

Statistical Theory - Solutions 1

Spring 2023. Email corrections to Kolyan Ray: kolyan.ray@ic.ac.uk

1. (i)

$$f_{a,b}(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} = \exp \{a \log x + b \log(1-x) - \log B(a,b) - \log x - \log(1-x)\},$$

so in the notation of the lecture notes, $c(\theta) = (a, b)$, $T(x) = (\log x, \log(1-x))$, $d(\theta) = \log B(a,b)$ and $S(x) = -\log x - \log(1-x)$.

(ii)

$$f_p(x) = p(1-p)^{x-1} = \exp \{x \log(1-p) + \log p - \log(1-p)\},$$

so $c(\theta) = \log(1-p)$, $T(x) = x$, $d(\theta) = \log((1-p)/p)$ and $S(x) = 0$.

(iii)

$$f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} = \exp \{\alpha \log x - \beta x + \alpha \log \beta - \log \Gamma(\alpha) - \log x\},$$

so $c(\theta) = (\alpha, \beta)$, $T(x) = (\log x, -x)$, $d(\theta) = -\alpha \log \beta + \log \Gamma(\alpha)$ and $S(x) = -\log x$.

2. The first moment is $E_\theta X_i = \theta/2$, so the method of moments estimator equates $\hat{\theta}_{MM}/2 = \bar{X}_n$, i.e. $\hat{\theta}_{MM} = 2\bar{X}_n$. It satisfies $E_\theta \hat{\theta}_{MM} = 2E_\theta X_1 = \theta$ and $\text{Var}_\theta(\hat{\theta}_{MM}) = 4\text{Var}_\theta(\bar{X}_n) = 4\text{Var}_\theta(X_1)/n$. Since $\text{Var}_\theta(X_1) = \theta^2/12$, this yields $\text{Var}_\theta(\hat{\theta}_{MM}) = \theta^2/(3n)$.

The likelihood is

$$\prod_{i=1}^n \frac{1}{\theta} \mathbb{1}_{[0,\theta]}(X_i) = \frac{1}{\theta^n} \mathbb{1}_{[0,\theta]}(\max_i X_i) = \frac{1}{\theta^n} \mathbb{1}\{0 \leq \max_i X_i \leq \theta\}.$$

This is a decreasing function of θ as long as the indicator function is one, so pick the smallest θ such that the indicator is 1, i.e. $\hat{\theta}_{ML} = \max_i X_i$. We have distribution function

$$P_\theta(\max_i X_i \leq t) = \prod_{i=1}^n P_\theta(X_i \leq t) = (t/\theta)^n,$$

and differentiating gives the pdf of $\max_i X_i$ as $g_\theta(t) = nt^{n-1}/\theta^n$ for $t \in [0, \theta]$. This yields $E_\theta \hat{\theta}_{ML} = \int_0^\theta nt^n/\theta^n dt = \frac{n}{n+1}\theta$ and $E_\theta \hat{\theta}_{ML}^2 = \int_0^\theta nt^{n+1}/\theta^n d\theta = \frac{n}{n+2}\theta^2$, which implies

$$\text{Var}_\theta(\hat{\theta}_{ML}) = E_\theta \hat{\theta}_{ML}^2 - (E_\theta \hat{\theta}_{ML})^2 = \frac{n}{(n+1)^2(n+2)}\theta^2.$$

Using the bias-variance decomposition, we compute the mean-squared errors for the estimators as $\text{MSE}_\theta(\hat{\theta}_{MM}) = \theta^2/(3n)$ and

$$\text{MSE}_\theta(\hat{\theta}_{ML}) = \left(\theta - \frac{n}{n+1}\theta\right)^2 + \frac{n}{(n+1)^2(n+2)}\theta^2 = \frac{2\theta^2}{(n+1)(n+2)}.$$

We see that $\text{MSE}_\theta(\hat{\theta}_{ML}) < \text{MSE}_\theta(\hat{\theta}_{MM})$ for all $\theta > 0$ and $n \geq 3$ (with equality if $n = 1, 2$). Thus while $\hat{\theta}_{ML}$ is biased, it has smaller MSE than the unbiased $\hat{\theta}_{MM}$.

3. In all cases, we will expand out the likelihood and use the factorization criterion to establish sufficiency.

(a)

$$\prod_{i=1}^n \frac{1}{B(a, b)} x_i^{a-1} (1-x_i)^{b-1} = \frac{1}{B(a, b)^n} \left(\prod_{i=1}^n x_i \right)^{a-1} \left(\prod_{i=1}^n (1-x_i) \right)^{b-1}.$$

We can rewrite the RHS as $g(\prod x_i, \prod(1-x_i), a, b)$ (taking $h(x) = 1$), which gives that $T(x) = (\prod x_i, \prod(1-x_i))$ is sufficient for (a, b) .

(b)

$$\prod_{i=1}^n \frac{1}{B(a, a)} x_i^{a-1} (1-x_i)^{a-1} = \frac{1}{B(a, a)^n} \left[\prod_{i=1}^n x_i (1-x_i) \right]^{a-1}.$$

We can rewrite the RHS as $g(\prod x_i (1-x_i), a)$ (taking $h(x) = 1$), which gives that $T(x) = \prod x_i (1-x_i)$ is sufficient for a .

(c)

$$\prod_{i=1}^n \frac{1}{\sigma} e^{-(x_i - \mu)/\sigma} 1_{[\mu, \infty)}(x_i) = \frac{1}{\sigma^n} e^{n\mu/\sigma} e^{-\frac{1}{\sigma} \sum_i x_i} 1_{[\mu, \infty)}(\min_i x_i),$$

where we have used that $\prod_{i=1}^n 1_{[\mu, \infty)}(x_i) = 1_{[\mu, \infty)}(\min_i x_i)$. Rewriting the RHS as $g(\min_i x_i, \sum_i x_i, \mu, \sigma)$ (taking $h(x) = 1$), we have that $(\min_i x_i, \sum_i x_i)$ is sufficient for (μ, σ) .

(d)

$$\prod_{i=1}^n \frac{\beta^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i} = \left(\frac{\beta^\alpha}{\Gamma(\alpha)} \right)^n \left(\prod_{i=1}^n x_i \right)^{\alpha-1} e^{-\beta \sum_i x_i}.$$

Rewrite the RHS as $g(\prod x_i, \sum x_i, \alpha, \beta)$ (taking $h(x) = 1$), which gives that $T(x) = (\prod x_i, \sum x_i)$ is sufficient for (α, β) .

4. (i)

$$f_\theta(x) = \prod_{i=1}^n f_{X_i, \theta}(x_i) = \prod_{i=1}^n e^{i\theta - x_i} 1\{\theta \leq \frac{x_i}{i}\} = e^{\theta \sum i - \sum x_i} 1\{\theta \leq \min_i \frac{x_i}{i}\}.$$

Writing $g(T, \theta) = e^{\theta \sum i} 1\{\theta \leq \min_i \frac{x_i}{i}\}$ and $h(x) = e^{-\sum x_i}$, the factorization criterion gives that $T = \min_i(X_i/i)$ is a sufficient statistic for θ .

(ii) Similarly to the last example,

$$\begin{aligned} f_\theta(x) &= \prod_{i=1}^n \frac{1}{2i\theta} 1\{-i(\theta-1) < x_i < i(\theta+1)\} \\ &= \frac{1}{(2\theta)^n} \left(\prod_{i=1}^n \frac{1}{i} \right) 1\{-(\theta-1) \leq \min_i \frac{x_i}{i}\} 1\{\max_i \frac{x_i}{i} < (\theta+1)\}. \end{aligned}$$

By the factorization criterion, $T = (\min_i(X_i/i), \max_i(X_i/i))$ is a sufficient statistic for θ .

5. From the definition of sufficiency, we must show that the conditional distribution of X_1, \dots, X_n given $T(x)$ depends on θ . Working out the conditional distribution explicitly,

$$P_\theta(X_1 = x_1, \dots, X_n = x_n | T = t) = \begin{cases} \frac{P_\theta(X_1 = x_1, \dots, X_n = x_n)}{P_\theta(T=t)} & \text{if } \sum_{i=1}^{n/2} x_i = t, \\ 0 & \text{otherwise.} \end{cases}$$

The second term is independent of θ , so consider the first case. Writing $s = \sum_{i=n/2+1}^n x_i$, the conditional probability equals

$$\frac{P_\theta(X_1 = x_1, \dots, X_n = x_n)}{P_\theta(T = t)} = \frac{\theta^{\sum_i x_i} (1-\theta)^{n-\sum_i x_i}}{\binom{n/2}{t} \theta^t (1-\theta)^{n/2-t}} = \frac{\theta^{s+t} (1-\theta)^{n-s-t}}{\binom{n/2}{t} \theta^t (1-\theta)^{n/2-t}} = \frac{\theta^s (1-\theta)^{n/2-s}}{\binom{n/2}{t}}.$$

This depends on θ for $s > 0$, and hence T is not sufficient for θ .

6. For minimal sufficiency, we consider the ratio

$$\frac{f_\theta(x)}{f_\theta(x')} = \frac{\prod_{i=1}^n \frac{(i\theta)^{x_i}}{x_i!} e^{-i\theta}}{\prod_{i=1}^n \frac{(i\theta)^{x'_i}}{x'_i!} e^{-i\theta}} = \left(\prod_{i=1}^n i^{x_i - x'_i} \frac{x'_i!}{x_i!} \right) \theta^{\sum x_i - \sum x'_i}.$$

This ratio is constant as a function of θ if and only if $\sum x_i = \sum x'_i$, and so $T(X) = \sum_i X_i$ is minimal sufficient for θ by a theorem in the lecture notes. By the independence, $T(X) \sim \text{Poisson}(\theta \sum_{i=1}^n i) = \text{Poisson}(n(n+1)\theta/2)$.

For the MLE, the log-likelihood equals

$$\ell_n(\theta) = \log \left(\prod_{i=1}^n \frac{(i\theta)^{x_i}}{x_i!} e^{-i\theta} \right) = \log \theta \sum_{i=1}^n x_i - \theta \sum_{i=1}^n i + \log h(x),$$

for some function $h(x)$ that does not depend on θ . If $\sum x_i > 0$, differentiating this twice,

$$\ell'_n(\theta) = \frac{1}{\theta} \sum_i x_i - \sum_i i, \quad \ell''_n(\theta) = -\frac{1}{\theta^2} \sum_i x_i.$$

Solving $\ell'_n(\theta) = 0$ gives $\hat{\theta} = (\sum_i x_i)/(\sum_i i)$, which is the MLE since $\ell''_n(\theta) < 0$ for all $\theta > 0$. It is unbiased since $E_\theta \hat{\theta} = E_\theta T(X)/\sum_i i = \theta$.

Note that if $\sum_i x_i = 0$, no MLE exists since then $\ell_n(\theta)$ is strictly increasing as $\theta \downarrow 0$, but $\theta = 0$ is not in the parameter space.

- 7.** (a) We have $E_\theta X_1 = \int_0^\infty x \theta e^{-\theta x} dx = 1/\theta$. Thus the method of moments estimator satisfies $1/\theta = \bar{X}_n$, i.e. $\hat{\theta}_{MM} = 1/\bar{X}_n = n/(\sum_{i=1}^n X_i)$.
- (b) Consider the ratio

$$\frac{f_\theta(x)}{f_\theta(x')} = \frac{\prod_{i=1}^n \theta e^{-\theta x_i}}{\prod_{i=1}^n \theta e^{-\theta x'_i}} = e^{\theta(\sum_i x'_i - \sum_i x_i)}.$$

This ratio does not depend on θ if and only if $\sum_i x_i = \sum_i x'_i$, and hence $T(X) = \sum_{i=1}^n X_i$ is minimal sufficient for θ . We have that $T(X) \sim \text{Gamma}(n, \theta)$. The log-likelihood equals

$$\ell_n(\theta) = \log \left[\theta^n e^{-\theta \sum_i x_i} \right] = n \log \theta - \theta \sum_i x_i.$$

Differentiating gives $\ell'_n(\theta) = n/\theta - \sum_i x_i$ and $\ell''_n(\theta) = -n/\theta^2$. Solving $\ell'_n(\theta) = 0$ gives $\hat{\theta} = n/T(X) = 1/\bar{X}_n$, which is the MLE since $\ell''_n(\theta) < 0$ for all $\theta > 0$.

We can now compute the moments:

$$E_\theta(1/T) = \int_0^\infty \frac{1}{t} \frac{\theta^n}{\Gamma(n)} t^{n-1} e^{-\theta t} dt = \frac{\theta^n}{\Gamma(n)} \frac{\Gamma(n-1)}{\theta^{n-1}} = \frac{\theta}{n-1},$$

$$E_\theta(1/T^2) = \int_0^\infty \frac{1}{t^2} \frac{\theta^n}{\Gamma(n)} t^{n-1} e^{-\theta t} dt = \frac{\theta^n}{\Gamma(n)} \frac{\Gamma(n-2)}{\theta^{n-2}} = \frac{\theta^2}{(n-1)(n-2)}.$$

Therefore, $E_\theta \hat{\theta} = \frac{n}{n-1} \theta \rightarrow \theta$ and hence $\hat{\theta}$ is biased for fixed n , but asymptotically unbiased. For the variance, $\text{Var}_\theta(\hat{\theta}) = \frac{n^2}{(n-1)^2(n-2)} \theta^2 \rightarrow 0$.

- (c) Let $\phi = h(\theta)$ and for notational simplicity let $g = h^{-1}$ so that $\theta = g(\phi)$. With $t = \sum_i x_i$, the ϕ -log-likelihood equals $\tilde{\ell}_n(\phi) = n \log g(\phi) - g(\phi)t$. Differentiating,

$$\tilde{\ell}'_n(\phi) = \frac{ng'(\phi)}{g(\phi)} - g'(\phi)t = 0,$$

so that if $g'(\phi) \neq 0$, we have solution $g(\phi) = n/t$. Since g is injective on its domain, we have $\hat{\phi} = g^{-1}(n/t) = h(n/t) = h(\hat{\theta})$ [this ‘invariance of MLE’ actually holds much more generally, as we will prove later in the course]. We use this expression to look for a suitable reparametrization. Since $T \sim \text{Gamma}(n, \theta)$, we have $E_\theta(T/n) = E_\theta(1/\hat{\theta}) = 1/\theta$, which suggests taking $\phi = h(\theta) = 1/\theta$. We can then verify directly that $\hat{\phi} = T/n$ is the MLE for ϕ and $E_\phi \hat{\phi} = E_\phi T/n = \phi$.

8. (i) Since $p = 1 - \sqrt{\theta}$,

$$f_p(x) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} = p^{\sum_i x_i} (1-p)^{n-\sum_i x_i} = (1-\sqrt{\theta})^{\sum_i x_i} (\sqrt{\theta})^{n-\sum_i x_i} = f_\theta(x).$$

Writing the RHS as $g(\sum x_i, \theta)$, we have that $T(X) = \sum_{i=1}^n X_i$ is sufficient for θ by the factorization criterion. For the MLE, writing $\ell_n = \log f$ and using the chain rule,

$$\begin{aligned} \frac{d\ell_n}{d\theta} &= \frac{d\ell_n}{dp} \frac{dp}{d\theta} = \frac{d}{dp} \{t \log p + (n-t) \log(1-p)\} \Big|_{p=1-\sqrt{\theta}} \frac{dp}{d\theta} \\ &= \left[\frac{t}{1-\sqrt{\theta}} - \frac{n-t}{\sqrt{\theta}} \right] \left(-\frac{1}{2\sqrt{\theta}} \right). \end{aligned}$$

Setting the RHS equal to zero and solving for θ yields $\sqrt{\theta} = (n-t)/n$ and hence $\hat{\theta} = (1-t/n)^2$. We now check this is a maximum over $\theta \in [0, 1]$. By rearranging the inequality, one can check that $\ell'_n(\theta) \geq 0$ if and only if $1-t/n \geq \sqrt{\theta}$, which implies $\hat{\theta}$ is the MLE.

- (ii) $E_p 1\{X_1 + X_2 = 0\} = P_p(X_1 + X_2 = 0) = P_p(X_1 = X_2 = 0) = (1-p)^2 = \theta$.
(iii) We use the estimator from the Rao-Blackwell theorem. Noting that $X_1, \dots, X_n \sim \text{Bernoulli}(1-\sqrt{\theta})$,

$$\begin{aligned} E[1\{X_1 + X_2 = 0\}|T=t] &= P_\theta(X_1 + X_2 = 0|T=t) \\ &= \frac{P_\theta(X_1 + X_2 = 0, \sum_{i=1}^n X_i = t)}{P_\theta(T=t)} \\ &= \frac{P_\theta(X_1 = 0)P_\theta(X_2 = 0)P_\theta(\sum_{i=3}^n X_i = t)}{P_\theta(T=t)} \\ &= \begin{cases} \frac{(\sqrt{\theta})^2 \binom{n-2}{t} (1-\sqrt{\theta})^t (\sqrt{\theta})^{n-2-t}}{\binom{n}{t} (1-\sqrt{\theta})^t (\sqrt{\theta})^{n-t}} & \text{if } t = 0, 1, \dots, n-2, \\ 0 & \text{otherwise.} \end{cases} \\ &= \begin{cases} \frac{(n-t)(n-t-1)}{n(n-1)} & \text{if } t = 0, 1, \dots, n-2, \\ 0 & \text{otherwise.} \end{cases} \\ &= \frac{(n-t)(n-t-1)}{n(n-1)} \quad \text{if } t = 0, 1, 2, \dots, n. \end{aligned}$$

Since $1\{X_1 + X_2 = 0\}$ is an unbiased estimator of θ^2 , then so is the estimator

$$\frac{(n-T)(n-T-1)}{n(n-1)} = \frac{(n - \sum_i X_i)(n - \sum_i X_i - 1)}{n(n-1)} = \left(1 - \frac{1}{n} \sum_i X_i\right) \left(1 - \frac{1}{n-1} \sum_i X_i\right)$$

by the Rao-Blackwell theorem.

- 9.** Unbiasedness follows from $E_\theta X_1 = \frac{3}{2}\theta$. For minimal sufficiency,

$$\frac{f_\theta(x)}{f_\theta(x')} = \frac{\prod_{i=1}^n \frac{1}{\theta} 1\{\theta \leq x_i \leq 2\theta\}}{\prod_{i=1}^n \frac{1}{\theta} 1\{\theta \leq x'_i \leq 2\theta\}} = \frac{1\{\theta \leq x_i \leq 2\theta, \text{ for all } i\}}{1\{\theta \leq x'_i \leq 2\theta, \text{ for all } i\}} = \frac{1\{\theta \leq \min_i x_i\} 1\{\max_i x_i \leq 2\theta\}}{1\{\theta \leq \min_i x'_i\} 1\{\max_i x'_i \leq 2\theta\}}$$

This ratio does not depend on θ if and only if $\min x_i = \min x'_i$ and $\max x_i = \max x'_i$. Thus $T = (\min x_i, \max x_i)$ is minimal sufficient for θ by a theorem in the lecture notes.

We know from the notes that the Rao-Blackwell estimator $E[\tilde{\theta}|T = t]$ is unbiased and has smaller variance than $\tilde{\theta}$ (since $\tilde{\theta}$ is not a function of T), so we will compute this. Set $x_{(1)} = \min x_i$ and $x_{(n)} = \max x_i$. Since X_1, \dots, X_n are i.i.d, they are each equally likely to be the smallest or largest value in the sample, and thus $P_\theta(X_1 = X_{(1)}) = P_\theta(X_1 = X_{(n)}) = 1/n$. Using the law of total expectation, for any $\theta \leq a \leq b \leq 2\theta$,

$$\begin{aligned} E_\theta[X_1|X_{(1)} = a, X_{(n)} = b] &= E_\theta[X_1|X_{(1)} = a, X_{(n)} = b, X_1 = X_{(1)}]P_\theta(X_1 = X_{(1)}) \\ &\quad + E_\theta[X_1|X_{(1)} = a, X_{(n)} = b, X_1 = X_{(n)}]P_\theta(X_1 = X_{(n)}) \\ &\quad + E_\theta[X_1|X_{(1)} = a, X_{(n)} = b, X_1 \neq X_{(1)}, X_{(n)}]P_\theta(X_1 \neq X_{(1)}, X_{(n)}) \end{aligned}$$

(we have split into the 3 cases where X_1 is the smallest, the largest or neither). The first conditional expectation is on the event $\{X_1 = \min X_i = a\}$ and thus equals a . Similarly, the second conditional expectation equals b . The third equals $E_\theta[X_1|a \leq X_1 \leq b]$, which we now compute. For any $a \leq x \leq b$,

$$P_\theta(X_1 \leq x|a \leq X_1 \leq b) = \frac{P_\theta(a \leq X_1 \leq x)}{P_\theta(a \leq X_1 \leq b)} = \frac{(x-a)/\theta}{(b-a)/\theta} = \frac{x-a}{b-a},$$

i.e. $X_1|a \leq X_1 \leq b \sim U[a, b]$. Therefore, $E_\theta[X_1|a \leq X_1 \leq b] = EU[a, b] = \frac{a+b}{2}$. Combining all of these,

$$E_\theta[X_1|X_{(1)} = a, X_{(n)} = b] = a \cdot \frac{1}{n} + b \cdot \frac{1}{n} + \frac{a+b}{2} \left(1 - \frac{2}{n}\right) = \frac{a+b}{2},$$

giving our estimator $\hat{\theta} = \frac{2}{3}E_\theta[X|T] = (\min_i X_i + \max X_i)/3$.

- 10.** We prove the following claim:

Claim: If T is sufficient for θ and there exists an MLE for θ , then there exists an MLE $\hat{\theta}$ for θ that is a function of T .

Proof. By the factorization criterion, we have can write $f_\theta(x) = g(T(x), \theta)h(x)$, so that the log-likelihood equals

$$\log f_\theta(x) = \log g(T(x), \theta) + \log h(x).$$

Maximizing this with respect to θ is equivalent to maximizing $\log g(T(x), \theta)$. Maximizing $\log g(T(x), \theta)$ thus means there exists an MLE $\hat{\theta}$ that is a function of $T(x)$. \square

If S is another sufficient statistic for θ , then by the claim there exists an MLE estimator $\hat{\theta}_S$ that is a function of S . But by uniqueness of the original MLE $\hat{\theta}_{ML}$, we must have $\hat{\theta}_{ML} = \hat{\theta}_S$ and thus $\hat{\theta}_{ML}$ is a function of S . Thus $\hat{\theta}_{ML}$ is both sufficient and a function of any sufficient statistic, and so is minimal sufficient.

- 11.** Note that X_i depends on u_1, \dots, u_{i-1} only through X_{i-1} , and thus the conditional distribution of X_i given X_1, \dots, X_{i-1} is the same as that given X_{i-1} . The likelihood therefore equals

$$f_\theta(x_1, \dots, x_n) = f_\theta(x_1) \prod_{i=2}^n f_\theta(x_i | x_1, \dots, x_{i-1}) = f_\theta(x_1) \prod_{i=2}^n f_\theta(x_i | x_{i-1}).$$

Since we have conditional distributions $X_i | X_{i-1} \sim N(\theta X_{i-1}, 1 - \theta^2)$ and $X_1 \sim N(0, 1)$, this equals

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \prod_{i=2}^n \frac{1}{\sqrt{2\pi(1-\theta^2)}} \exp\left(-\frac{(x_i - \theta x_{i-1})^2}{2(1-\theta^2)}\right) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} (1-\theta^2)^{\frac{n-1}{2}}} \exp\left\{-\frac{1}{2(1-\theta^2)} \sum_{i=2}^n (x_i - \theta x_{i-1})^2\right\} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} (1-\theta^2)^{\frac{n-1}{2}}} \exp\left\{-\frac{1}{2(1-\theta^2)} \sum_{i=2}^n (x_i^2 - 2\theta x_i x_{i-1} + \theta^2 x_{i-1}^2)\right\}. \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} (1-\theta^2)^{\frac{n-1}{2}}} \exp\left\{-\frac{1}{2(1-\theta^2)} \sum_{i=2}^n x_i^2 + \frac{\theta}{(1-\theta^2)} \sum_{i=2}^n x_i x_{i-1} - \frac{\theta^2}{2(1-\theta^2)} \sum_{i=2}^n x_{i-1}^2\right\}. \end{aligned}$$

We can write this as a function of θ and $T(x) = (\sum_{i=2}^n x_i^2, \sum_{i=2}^n x_i x_{i-1}, \sum_{i=1}^{n-1} x_i^2)$, and so by the factorization criterion, T is sufficient for θ .