STAT 30100 : Mathematical Statistics-1

Winter 2020

•

Homework 6

•

Solutions by

Jinhong Du

12243476

STAT 30100, Homework 6

1. Suppose $X_1, \dots, X_n$ are i.i.d. Unif$(\theta, \theta+1)$. Then $(X_{(1)}, X_{(n)})$ is minimal sufficient. Prove this carefully using the version of the Lehmann-Scheffé Theorem given in class, which allows the support of the distribution to depend on the parameter. (Do not use Theorem 6.2.13 in Casella and Berger, which is stated in such a way that it does not allow for the support to depend on the parameter.)

Proof. Since $X_1, \dots, X_n \overset{iid}{\sim} \text{Unif}(\theta, \theta+1)$, the joint density of $\boldsymbol{X} = (X_1, \dots, X_n)^\top$ is $f_\theta(\boldsymbol{x}) = \mathbb{1}_{\{\theta < x_1, \dots, x_n < \theta+1\}} = \mathbb{1}_{\{\theta < x_{(1)} < x_{(n)} < \theta+1\}}$. The experiment is given by $(\mathcal{X}, \mathcal{A}, \{f_\theta : \theta \in \Theta\})$ where $\mathcal{X} = \{\boldsymbol{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n : x_{(n)} - x_{(1)} < 1\}$, $\mathcal{A}$ is the set of all subsets of $\mathcal{X}$, and $\Theta = \mathbb{R}$. For $\boldsymbol{x} \in \mathcal{X}$, define $\Theta_{\boldsymbol{x}} = \{\theta : f_\theta(\boldsymbol{x}) > 0\}$, then $\Theta_{\boldsymbol{x}} = \{\theta : x_{(n)} - 1 < \theta < x_{(1)}\} \neq \varnothing$. So for any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, $\Theta_{\boldsymbol{x}} = \Theta_{\boldsymbol{y}}$ if and only if $(x_{(1)}, x_{(n)}) = (y_{(1)}, y_{(n)})$, and when this holds, $\frac{f_\theta(\boldsymbol{x})}{f_\theta(\boldsymbol{y})} = 1$ for all $\theta \in \Theta$. Then by the Lehmann-Scheffé Theorem, we have $(X_{(1)}, X_{(n)})$ is the minimal sufficient statistic for θ. $\square$

2. (Casella and Berger Problem 6.9) For each of the following distributions let $X_1, \dots, X_n$ be a random sample. Find a minimal sufficient statistic for θ.

   (a) $f(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty. \quad$ (normal)

   Since $fx|\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot e^{-\frac{\theta^2}{2}} \cdot e^{\theta x}$, the densities are in the 1-parameter exponential family where $t_1(x) = x$. Then by Theorem 6.2.10, $T(\boldsymbol{X}) = \sum_{i=1}^n X_i$ is a sufficient statistic for θ. Since this is a full exponential family and the parameter space Θ contains a open set in $\mathbb{R}$, by Theorem 6.2.25 we have $T(\boldsymbol{X}) = \sum_{i=1}^n X_i$ is a complete statistic for θ. Therefore, $T(\boldsymbol{X}) = \sum_{i=1}^n X_i$ is a minimal sufficient statistic for θ.

   (b) $f(x|\theta) = e^{-(x-\theta)}, \quad \theta < x < \infty, \quad -\infty < \theta < \infty. \quad$ (location exponential)

   The joint density of $\boldsymbol{X}$ is given by $f_\theta(\boldsymbol{x}) = e^{-\sum_{i=1}^n (x_i - \theta)} \mathbb{1}_{\{x_{(1)} > \theta\}}$. Define $\Theta_{\boldsymbol{x}} = \{\theta : f_\theta(x) > 0\} = \{\theta : \theta < x_{(1)}\}$.
   Since $\forall\ \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, $\Theta_{\boldsymbol{x}} = \Theta_{\boldsymbol{y}}$ and $\frac{f_\theta(\boldsymbol{x})}{f_\theta(\boldsymbol{y})} = e^{-\sum_{i=1}^n (x_{(i)} - y_{(i)})} \frac{\mathbb{1}_{\{x_{(1)} > \theta\}}}{\mathbb{1}_{\{y_{(1)} > \theta\}}}$ will be constant as a function of θ if and only if $x_{(1)} = y_{(1)}$, by Lehmann-Scheffé Theorem, we have $X_{(1)}$ is a minimal sufficient statistic for θ.

   (c) $f(x|\theta) = \frac{e^{-(x-\theta)}}{\left(1+e^{-(x-\theta)}\right)^2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty. \quad$ (logistic)

   The joint density of $\boldsymbol{X}$ is given by $f_\theta(\boldsymbol{x}) = \prod_{i=1}^n \frac{e^{-(x_i - \theta)}}{(1+e^{-(x_i-\theta)})^2}$. Define $\Theta_{\boldsymbol{x}} = \{\theta : f_\theta(\boldsymbol{x}) > 0\} = \mathbb{R}$.
   Since $\forall\ \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, $\Theta_{\boldsymbol{x}} = \Theta_{\boldsymbol{y}}$ and $\frac{f_\theta(\boldsymbol{x})}{f_\theta(\boldsymbol{y})} = e^{-\sum_{i=1}^n (x_i - y_i)} \prod_{i=1}^n \frac{(1+e^{-(y_i-\theta)})^2}{(1+e^{-(x_i-\theta)})^2}$ will be constant as a function of θ if and only if $(x_{(1)}, \dots, x_{(n)}) = (y_{(1)}, \dots, y_{(n)})$, by Lehmann-Scheffé Theorem, we have $(X_{(1)}, \dots, X_{(n)})$ is a minimal sufficient statistic for θ.

(d) $f(x|\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty. \quad$ (Cauchy)

The joint density of $\boldsymbol{X}$ is given by $f_\theta(\boldsymbol{x}) = \prod_{i=1}^n \frac{1}{\pi[1+(x_i-\theta)^2]}$. Define $\Theta_{\boldsymbol{x}} = \{\theta : f_\theta(x) > 0\} = \{\theta : \theta \neq x_i,\ i = 1, \ldots, n\}$.
Since $\forall\ \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, $\Theta_{\boldsymbol{x}} = \Theta_{\boldsymbol{y}}$ and $\frac{f_\theta(\boldsymbol{x})}{f_\theta(\boldsymbol{y})} = \prod_{i=1}^n \frac{1+(y_i-\theta)^2}{1+(x_i-\theta)^2}$ will be constant as a function of θ if and only if $(x_{(1)}, \ldots, x_{(n)}) = (y_{(1)}, \ldots, y_{(n)})$, by Lehmann-Scheffé Theorem, we have $(X_{(1)}, \ldots, X_{(n)})$ is a minimal sufficient statistic for θ.

(e) $f(x|\theta) = \frac{1}{2}e^{-|x-\theta|}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty. \quad$ (double exponential)

The joint density of $\boldsymbol{X}$ is given by $f_\theta(\boldsymbol{x}) = \frac{1}{2^n}e^{-\sum_{i=1}^n |x_i-\theta|}$. Define $\Theta_{\boldsymbol{x}} = \{\theta : f_\theta(x) > 0\} = \mathbb{R}$.
Since $\forall\ \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, $\Theta_{\boldsymbol{x}} = \Theta_{\boldsymbol{y}}$ and $\frac{f_\theta(\boldsymbol{x})}{f_\theta(\boldsymbol{y})} = e^{\sum_{i=1}^n (|y_{(i)}-\theta| - |x_{(i)}-\theta|)}$ will be constant as a function of θ if and only if $(x_{(1)}, \ldots, x_{(n)}) = (y_{(1)}, \ldots, y_{(n)})$, by Lehmann-Scheffé Theorem, we have $(X_{(1)}, \ldots, X_{(n)})$ is a minimal sufficient statistic for θ.

3. Let $X_1, \ldots, X_n$ be indepdendent but not identically distributed random variables with X_i having pdf

$$f_i(x_i) = \frac{1}{2i\theta} \text{ for } -i(\theta-1) < x_i < i(\theta+1),$$

where $\theta > 0$. Find a minimal sufficient statistic for θ.

Proof. Since the joint density for $\boldsymbol{X}$ is given by

$$\begin{aligned} f_\theta(\boldsymbol{x}) &= \prod_{i=1}^n \frac{1}{2i\theta} \mathbb{1}_{\{-i(\theta-1)<x_i<i(\theta+1)\}} \\ &= \prod_{i=1}^n \frac{1}{2i\theta} \mathbb{1}_{\{\left|\frac{x_i}{i}-1\right|<\theta\}} \\ &= \left(\prod_{i=1}^n \frac{1}{2i\theta}\right) \cdot \mathbb{1}_{\left\{\max_i \left|\frac{x_i}{i}-1\right|<\theta\right\}} \end{aligned}$$

by Fisher-Neymann Factorization Theorem, we have $\max_i \left|\frac{X_i}{i} - 1\right|$ is a sufficient statistic for θ.
Define $\Theta_{\boldsymbol{x}} = \{\theta : f_\theta(\boldsymbol{x}) > 0\} = \left\{\theta : \theta > \max_i \left|\frac{X_i}{i} - 1\right|\right\}$. Since $\forall\ \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, $\Theta_{\boldsymbol{x}} = \Theta_{\boldsymbol{y}}$ and $\frac{f_\theta(\boldsymbol{x})}{f_\theta(\boldsymbol{y})} = \frac{\mathbb{1}_{\left\{\max_i \left|\frac{x_i}{i}-1\right|<\theta\right\}}}{\mathbb{1}_{\left\{\max_i \left|\frac{y_i}{i}-1\right|<\theta\right\}}}$ will be a constant as a function of θ if and only if $\max_i \left|\frac{x_i}{i} - 1\right| = \max_i \left|\frac{y_i}{i} - 1\right|$, by Lehmann-Scheffé Theorem, we have $\max_i \left|\frac{X_i}{i} - 1\right|$ is a minimal sufficient statistic for θ. □

4. (Casella and Berger Problem 6.25) We have seen a number of theorems concerning sufficiency and related concepts for exponential families. Theorem 5.2.11 gave the distribution of a statistic whose sufficiency is characterized in Theorem 6.2.10 and completeness in Theorem 6.2.25. But if the family is curved, the open set condition of Theorem 6.2.25 is not satisfied. In such cases, is the sufficient statistic of Theorem 6.2.10 also minimal? By applying Theorem 6.2.13 to $T(x)$ of Theorem 6.2.10, establish the following:

(a) The statistic $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ is sufficient, but not minimal sufficient, in the $\mathcal{N}(\mu, \mu)$ family.

Proof. The joint density of $\boldsymbol{X}$ is $f_\mu(\boldsymbol{x}) = \frac{1}{(2\pi\mu)^{\frac{n}{2}}} e^{-\frac{1}{2\mu}\sum_{i=1}^n (x_i-\mu)^2} = e^{-\sum_{i=1}^n x_i} \cdot \frac{1}{(2\pi\mu)^{\frac{n}{2}}} e^{-\frac{\mu}{2}} \cdot e^{-\frac{1}{2\mu}\sum_{i=1}^n x_i^2}$, which is in the 1-parameter exponential family with parameter space $\mathbb{R}^+$. By Theorem 6.2.10 and Theorem 6.2.25, $\sum_{i=1}^n X_i^2$ is the complete sufficient statistic for μ. Thus by Theorem 6.2.28, we have $\sum_{i=1}^n X_i^2$ is the minimal sufficient statistic for μ.
By ization Theorem, since $f_\mu(\boldsymbol{x}) = e^{-\sum_{i=1}^n x_i} \frac{1}{(2\pi\mu)^{\frac{n}{2}}} e^{-\frac{\mu}{2}} e^{-\frac{1}{2\mu}\sum_{i=1}^n x_i^2} = g((\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2)|\mu)$, $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ is a sufficient statistic for μ.
It is obvious that $\sum_{i=1}^n X_i^2$ is not a one-to-one function of $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$. Therefore, $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ is not minimal sufficient. $\square$

(b) The statistic $\sum_{i=1}^n X_i^2$ is minimal sufficient in the $\mathcal{N}(\mu, \mu)$ family.

Proof. The joint density of $\boldsymbol{X}$ is $f_\mu(\boldsymbol{x}) = \frac{1}{(2\pi\mu)^{\frac{n}{2}}} e^{-\frac{1}{2\mu}\sum_{i=1}^n (x_i-\mu)^2} = e^{-\sum_{i=1}^n x_i} \cdot \frac{1}{(2\pi\mu)^{\frac{n}{2}}} e^{-\frac{\mu}{2}} \cdot e^{-\frac{1}{2\mu}\sum_{i=1}^n x_i^2}$, which is in the 1-parameter exponential family. By Theorem 6.2.10 and Theorem 6.2.25, $\sum_{i=1}^n X_i^2$ is the complete sufficient statistic for μ. Thus by Theorem 6.2.28, we have $\sum_{i=1}^n X_i^2$ is the minimal sufficient statistic for μ. $\square$

(c) The statistic $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ is minimal sufficient in the $\mathcal{N}(\mu, \mu^2)$ family.

Proof. The joint density of $\boldsymbol{X}$ is $f_\mu(\boldsymbol{x}) = \frac{1}{(2\pi\mu^2)^{\frac{n}{2}}} e^{-\frac{1}{2\mu^2}\sum_{i=1}^n (x_i-\mu)^2} = \frac{1}{(2\pi\mu^2)^{\frac{n}{2}}} e^{-\frac{1}{2}} \cdot e^{-\frac{1}{\mu}\sum_{i=1}^n x_i - \frac{1}{2\mu^2}\sum_{i=1}^n x_i^2}$, which is in the 2-parameter exponential family. By Theorem 6.2.10 and Theorem 6.2.25, $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ is the complete sufficient statistic for μ. Thus by Theorem 6.2.28, we have $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ is the minimal sufficient statistic for μ. $\square$

5. (Casella and Berger Problem 6.11) Refer to the pdfs given in Exercise 6.9. For each, let $X_{(1)} < \ldots < X_{(n)}$ be the ordered sample, and define $Y_i = X_{(n)} - X_{(i)}$, $i = 1, \ldots, n-1$.

(a) For each of the pdfs in Exercise 6.9, verify that the set $(Y_1, \ldots, Y_{n-1})$ is ancillary for θ. Try to prove a general theorem, like Example 6.2.18, that handles all these families at once.

Proof. Let $X_1, \ldots, X_n$ be iid observations from a location parameter family with cdf $F(x-\theta)$, $-\infty < \theta < \infty$. We work with $Z_1, \ldots, Z_n$ iid observations from $F(x)$ (corresponding to $\theta = 0$) with $X_1 = Z_1 + \theta, \ldots, X_n = Z_n + \theta$. Thus the cdf of the range statistic, Y_i, is

$$\begin{aligned} F_{Y_i}(r;\theta) &= \mathbb{P}_\theta(Y_i \le r) \\ &= \mathbb{P}_\theta(X_{(n)} - X_{(i)} \le r) \\ &= \mathbb{P}_\theta(Z_{(n)} - Z_{(i)} \le r) \\ &= \mathbb{P}(Z_{(n)} - Z_{(i)} \le r), \end{aligned}$$

which does not depend on θ. Thus $(Y_1, \ldots, Y_{n-1})$ is an ancillary statistic for all these location families. $\square$

(b) In each case determine whether the set $(Y_1, \ldots, Y_{n-1})$ is independent of the minimal sufficient statistic.

(a) Since $\sum_{i=1}^n X_i$ is a complete and minimal sufficient statistic for θ, by Basu Theorem, it is independent of the ancillary statistic $(Y_1, \ldots, Y_{n-1})$.
(b) The density for $X_{(1)}$ is given by

$$F_{X_{(1)}}(t) = 1 - [1 - F(x)]^n = [1 - e^{n\theta}e^{-nx}]$$
$$f_{X_{(1)}}(t) = ne^{n\theta}e^{-nx}\mathbb{1}_{\{x>\theta\}}$$

Suppose that g is a function such that $\mathbb{E}_\theta[g(X_{(1)})] = 0 \ \forall\ \theta$, then

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}\theta}\mathbb{E}_\theta[g(X_{(1)})] &= \frac{\mathrm{d}}{\mathrm{d}\theta}\int_\theta^\infty g(t)ne^{n(\theta-t)}\mathrm{d}t \\ &= e^{n\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\int_\theta^\infty g(t)ne^{-nt}\mathrm{d}t + \frac{\mathrm{d}e^{n\theta}}{\mathrm{d}\theta}\cdot\int_\theta^\infty g(t)ne^{n(\theta-t)}\mathrm{d}t \\ &= -g(\theta), \end{aligned}$$

which implies that $\mathbb{P}(g(X_{(1)}) = 0) = 1 \ \forall\ \theta \in \mathbb{R}$. So $X_{(1)}$ is complete. Then by Basu's Theorem, $X_{(1)}$ and $(Y_1, \ldots, Y_{n-1})$ are independent.
(c) (d) (e) Since $Y_i = X_{(n)} - X_{(i)}$, they are not independent.

6. (Casella and Berger Problem 6.40) Let $X_1, \ldots, X_n$ be iid observations from a location-scale family. Let $T_1(X_1, \ldots, X_n)$ and $T_2(X_1, \ldots, X_n)$ be two statistics that both satisfy

$$T_i\left(ax_1 + b, \ldots, ax_n + b\right) = aT_i\left(x_1, \ldots, x_n\right)$$

for all values of $x_1, \ldots, x_n$ and b and for any $a > 0$.

(a) Show that $\frac{T_1}{T_2}$ is an ancillary statistic.

Proof. Suppose that $X_1, \ldots, X_n$ comes from the location-scale family $\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right)$ with the standard pdf f. Let $Z_1, \ldots, Z_n$ be an iid sample from f with $X_i = \frac{Z_i - \mu}{\sigma}$. The distribution function for $\frac{T_1}{T_2}$ is

$$\begin{aligned} \mathbb{P}\left(\frac{T_1(X_1, \ldots, X_n)}{T_2(X_1, \ldots, X_n)} \le t\right) &= \mathbb{P}\left(\frac{T_1\left(\frac{Z_1-\mu}{\sigma}, \ldots, \frac{Z_n-\mu}{\sigma}\right)}{T_2\left(\frac{Z_1-\mu}{\sigma}, \ldots, \frac{Z_n-\mu}{\sigma}\right)} \le t\right) \\ &= \mathbb{P}\left(\frac{T_1(Z_1, \ldots Z_n)}{T_2(Z_1, \ldots Z_n)} \le t\right), \end{aligned}$$

which does not depend on μ and σ. Thus, $\frac{T_1}{T_2}$ is an ancillary statistic. $\square$

(b) Let R be the sample range and S be the sample standard deviation. Verify that R and S satisfy the above condition so that $\frac{R}{S}$ is an ancillary statistic.

Proof. Since

$$\begin{aligned} R(ax_1+b,\ldots,ax_n+b) &= (ax_{(n)}+b)-(ax_{(1)}+b) \\ &= a(x_{(n)}-x_{(1)}) \\ &= aR(x_1,\ldots,x_n) \end{aligned}$$

$$\begin{aligned} S(ax_1+b,\ldots,ax_n+b) &= \sqrt{\sum_{i=1}^{n}\left[(ax_i+b)-(a\overline{x}_n+b)\right]^2} \\ &= a\sqrt{\sum_{i=1}^{n}(x_i-\overline{x}_n)^2} \\ &= aS(x_1,\ldots,x_n), \end{aligned}$$

R and S satisfy the above condition so that $\frac{R}{S}$ is an ancillary statistic. $\square$

7. (Casella and Berger Problem 6.15) Let $X_1,\ldots,X_n$ be iid $\mathcal{N}(\theta,a\theta^2)$, where a is a known constant and $a>0$.

(a) Show that the parameter space does not contain a two-dimensional open set.

Proof. The pdf of X_i is $f(x|\theta)=\frac{1}{\sqrt{2\pi a\theta^2}}e^{-\frac{1}{2a\theta^2}(x-\theta)^2}=e^{-\frac{1}{2a}}\cdot\frac{1}{\sqrt{2\pi a\theta^2}}\cdot e^{\frac{1}{a\theta}x-\frac{1}{2a\theta^2}x^2}$, which is in the 2-parameter exponential family with parameter space $\left\{\left(\frac{1}{a\theta},-\frac{1}{2a\theta^2}\right):\theta\neq 0\right\}$. The parameter space does not contain a two-dimensional open set, as it consists of only the points on a parabola. $\square$

(b) Show that the statistic $T=(\overline{X},S^2)$ is a sufficient statistic for θ, but the family of distributions is not complete.

Proof. By Theorem 6.2.10, we have $T'=\left(\sum_{i=1}^n X_i,\sum_{i=1}^n X_i^2\right)$ is a sufficient statistic for θ. Let $g(x_1,x_2)=\left(\frac{1}{n}x_1,\frac{n}{n-1}\left(\frac{1}{n}x_2-\frac{1}{n^2}x_1^2\right)\right)$. Since g is a one-to-one function with $g^{-1}(y_1,y_2)=\left(ny_1,n\left(\frac{n}{n-1}+y_1^2\right)\right)$, we have that $T=(\overline{X},S^2)=g(T')$ is also a sufficient statistic for θ.
Since $(n-1)\frac{S^2}{a\theta^2}\sim\chi^2_{n-1}$, we have $\mathbb{E}_\theta(S^2)=a\theta^2$. Also, $\mathbb{E}_\theta(\overline{X}^2)=\mathrm{Var}_\theta(\overline{X})+[\mathbb{E}_\theta(\overline{X})]^2=\frac{1}{n}a\theta^2+\theta^2$. Let $h(x_1,x_2)=x_1^2-\frac{1}{a}\left(\frac{a}{n}+1\right)x_2$, then $h(\overline{X},S^2)=0$. Since $\overline{X}$ and S^2 are independent, $\mathbb{P}_\theta(h(T)=0)<1$ for some θ (Otherwise, $\overline{X}^2=cS^2$ for some constant c almost surely). So T is not complete. $\square$

8. (Casella and Berger Problem 6.21) Let X be one observation from the pdf

$$f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, \quad x = -1, 0, 1, \quad 0 \leq \theta \leq 1.$$

(a) Is X a complete sufficient statistic?

Since $f(x|x=t) = \begin{cases} 1 & , \ t \in \{-1,0,1\} \\ 0 & , \ \text{otehrwise} \end{cases}$, which is free of θ, and thus X is a sufficient statistic.
Notice that $\mathbb{E}_\theta(X) = -1 \times \frac{\theta}{2} + 0 \times (1-\theta) + 1 \times \frac{\theta}{2} = 0$, let $g(x) = x$, then $\mathbb{E}_\theta[g(X)] = 0$ for all θ while $\mathbb{P}_\theta(g(X) = 0) = 1 - \theta$ which is not equal to 1 for all θ. Therefore, X is not complete. So X is not a complete sufficient statistic.

(b) Is $|X|$ a complete sufficient statistic?

Since $f(x|\theta) = \theta \mathbb{1}_{\{|x|=1\}} + (1-\theta)\mathbb{1}_{\{|x|=0\}}$, by Fisher-Neymann Factorization Theorem, we have $|X|$ is a sufficient statistic. Suppose that g is a function such that $\mathbb{E}_\theta[g(|X|)] = 0 \ \forall \ \theta$, then

$$0 = \mathbb{E}_\theta[g(|X|)] = g(1)\theta + g(0)(1-\theta) = g(0) + \theta[g(1) - g(0)].$$

which implies $g(1) = g(0) = 0$. Therefore, $\mathbb{P}_\theta(g(|X|) = 0) = 1$ since the support of $|X|$ is $\{0, 1\}$. So $|X|$ is complete. Thus, $|X|$ is a complete sufficient statistic.

(c) Does $f(x|\theta)$ belong to the exponential class?

For $\theta \in (0,1)$, $f(x|\theta) = \mathbb{1}_{\{x \in \{-1,0,1\}\}} e^{|X| \ln \frac{\theta}{2} + (1-|X|)\ln(1-\theta)} = \mathbb{1}_{\{x \in \{-1,0,1\}\}} \cdot (1-\theta) \cdot e^{|X| \ln \frac{\theta}{2(1-\theta)}}$.
For $\theta = 0$, $f(x|\theta) = \frac{1}{2}\mathbb{1}_{\{|x|=1\}}$. For $\theta = 1$, $f(x|\theta) = \mathbb{1}_{\{x=0\}}$.
Sicne the support of X depends on θ, $f(x|\theta)$ is not in the exponential class.