# Solution.

## 1

$X \sim f_\theta(x) = P(X = x) = -(x \log\theta)^{-1}(1-\theta)^x, x = 1, 2, ..,$

(a) For any $t > 1$. Compute

$$\int_0^t g_\theta(x)dx = \sum_{x=1}^{[t]-1} -(x\log\theta)^{-1}(1-\theta)^x + -(t-[t])([t]\log\theta)^{-1}(1-\theta)^{[t]}$$

which equals to

$$\sum_{x=1}^{[t]-1} P(X = x) + P(X = [t], U \le t - [t]) = P(X + U \le t)$$

(b) For any $0 < \theta_1 < \theta_2 < 1$,

$$\frac{g_{\theta_2}(t)}{g_{\theta_1}(t)} = \left(\frac{1-\theta_2}{1-\theta_1}\right)^{[t]} \frac{\log\theta_1}{\log\theta_2}$$

which is a nondecreasing function in $t$ .

(c) Since $\{g_\theta\}$ has MLR in $X + U$, the UMP test for $H_0 : \theta \le \theta_0$ v.s. $H_1 : \theta > \theta_0$ for a given $\theta_0$ is $T = 1\{X + U > c\}$ (since $X + U$ has Lebesgue density, $P(X + U = t) = 0$). To have size $\alpha$, $c = c(\theta_0)$ is determined by

$$\alpha = \int_{c(\theta_0)}^{\infty} g_{\theta_0}(t)dt$$

(d) Note that for any $\theta \in (0, 1)$ the c.d.f of $X + U$ : $F_{X+U,\theta}(t)$ is a strictly increasing continuous function in $t$ , then $F_{X+U,\theta}(X + U) \sim U[0, 1]$ . From Lemma 6.3 $F_{X+U,\theta}(X + U)$ is a continuous nonincreasing function in $\theta$.Then $[\underline{\theta}, 1)$ is a lower confidence bound for $\theta$ with confidence coefficient $1 - \alpha$, where

$$\underline{\theta} = \inf\{\theta : F_{X+U,\theta}(X + U) \le 1 - \alpha\}.$$

(e) Since $T_{\theta_0} = 1\{X + U > c(\theta_0)\}$ is the size $\alpha$ UMP test for $H_0 : \theta = \theta_0$ v.s. $H_1 : \theta > \theta_0$, the UMA lower confidence interval for $\theta$ with confidence coefficient $1 - \alpha$ is

$$\{\theta \in (0, 1) : \ c(\theta) \ge X + U\}$$

Since $c(\theta) \ge X + U$ is equivalent to $F_{X+U,\theta}(c(\theta)) \ge F_{X+U,\theta}(X + U)$ which is $1 - \alpha \ge F_{X+U,\theta}(X + U)$. So the confidence interval in (d) and (e) are the same.

## 2

Since the p.d.f. $\frac{1}{2\sigma} f(\frac{|x-\mu|}{\sigma})$ is symmetric around $\mu$, the mean and median are $\mu$. Denote $\bar{X}_n$ as sample mean and $m_n$ as sample median, from theorems

$$\sqrt{n}(\bar{X}_n - \mu) \to_d N(0, \sigma_1^2)$$

and

$$\sqrt{n}(m_n - \mu) \to_d N(0, \sigma_2^2)$$

where $\sigma_1^2 = Var(X_1) = 2\int_0^\infty \frac{t^2}{2\sigma} f(\frac{t}{\sigma})dt = 2\int_0^\infty \frac{t^2}{2\sigma} f(\frac{t}{\sigma})dt = \sigma^2 \int_0^\infty t^2 f(t)dt$ and $\sigma_2^2 = \frac{\sigma^2}{(f(0))^2}$ Thus $e_{\bar{X},m_n} = \frac{1}{(f(0))^2 \int_0^\infty t^2 f(t)dt}$

**3**

(a) $l(\theta) = 2^{-n}(1-\theta^2)^n e^{\theta \sum X_i - \sum |X_i|}$ . $\frac{\partial}{\partial\theta} \log l(\theta) = \sum X_i - \frac{2n\theta}{1-\theta^2}$ So we have two sequence of solutions

$$\hat{\theta}_n^{\pm} = -(\bar{X}_n)^{-1} \pm (\bar{X}_n)^{-1}\sqrt{(\bar{X}_n)^2 + 1}\ a.s.$$

where $\bar{X}_n = \sum_{i=1}^n X_i/n$ . Since $\theta \in (-1,1)$ only one sequence of solutions should be kept :

$$\hat{\theta}_n^{+} = -(\bar{X}_n)^{-1} + (\bar{X}_n)^{-1}\sqrt{(\bar{X}_n)^2 + 1}\ a.s.$$

From LLN $\bar{X}_n \rightarrow_{a.s.} EX_1 = \frac{2\theta}{1-\theta^2}$ . Then for $\theta \neq 0$, $\hat{\theta}_n^+ \rightarrow_{a.s.} \theta$ which is consistent . For $\theta = 0$, $\hat{\theta}_n^+ \rightarrow_{a.s} 0$ which is also consistent. Since the distribution belongs to a full rank exponential family, the regularity condition in Thm 4.16 holds. Thus from Thm 4.17

$$\sqrt{n}(\hat{\theta}_n^+ - \theta) \rightarrow N(0, I_1^{-1}(\theta))$$

where $I_1(\theta) = \frac{1+\theta^2}{(1-\theta^2)^2}$

(b) The population belongs to one parameter exponential family with $\theta$ as the natural parameter. a UMPU test for testing $H_0 : \theta = 0$ v.s. $H_1 : \theta \neq 0$ is of form $T = 1\{\bar{X}_n < c_1 \text{ or } \bar{X}_n > c_2\}$ where $c_1, c_2$ are constants determined by

$$E_{\theta=0} T = \alpha,\ E_{\theta=0}(\bar{X}_n T) = \alpha E_{\theta=0}(\bar{X}_n) = 0.$$

Since when $\theta = 0$, the density is symmetric around 0. We can choose $-c_1 = c_2 = c > 0$, then the second condition is automatically satisfied. Then $T = 1\{\bar{X}_n < -c \text{ or } \bar{X}_n > c\}$ and $c$ satisfies

$$\int_{\{x_1+..+x_n > nc\}} 2^{-n} e^{-\sum_{i=1}^n |x_i|} dx_1..x_n = \alpha/2$$

(c)$l_n(\theta) = 2^{-n}(1-\theta^2)^n e^{\theta \sum X_i - \sum |X_i|}$ . $s_n(\theta) = \frac{\partial}{\partial\theta} \log l(\theta) = \sum X_i - \frac{2n\theta}{1-\theta^2}$. $I_n(\theta) = \frac{2n+2n\theta^2}{(1-\theta^2)^2}$. MLE $\hat{\theta}_n = -(\bar{X}_n)^{-1} + (\bar{X}_n)^{-1}\sqrt{(\bar{X}_n)^2+1}$. Consider testing $H_0 : \theta = \theta_0$ v.s. $H_1 : \theta \neq \theta_0$. The LRT statistics is

$$\lambda_n(\mathbf{X}, \theta_0) = \frac{l_n(\theta_0)}{l_n(\hat{\theta}_n)} = \left( \frac{(\bar{X}_n)^2(1-\theta_0^2)}{2(\sqrt{(\bar{X}_n)^2+1} - 1)} \right)^n e^{n(\theta_0 \bar{X}_n + 1 - \sqrt{(\bar{X}_n)^2+1})}$$

The Wald test statistics is $W_n(\mathbf{X}, \theta_0) = (\hat{\theta}_n - \theta_0)^2 I_n(\hat{\theta}_n)$ and the Rao test statistics is $R_n(\mathbf{X}, \theta_0) = s_n^2(\theta_0)/I_n(\theta_0) = \frac{\left((1-\theta_0^2)\bar{X}_n - 2\theta_0\right)^2}{2(1+\theta_0^2)}$ . Then we can construct three $1-\alpha$ asymptotically correct confidence sets by inverting the acceptance regions

$$C_1(\mathbf{X}) = \{\theta \in (-1,1), :\ \lambda_n(\mathbf{X}, \theta) \geq e^{-\chi^2_{1,\alpha}/2}\}$$

$$C_2(\mathbf{X}) = \{\theta \in (-1,1), :\ W_n(\mathbf{X}, \theta) \leq \chi^2_{1,\alpha}\}$$

$$C_3(\mathbf{X}) = \{\theta \in (-1,1), : \; R_n(\mathbf{X},\theta) \le \chi^2_{1,\alpha}\}$$

(d) The Bayes estimator $\tilde{\theta}$ under the square error loss is the posterior mean $E(\theta|X)$. Thus

$$\tilde{\theta}(X) = \frac{\int_{-1}^{1} \theta(1-\theta^2)e^{\theta X - |X|}d\theta}{\int_{-1}^{1}(1-\theta^2)e^{\theta X - |X|}d\theta} = \frac{e^{2X}}{2} - 7$$