beta \cdot x_{i-2} + \varepsilon_{i-1}\right)$$

$$= \dots$$

$$\varepsilon_i + \beta_1\varepsilon_{i-1} + \beta^2\varepsilon_{i-2} + \dots = x_i - \mu$$

$\beta = b, \qquad \mu = \mu_0$, then

$$\Rightarrow \mathrm{VAR}(\mathrm{LHS}) = \sigma^2\left(1+\beta^2+\beta^4+\cdots\right) = \sigma^2\left(\sum_{j=1}^{i-1} b^{2j}\right)$$

$$\text{So, } E\left(U_i^{\,2}\right) = E\left\{\frac{x_i-\mu}{\left(\sum_{j=0}^{i-1} b^{2j}\right)^{1/2}}\right\}^2 = \frac{E(x_i-\mu)^2}{\left(\sum_{j=0}^{i-1} b^{2j}\right)} = \sigma^2$$

, and the MOM estimates of σ^2 is $\frac{1}{n}\sum U_i^2$.

, where

$$U_i = (x_i - \mu_0)\Big/\left(\sum_{j=0}^{i-1} b^{2j}\right)^{1/2}$$

(c).

Since $\gamma(1) = \beta\gamma(0) \Rightarrow \hat{\beta} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)}$

$$\hat{\gamma}_1 = \frac{\sum_{i=2}^{n}\left(x_i - \overline{x}\right)\left(x_{i-1} - \overline{x}\right)}{n-2}, \ \hat{\gamma}_0 = \hat{\sigma}^2 .$$

So, a MOM estimate of β is $\frac{\hat{\gamma}_1}{\hat{\sigma}^2}$.

Problem 2.2.15

Suppose that $T(X)$ is sufficient for θ and that $\hat{\theta}(X)$ is an MLE of θ. Show that $\hat{\theta}$ depends on X through $T(X)$ only provided that $\hat{\theta}$ is unique.

Hint: Use the factorization theorem (Theorem 1.5.1)

<u>Solution:</u>

By factorization theorem $f\left(x \mid \theta\right) = g\left(T(x), \theta\right)\ln(x)$, so

$$R = \frac{f\left(x \mid \theta\right)}{f\left(x \mid \theta_0\right)} = \frac{g\left(T(x), \theta\right)}{g\left(T(x), \theta_0\right)}$$

If $\hat{\theta}$ maximize *L*, then $\hat{\theta}$ also maximize *R*, from the above equation, we can see that $\hat{\theta}$ depends on *X* only through *T(X)* by $g\left(T(x), \theta\right)$.

Problem 2.2.16

(a) Let $X \sim P_\theta, \theta \in \Theta$ and let $\hat{\theta}$ denote the MLE of θ. Suppose that h is a one-to-one function from Θ onto $h(\Theta)$. Define $\eta = h(\theta)$ and let $f(x,\eta)$ denote the density or frequency function of X in terms of η (i.e., reparameterize the model using η). Show that the MLE of η is $h(\hat{\theta})$ (i.e., MLEs are unaffected by reparameterization, they are equivariant under one-to-one transformations).

(b) Let $P = \{P_\theta : \theta \in \Theta\}, \Theta \subset \mathbb{R}^p, p \geq 1$, be a family of models for $X \in \aleph \subset \mathbb{R}^d$. Let q be a map from Θ onto Ω, $\Omega \subset \mathbb{R}^k, 1 \leq k \leq p$. Show that if $\hat{\theta}$ is a MLE of θ, then $q(\hat{\theta})$ is an MLE of $\omega = q(\theta)$.

<u>Solution:</u>

(a)

Since h is one-to-one function from Θ onto $h(\Theta)$, then h is inversable, i.e. h^{-1} exists.

And thus $f(x,\theta)$ can also be written as a form of $g(x,\eta) \hat{=} f(x, h^{-1}(\eta))$.

Suppose that $\hat{\theta}$ is MLE, $\forall \theta \in \Theta$, $f(x,\hat{\theta}) \geq f(x,\theta)$, i.e. $f(x, h^{-1}(\hat{\eta})) \geq f(x, h^{-1}(\eta))$

$\Rightarrow g(x,\hat{\eta}) \geq g(x,\eta), \forall \eta \in h(\theta)$, where $\hat{\eta} = h(\hat{\theta})$. So $\hat{\eta}$ is the MLE.

(b)

Let $\Theta(\omega) = \{\theta \in \Theta : q(\theta) = \omega\}$, then $\{\Theta(\omega) : \omega \in \Omega\}$ is a portion of Θ, and also $\hat{\theta}$ must in one $\Theta(\omega)$, denote $\Theta(\hat{\omega})$, that is, $q(\Theta(\hat{\omega})) = q(\hat{\theta}) = \hat{\omega}$.

Let

$$\hat{w}_{MLE} = \arg\sup_{w \in \Omega} \sup\{L_X(\theta), \theta \in \Theta(\omega)\}$$

, since

$$\sup_{\omega \in \Omega} \sup\{L_X(\theta) : \theta \in \Theta(w)\} = L_X(\hat{\theta})$$

, then

$$\hat{w}_{MLE} = q(\Theta(\hat{\omega})) = q(\hat{\theta}) = \hat{w}$$

i.e. $\hat{w}_{MLE} = \hat{\omega} = q(\hat{\theta})$.

Problem 2.2.19

Let $X_1, ..., X_n$ be independently distributed with X_i having a $N(\theta_i, 1)$ distribution, $1 \leq i \leq n$.

(a) Find maximum likelihood estimates of the assumption that these quantities vary freely.

(b) Solve the problem of part (a) for $n = 2$ when it is known that $\theta_1 \leq \theta_2$. A general solution of this and related problems may be found in the book by Barlow, Bartholomew, Bremner, and Brunk (1972).

Solution:

(a).

$$f(x \mid \theta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{\sum (x_i - \theta_i)^2}{2}\right)$$

So, MLE of θ is $\hat{\theta}_i = x_i$.

(b).

$x_1 \le x_2 \Rightarrow$ The MLE of θ_1, θ_2 is $\hat{\theta}_1 = x_1$, $\hat{\theta}_2 = x_2$.

If $x_1 > x_2$, since $a^2 + b^2 \le 2ab$; and the equality holds if a=b.

So, the MLE of θ_1, θ_2 is $\theta_1 = \theta_2 = \dfrac{x_1 + x_2}{2}$.

Problem 2.2.21 (Kiefer-Wolfowitz)

Suppose $(X_1,..., X_n)$ is a sample from a population with density

$$f(x,\theta) = \frac{9}{10\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right) + \frac{1}{10}\varphi(x-\mu)$$

, where φ is the standard normal density and

$$\theta = (\mu, \sigma^2) \in \Theta = \left\{(\mu, \sigma^2) : -\infty < \mu < \infty, 0 < \sigma^2 < \infty\right\}.$$

Show that maximum likelihood estimates do not exist, but that

$$\sup_\sigma p(x, \hat{\mu}, \sigma^2) = \sup_{\mu,\sigma} p(x, \mu, \sigma^2)$$

if, and only if, $\hat{\mu}$ equals one of the numbers $x_1,..., x_n$. Assume that $x_i \ne x_j$ for $i \ne j$ and that $n \ge 2$.

Solution:

Let $\hat{\mu} = x_i, \qquad 1 \le i \le n$

(a).

Consider $\sigma^2 \to 0$ across line $\ell = \{(x_1, \sigma^2) : \sigma^2 > 0\} \subset \Theta$, then

$$f_1(x_1, \theta) = \frac{1}{\sqrt{2\pi}}\left(\frac{9}{10}\frac{1}{\sigma} + \frac{1}{10}\right) \to \infty$$

, as $\sigma^2 \to 0$. And also

$$f_i(x_i, \theta) = \frac{1}{\sqrt{2\pi}}\left(\frac{9}{10}\frac{1}{\sigma}e^{-\frac{(x_1-x_i)^2}{2\sigma^2}} \frac{1}{10}e^{-\frac{(x_1-x_i)^2}{2}}\right) \to \frac{1}{\sqrt{2\pi}}\frac{1}{10}e^{-\frac{(x_1-x_i)^2}{2}}$$

, as $\sigma^2 \to 0$. So,

$L(x \mid \theta) \to \infty$, when $\sigma \to 0$.

And thus the likelihood function is unbounded, which means that the MLE $\left(\mu,\sigma^2\right)$ in Θ doesn't exist.

(b).

From above, we have

$$\sup_{\sigma} p\left(x,\hat{\mu},\sigma\right) = s\,\text{up}_{\sigma,\mu}\, p\left(x,\mu,\sigma^2\right) = \infty$$

If $\sup_{\sigma} p\left(x,\mu,\sigma\right) = 0$, then $\hat{\mu}$ must equal to some x_i.

If $\hat{\mu} \neq x_{i,} \quad \forall \quad i.\text{st} \quad 1 \le i \le n$, then by

$$\frac{1}{\sigma}\exp\left(-\frac{M}{\sigma}\right) \to 0$$

, as $\sigma \to 0$, where

$$M = \max_i \left\{\left(x_i - \mu\right)^2\right\},$$

, and hence

$$\sup_{\sigma} p\left(x,\hat{\mu},\sigma\right) < \infty, \text{ if } \hat{\mu} \neq x_i.$$

So, $\hat{\mu}$ must equal to some x_i.