

Name:

SSN:

Section #

all steps for each problem to receive full credit.

I pledge my honor that I have abided by the Honor System. \_\_\_\_\_

(Signature)

1. Let  $X_1, \dots, X_n$  be a random sample from a distribution with probability density function

$$\text{(pdf), } f(x; \theta) = \frac{\theta \exp\{-|x|^{\theta}\}}{2\Gamma(1/\theta)}, x \text{ on the real line, where } \theta > 0. \text{ Suppose that}$$

$\Omega = \{ \theta : \theta = 1, 2 \}$ . Consider the hypotheses  $H_0 : \theta = 1$  (a double exponential distribution) versus  $H_1 : \theta = 2$  (a normal distribution, here  $\Gamma(1/2) = \sqrt{\pi}$ ). Derive the maximum likelihood test and

show that it is based on  $\sum_{i=1}^n X_i^2 - \sum_{i=1}^n |X_i|$ . How would one setup the test (find the critical values

for the rejection region of  $H_0$ ) based on large sample size? (15 points)

(a)  $H_0: \theta = 1$   $\hat{\alpha}$  significance level  $\alpha$

$$H_1: \theta \neq 1 (\theta = 2)$$

$$\Lambda = \frac{\max_{\theta=1,2} L(\theta)}{L(1)} = \frac{L(1)}{L(1) \vee L(2)} < c < 1$$

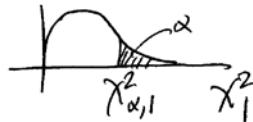
$$\Rightarrow \Lambda = \frac{L(1)}{L(2)} < c, \text{ where } L(\theta) = \frac{\theta^n \exp\left\{-\sum_{i=1}^n |x_i|^{\theta}\right\}}{2^n \Gamma(1/\theta)^n}.$$

$$\frac{L(1)}{L(2)} = \frac{\exp\left\{-\sum_{i=1}^n |x_i|\right\}}{\frac{2^n}{2^n} \exp\left\{-\sum_{i=1}^n x_i^2\right\}} < c$$

$$\exp\left\{\sum_{i=1}^n x_i^2 - \sum_{i=1}^n |x_i|\right\} < 2^n \frac{c}{(\pi)^{n/2}}$$

(b) Reject  $H_0$  if  $\sum_{i=1}^n x_i^2 - \sum_{i=1}^n |x_i| < c^* = \ln \int \frac{2^n C}{(\pi)^{n/2}}$

$$-2 \ln \Lambda \stackrel{d}{\rightarrow} \chi^2_1, \text{ Reject } H_0 \text{ if } -2 \ln \Lambda > \chi^2_{\alpha, 1}$$



2. Let  $X$  have a gamma distribution with  $\alpha = 3$  and  $\beta = \theta > 0$ .

a. Find the Fisher information  $I(\theta)$ .

b. If  $X_1, \dots, X_n$  is a random sample from this distribution, show that the derived MLE (Maximum Likelihood Estimator) of  $\theta$  is an efficient estimator of  $\theta$ .

c. What is the asymptotic distribution of  $\sqrt{n}(\hat{\theta} - \theta)$ ? (20 points)

$$(a) f(x, \theta) = \frac{x^{3-1} e^{-x/\theta}}{\Gamma(3) \theta^3}, x > 0, \ln f(x, \theta) = \frac{2 \ln x - x - 3 \ln \theta}{\theta} - \ln \Gamma(3)$$

$$\frac{\partial}{\partial \theta} \ln f(x, \theta) = \frac{x - 3}{\theta^2}, \quad \frac{\partial^2}{\partial \theta^2} \ln f(x, \theta) = -\frac{2x}{\theta^3} + \frac{3}{\theta^2}$$

$$I(\theta) = -E \frac{\partial^2}{\partial \theta^2} \ln f(x, \theta) = \frac{2EX}{\theta^3} + \frac{-3}{\theta^2} = \frac{(2)3\theta}{\theta^3} - \frac{3}{\theta^2} \\ = \frac{3}{\theta^2} \Rightarrow nI(\theta) = \frac{3n}{\theta^2} - (\star)$$

$$(b) L(\theta) = \frac{(\prod x_i)^2}{2^n \theta^{3n}} e^{-\sum x_i/\theta} \quad l(\theta) = \ln \frac{(\prod x_i)^2}{z^n} - \frac{\sum x_i}{\theta} - 3n \ln \theta.$$

$$l'(\theta) = \frac{\sum x_i}{\theta^2} - \frac{3n}{\theta} \Rightarrow \hat{\theta} = \bar{X}, E \bar{X} = EX = \frac{3\theta}{3} = \theta.$$

$$l''(\theta) = -\frac{2\sum x_i}{\theta^3} + \frac{3n}{\theta^2} \quad l''(\hat{\theta}) = -\frac{2(3n)}{\theta^3} + \frac{3n}{\theta^2} = \frac{-3n}{\theta^2} < 0.$$

Thus,  $\hat{\theta} = \bar{X}$  is M.L.E.  $\text{Var}(\bar{X}) = \frac{1}{9} \text{Var}(X) = \frac{3\theta^2}{9n} = \frac{\theta^2}{3n}$   
 $= \frac{1}{nI(\theta)}$  from (\*). Hence  $\hat{\theta}$  is efficient estimator.

$$(c) \sqrt{n} \left( \bar{X} - \theta \right) \xrightarrow{d} N(0, \frac{1}{I(\theta)}) = N(0, \frac{\theta^2}{3})$$

(Also from C.L.T.  $\sqrt{n}(\bar{X} - 3\theta) \xrightarrow{d} N(0, 3\theta^2)$ ,)  
which confirms the result.

3. It is known that the probability  $p$  of tossing head on an unbalanced coin is either  $\frac{1}{4}$  or  $\frac{3}{4}$ . The coin is tossed twice and a value for  $Y$ , the number of heads, is observed. Based on the observed data  $Y$ , find the MLE. (Also find the MLE of  $P(Y \geq 1)$ ). (15 points)

(a)  $Y$ : # of heads out of  $n=2$

$$L(p) = P(Y=y) = \binom{2}{y} p^y (1-p)^{2-y}$$

$$\hat{p} = \text{MLE} = \begin{cases} \frac{1}{4} & \text{if } L\left(\frac{1}{4}\right) > L\left(\frac{3}{4}\right) \equiv y < 1 \\ \frac{3}{4} & \text{if } L\left(\frac{1}{4}\right) < L\left(\frac{3}{4}\right) \equiv y > 1 \end{cases}$$

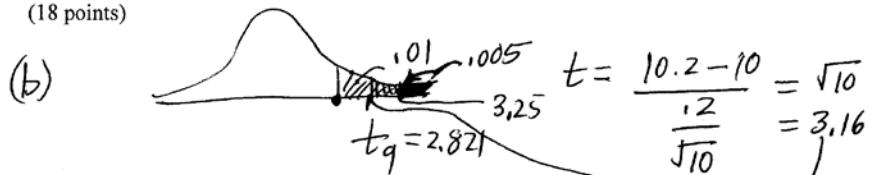
$$L\left(\frac{1}{4}\right) = L\left(\frac{3}{4}\right)$$

$$\binom{2}{y} \left(\frac{1}{4}\right)^y \left(\frac{3}{4}\right)^{2-y} > \binom{2}{y} \left(\frac{3}{4}\right)^y \left(\frac{1}{4}\right)^{2-y}$$

$$3^{2-y} > 3^y$$

$$\Rightarrow 2-y > y \Rightarrow y < 1 \quad \text{MLE } P(Y \geq 1) = \frac{1 - (1-\hat{p})^2}{1 - (1-\hat{p})^2}, \text{ where } \hat{p} \text{ is from 3(a).}$$

4. Assume that the weight of cereal in a "10-ounce box" is  $N(\mu, \sigma^2)$ . To test  $H_0: \mu = 10$  against  $H_1: \mu > 10$ . A random sample of size  $n = 10$  is taken and observed that  $\bar{x} = 10.2$  and  $s = 0.2$ . Does one reject  $H_0$  at the 1% significance level? What is the approximate p-value of the test? Justify the distribution of the test statistics used under  $H_0$ . (18 points)



$$.01 > \text{p-value} > .005$$

"  $\boxed{\text{p-value}} = P(t_q > 3.16)$

(a) Since p-value < .01 Reject  $H_0$ .

$$(c) t = \frac{\bar{X} - 10}{S / \sqrt{n}} = \left( \frac{\bar{X} - 10}{\frac{\sigma}{\sqrt{n}}} \right) / \sqrt{\frac{(n-1)S^2}{(n-1)\sigma^2}}$$

$\frac{\bar{X} - 10}{\sigma / \sqrt{n}} \sim N(0, 1)$  and is independent of  $S^2$ .  
 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ . Hence  $t \sim t_{n-1=9}$ .

5. Let  $Y_1 < Y_2 < Y_3 < Y_4$  be the order statistics of a random sample of size  $n = 4$  from a distribution with pdf  $f(x; \theta) = \frac{1}{\theta}, 0 < x < \theta$ , zero elsewhere. The hypothesis  $H_0: \theta = 2$  is rejected and  $H_1: \theta > 2$  is accepted if the observed  $Y_4 \geq c$ . Find the constant  $c$  so that the significance level is 1%. Determine the power function of the test. (15 points)

$$\begin{aligned}
 (a) .01 &= P(Y_4 \geq c | \theta = 2) = 1 - P(Y_4 < c | \theta = 2) \\
 &= 1 - [P(X_1 < c | \theta = 2)]^n = 1 - \left(\frac{c}{2}\right)^n, n=4 \\
 \Rightarrow \left(\frac{c}{2}\right)^4 &= .99 \quad c^4 = (.99)16 \quad c = 2\sqrt[4]{.99} \\
 &= 1.99498114
 \end{aligned}$$

$$\begin{aligned}
 (b) \gamma(\theta) &= P(Y_4 \geq 1.99498114), \theta > 2 \\
 &= 1 - P(X_1 < 1.99498114)^4 = 1 - \left(\frac{1.99498114}{\theta}\right)^4 = 1 - \frac{(.99)16}{\theta^4} \\
 &= 1 - \frac{15.84}{\theta^4}
 \end{aligned}$$

6. Obtain the probability that an observation is a potential outlier for the Laplace distribution given by  $f(x) = (1/2) \exp\{-|x|\}$ . (17 points)

$$\begin{aligned}
 P(X < \xi_{.25}) &= .25 \\
 \text{Ans} &= 1 - (1 - .0625) = .0625 \\
 &\quad \left| \begin{array}{l} \frac{1}{2} \int_{-\infty}^{\xi_{.25}} e^{-|x|} dx = \frac{e^{-x}}{2} \Big|_{-\infty}^{\xi_{.25}} = .25 \\ \xi_{.25} = .5 \\ \xi_{.25} = -0.693147181 \end{array} \right. \\
 \frac{1}{2} + P(0 < X < \xi_{.75}) &= .75
 \end{aligned}$$

$$\begin{aligned}
 P(0 < X < \xi_{.75}) &= .25 \\
 h &= \xi_{.75} - \xi_{.25} = 1.386294361 \quad \frac{1}{2} \int_0^{\xi_{.75}} e^{-|x|} dx = .25 \quad \left. \frac{e^{-x}}{-1} \right|_0^{\xi_{.75}} = .5 \\
 &\quad \left. \begin{array}{l} 1.5h = 2.079441542 \\ 1 - e^{-\xi_{.75}} = .5 \\ e^{-\xi_{.75}} = .5 \end{array} \right. \\
 P(-2.775 < X < 2.772588723) &= \int_{-2.775}^{2.772588723} e^{-|x|} dx = \left. \frac{e^{-x}}{-1} \right|_{-2.775}^{2.772588723} = 1 - e^{-2.772588723} = 1 - e^{-2.775} = 1 - .0625
 \end{aligned}$$