Matthew J.

Use appropriate statistical methods to answer these questions. **Make sure to justify your answers and steps**! An inadequately justified answer will receive no points or have points deducted.

Suppose that $X_1, \dots, X_n \overset{i.i.d.}{\sim} f(x \mid b)$ with

$$f(x \mid b) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right)$$

with $b > 0$.

1. Determine a complete sufficient statistic for b.

$$f(x|b) = \frac{1}{2b} \exp\left(\frac{-|x|}{b}\right)$$

3.4 Exponential Families

A family of pdfs or pmfs is called an *exponential family* if it can be expressed as

$$(3.4.1) \qquad f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta}) \exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta}) t_i(x)\right).$$

$\therefore$ we can rewrite as:

$$f(x|b) = \underbrace{1}_{h(x)} \cdot \underbrace{\frac{1}{2b}}_{c(\theta)} \cdot \underbrace{\exp\left(\frac{-|x|}{b}\right)}_{\substack{w_i(\theta) = -\frac{1}{b} \\ t_i(x) = |x|}}$$

where $b = \theta$

$\therefore f(x|b)$ is in the exponential family

Theorem 6.2.25 (Complete statistics in the exponential family) *Let $X_1, \ldots, X_n$ be iid observations from an exponential family with pdf or pmf of the form*

$$f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta}) \exp\left(\sum_{j=1}^{k} w(\theta_j)t_j(x)\right), \tag{6.2.7}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$. Then the statistic

$$T(\mathbf{X}) = \left(\sum_{i=1}^{n} t_1(X_i), \sum_{i=1}^{n} t_2(X_i), \ldots, \sum_{i=1}^{n} t_k(X_i)\right)$$

is complete as long as the parameter space Θ contains an open set in $\Re^k$.

using this, we can see that $T(x)$ is complete if Θ contains an open set.

we are told $b > 0$ (open set in $\mathbb{R}^k$).

$$\therefore \quad T(x) = \left(\sum_{i=1}^{n} t_1(X_i), \ldots, \sum_{i=1}^{n} t_k(X_i)\right)$$

$$= \left(\sum_{i=1}^{n} |X_i|, \ldots, \sum_{i=1}^{n} |X_i|\right)$$

$$= \sum_{i=1}^{n} |X_i|$$

$\therefore T(x) = \sum_{i=1}^{n} |X_i|$ is a complete stat for b.

Now we will show $T(x)$ is sufficient with:

Theorem 6.2.6 (Factorization Theorem) *Let $f(\mathbf{x}|\theta)$ denote the joint pdf or pmf of a sample $\mathbf{X}$. A statistic $T(\mathbf{X})$ is a sufficient statistic for θ if and only if there exist functions $g(t|\theta)$ and $h(\mathbf{x})$ such that, for all sample points $\mathbf{x}$ and all parameter points θ,*

$$f(\mathbf{x}|\theta) = g(T(\mathbf{x})|\theta)h(\mathbf{x}). \tag{6.2.3}$$

$$\therefore f(X_1, \ldots, X_n|\theta) = \prod_{i=1}^{n} 1 \cdot \frac{1}{2b} \cdot \exp\left(\frac{-|X_i|}{b}\right)$$

$$= \underbrace{\left(\frac{1}{2b}\right)^n \exp\left(-\frac{1}{b}\underbrace{\sum_{i=1}^{n}|X_i|}_{T(x)}\right)}_{g(T(x)|\theta)} \cdot \underbrace{1}_{h(x)}$$

$\therefore$ We have shown that $T(x) = \sum_{i=1}^{n}|X_i|$ is complete and sufficient for b.

2. Determine the Method of Moments estimator for b^2. Determine the mean and variance of the estimator.

from our table we can see that:

Double exponential(μ, σ)	
pdf	$f(x\|\mu,\sigma) = \frac{1}{2\sigma}e^{-\|x-\mu\|/\sigma}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$
mean and variance	$\mathrm{E}X = \mu, \quad \mathrm{Var}\,X = 2\sigma^2$
mgf	$M_X(t) = \frac{e^{\mu t}}{1-(\sigma t)^2}, \quad \|t\| < \frac{1}{\sigma}$
notes	Also known as the *Laplace* distribution.

$$\therefore \quad f(x|\mu,\sigma) = \frac{1}{2\sigma}\exp\left(-\frac{|x-\mu|}{\sigma}\right) = f(x|b)$$

where $\mu = 0$ & $\sigma = b$

$$\therefore \quad E[X] = 0 \quad \& \quad E[X^2] = Var[X] = 2b^2$$

$$E[X^2] \approx \frac{\sum_{i=1}^{n} X_i^2}{n}$$

$$\therefore \quad 2\tilde{b}^2 = \frac{\sum_{i=1}^{n} X_i^2}{n}$$

$$\tilde{b}^2 = \frac{\sum_{i=1}^{n} X_i^2}{2n} = b^2_{MoM}$$

to find the mean:

$$\frac{\sum_{i=1} \ldots}{2n} = \ldots_{MoM}$$

to find the mean:

$$E[\hat{b}^2] = E\left[\frac{\sum_{i=1}^{n} X_i^2}{2n}\right]$$

$$= \frac{1}{2n} \sum_{i=1}^{n} E[X_i^2]$$

$$= \frac{1}{\cancel{2n}} \cdot \cancel{n} \cdot \cancel{2}b^2$$

$$E[\hat{b}^2] = b^2 \qquad \text{(unbiased)}$$

finding variance:

$$Var[\hat{b}^2] = Var\left[\frac{\sum_{i=1}^{n} X_i^2}{2n}\right]$$

$$= \frac{1}{4n^2} Var\left[\sum_{i=1}^{n} X_i^2\right]$$

$$= \frac{1}{4n^2} \cdot n \cdot Var[X_i^2]$$

$$= \frac{1}{4n} \cdot E[X_i^4] - E[X_i^2]^2$$

$$\therefore Var[\hat{b}^2] = \frac{5b^4}{n}$$

3. Determine the Maximum Likelihood estimator (MLE) of b^2. Determine the mean of the estimator. Recall that $\text{var}(Y_i) = E(Y_i^2) - [E(Y_i)]^2$.

Definition 7.2.4 For each sample point $\mathbf{x}$, let $\hat{\theta}(\mathbf{x})$ be a parameter value at which $L(\theta|\mathbf{x})$ attains its maximum as a function of θ, with $\mathbf{x}$ held fixed. A *maximum likelihood estimator* (MLE) of the parameter θ based on a sample $\mathbf{X}$ is $\hat{\theta}(\mathbf{X})$.

$$\therefore L(\theta|x) = L(f(x_i|b))$$

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

$$= \prod_{i=1} \frac{1}{2b} \exp\left(\frac{\quad}{b} \right)$$

$$= \left(\frac{1}{2b}\right)^n \prod_{i=1}^{n} \exp\left(\frac{-|X_i|}{b}\right)$$

$$= \left(\frac{1}{2b}\right)^n \exp\left(-\frac{1}{b}\sum_{i=1}^{n} |X_i|\right)$$

take the log:

$$\log(L(f(X_i|b))) = \log\left(\left(\frac{1}{2b}\right)^n e^\wedge\left(-\frac{1}{b}\sum_{i=1}^{n}|X_i|\right)\right)$$

$$= n\log\left(\frac{1}{2b}\right) + \left(-\frac{1}{b}\sum_{i=1}^{n}|X_i|\right)$$

take derivative:

$$d/db\left(\log(L(f(X_i|b)))\right) = d/db\left(n\log\left(\frac{1}{2b}\right) - \frac{1}{b}\sum_{i=1}^{n}|X_i|\right)$$

$$= -\frac{n}{b} + \frac{1}{b^2}\sum_{i=1}^{n}|X_i|$$

set equal to 0:

$$\frac{n}{b} = \frac{1}{b^2}\sum_{i=1}^{n}|X_i|$$

solve for b:

$$\frac{nb^2}{b} = \sum_{i=1}^{n}|X_i|$$

$$nb = \sum_{i=1}^{n}|X_i|$$

$$\therefore \hat{b} = b_{MLE} = \frac{\sum_{i=1}^{n}|X_i|}{n}$$

$$\therefore \hat{b}^2 = b^2_{MLE} = \left(\frac{\sum_{i=1}^{n}|X_i|}{n}\right)^2$$

finding mean:

$$Var(\hat{b}) = \underset{?}{E[\hat{b}^2]} - [E[\hat{b}]]^2$$

$$Var(\hat{b}) + [E[\hat{b}]]^2 = E[\hat{b}^2]$$

$$Var\left[\frac{\sum_{i=1}^{n}|X_i|}{n}\right] + \left[E\left[\frac{\sum_{i=1}^{n}|X_i|}{n}\right]\right]^2 = E[\hat{b}^2]$$

check MLE with second derivative:

$$\frac{d}{db}\left(\frac{-n}{b} + \frac{1}{b^2}\sum_{i=1}^{n}|X_i|\right) = \frac{n}{b^2} - \frac{2}{b^3}\sum_{i=1}^{n}|X_i|$$

plug in critical point $\hat{b}$:

$$\frac{n}{\hat{b}^2} - \frac{2}{\hat{b}^3}\sum_{i=1}^{n}|X_i| = \frac{n}{\hat{b}^2} - \frac{2n}{\hat{b}^2}$$

$$= -\frac{n}{\hat{b}^2}$$

we are told $b > 0$, and thus $\hat{b}$ & $\hat{b}^2$ must also be > 0. also n must be > 0.

$\therefore -\frac{n}{\hat{b}^2} \Rightarrow -\left(\frac{+}{+}\right) = -$. $\frac{-n}{\hat{b}^2}$ is always negative, indicating $\hat{b}^2 = \left(\frac{\sum_{i=1}^{n}|X_i|}{n}\right)^2$ is indeed a max.

4. Determine the CRLB for any unbiased estimator of b^2.

we can use:

Theorem 7.3.9 (Cramér–Rao Inequality) *Let $X_1, \ldots, X_n$ be a sample with pdf $f(\mathbf{x}|\theta)$, and let $W(\mathbf{X}) = W(X_1, \ldots, X_n)$ be any estimator satisfying*

$$\frac{d}{d\theta} \mathrm{E}_\theta W(\mathbf{X}) = \int_{\mathcal{X}} \frac{\partial}{\partial\theta} [W(\mathbf{x}) f(\mathbf{x}|\theta)] \; d\mathbf{x}$$

(7.3.4) *and*

$$\mathrm{Var}_\theta W(\mathbf{X}) < \infty.$$

Then

$$\mathrm{Var}_\theta \, (W(\mathbf{X})) \geq \frac{\left(\frac{d}{d\theta} \mathrm{E}_\theta W(\mathbf{X})\right)^2}{\mathrm{E}_\theta \left(\left(\frac{\partial}{\partial\theta} \log f(\mathbf{X}|\theta)\right)^2\right)}. \tag{7.3.5}$$

However in problem (1), we showed $f(x|b)$ is part of the exponential family.

$\therefore$ We can use:

Lemma 7.3.11 *If $f(x|\theta)$ satisfies*

$$\frac{d}{d\theta}\mathrm{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right) = \int \frac{\partial}{\partial\theta}\left[\left(\frac{\partial}{\partial\theta}\log f(x|\theta)\right) f(x|\theta)\right] dx$$

(true for an exponential family), then

$$\mathrm{E}_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right)^2\right) = -\mathrm{E}_\theta\left(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right).$$

$$\log(f(x|b)) = \log\left(\frac{1}{2b}\exp\left(\frac{-|x|}{b}\right)\right)$$

$$= \log\left(\frac{1}{2b}\right) - \frac{-|X|}{b}$$

$$\partial/\partial b(\log(f(x|b))) = 2b\cdot\frac{1}{2}\cdot(-1)\cdot\frac{1}{b^2} + \frac{|X|}{b}$$

$$= -\frac{b}{b^2} + \frac{|X|}{b^2}$$

$$= -\frac{b}{b^2} + \frac{|x|}{b^2}$$

$$= \frac{-1}{b} + \frac{|x|}{b^2}$$

$$\frac{\partial^2}{\partial b^2}(\log(f(x|b))) = \frac{\partial}{\partial b}\left(\frac{-1}{b} + \frac{|x|}{b^2}\right)$$

$$= \frac{1}{b^2} - 2\frac{|x|}{b^3}$$

$$-E_b\left[\frac{\partial^2}{\partial b^2}\log(f(x|b))\right] = -E_b\left[\frac{1}{b^2} - \frac{2|x|}{b^3}\right]$$

$$= \frac{1}{b^2} - \frac{2E_b[|x|]}{b^3}$$

$$= \frac{1}{b^2} - \frac{2(0)}{b^3}$$

$$= \frac{1}{b^2}$$

$$\therefore E_b = \frac{1}{b^2} \leftarrow \star$$

$$E_\theta[W(x)] = E_b[b^2]$$

$$= b^2 \quad (\text{since unbiased})$$

$$\frac{d}{d\theta}(E_\theta[W(x)]) = \frac{d}{db}(E_b[W(x)])$$

$$= \frac{d}{db}(E_b[b^2])$$

$$= 2b$$

$$\left(\frac{d}{d\theta}(E_\theta[W(x)])\right)^2 = \left(\frac{d}{db}(E_b[W(x)])\right)^2$$

$$= (2b)^2$$

$$= 4b^2 \leftarrow \Delta$$

$$= 4b^2 \leftarrow \Delta$$

$$\text{Var}_\theta(W(X)) \geq \frac{\left(\frac{d}{d\theta} E_\theta[W(X)]\right)^2 \leftarrow \Delta}{n E_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(x|\theta)\right)^2\right) \leftarrow *}$$

(iid case)

$$\therefore \text{Var}_{b^2}(W(X)) \geq \frac{4b^2}{n \frac{1}{b^2}}$$

$$\geq \frac{4b^2}{1} \cdot \frac{b^2}{1} \cdot \frac{1}{n}$$

$$\geq \frac{4b^4}{n}$$

$\therefore$ CRLB for unbiased estimator of b^2

$$= \frac{4b^4}{n}$$

5. Determine the UMVUE for b^2.

We will use:

Theorem 7.5.1 (Lehmann–Scheffé) *Unbiased estimators based on complete sufficient statistics are unique.*

from part (1), we found that a complete sufficient for b is:

$$T(X) = \sum_{i=1}^{n} |X_i|$$

given $X_i \sim f(x|b) = \frac{1}{2b}\exp\left(\frac{-|x|}{b}\right)$,

then $|X_i| \sim \text{exponential}(\lambda = 1/b)$

we also know that the sum of independent exponential R.V.'s $\sim$ gamma

$$\therefore \quad T(x) = \sum_{i=1}^{n} |X_i| \sim \text{gamma}(n, b)$$

Gamma(α, β)	
pdf	$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, \quad 0 \le x < \infty, \quad \alpha, \beta > 0$
mean and variance	$\mathrm{E}X = \alpha\beta, \quad \mathrm{Var}\,X = \alpha\beta^2$

where $\alpha = n$ & $\beta = b$

$$\therefore \quad E[T(x)] = nb$$

$$\therefore \quad E\left[\frac{T(x)}{n}\right] = b$$

$$E[X^2] = 2b^2$$

$$E\left[\frac{X^2}{2}\right] = b^2$$

$$\therefore \quad E\left[\frac{(T(x))^2}{2n}\right] = b^2$$

$$\therefore \quad \frac{\sum_{i=1}^{n} |X_i|^2}{2n} = \frac{T^2}{2n} \quad \text{is UMVUE for } b^2$$

6. Does the variance of the UMVUE attain the CRLB? Why?

$$Var\left[\frac{T^2}{2n}\right] = Var\left[\frac{\sum_{i=1}^{n}|X_i|^2}{2n}\right]$$

$$= \frac{1}{4n^2} Var\left[\sum_{i=1}^{n}|X_i|^2\right]$$

$$Var(g(T)) \approx (g'(E[T]))^2 \, Var(T)$$

$$g'(T) = \frac{dT^2}{dT} \cdot \frac{1}{2n}$$

$$= \frac{2T}{2n}$$

$$= \frac{T}{n}$$

$$g'(E[T]) = \frac{nb}{n}$$

$$= b$$

$$\therefore \; Var\left(\frac{T^2}{2n}\right) \approx b^2 \cdot nb^2 = nb^4$$

$$Var\left(\frac{T^2}{2n}\right) \geq CRLB$$

$$nb^4 > \underline{4b^4}$$

$$nb^4 > \frac{4b^4}{n}$$

No, the variance of the UMVUE does not attain CRLB

7. Construct a likelihood ratio test for testing $H_0: b = 1$ versus $H_1: b \neq 1$. Simplify as much as possible.

from part 3 we found that

$$b_{MLE} = \hat{b} = \frac{\sum_{i=1}^{n} |X_i|}{n}$$

we also found:

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

$$= \left(\frac{1}{2b}\right)^n \exp\left(\frac{-1}{b} \sum_{i=1}^{n} |X_i|\right)$$

Definition 8.2.1 The *likelihood ratio test statistic* for testing $H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_0^c$ is

$$\lambda(\mathbf{x}) = \frac{\sup_{\Theta_0} L(\theta|\mathbf{x})}{\sup_{\Theta} L(\theta|\mathbf{x})}.$$

A *likelihood ratio test* (LRT) is any test that has a rejection region of the form $\{\mathbf{x}: \lambda(\mathbf{x}) \leq c\}$, where c is any number satisfying $0 \leq c \leq 1$.

$$\therefore \lambda = \frac{L(b=1)}{L(b=\hat{b})}$$

$$L(b = b$$

$$= \frac{\left(\frac{1}{2\cdot 1}\right)^n \exp\left(\frac{-1}{1}\sum_{i=1}^{n}|X_i|\right)}{\left(\frac{1}{2\hat{b}}\right)^n \exp\left(\frac{-1}{\hat{b}}\sum_{i=1}^{n}|X_i|\right)}$$

$$= \frac{\left(\frac{1}{2}\right)^n \exp\left(-\sum_{i=1}^{n}|X_i|\right)}{\left(\frac{1}{2\hat{b}}\right)^n \exp\left(\frac{-1}{\hat{b}}\sum_{i=1}^{n}|X_i|\right)}$$

$$\left(\frac{1}{2}\right)^n \Big/ \left(\frac{1}{2\hat{b}}\right)^n = (2\hat{b})^n/2^n$$

$$= \hat{b}^n$$

$$\exp\left(-\sum_{i=1}^{n}|X_i|\right)\Big/\exp\left(\frac{-1}{\hat{b}}\sum_{i=1}^{n}|X_i|\right) = \exp\left(\left(1-\frac{1}{\hat{b}}\right)\sum_{i=1}^{n}|X_i|\right)$$

$$\therefore \lambda = \hat{b}^n \exp\left(\left(1-\frac{1}{\hat{b}}\right)\sum_{i=1}^{n}|X_i|\right)$$