STA 261S, Winter 2004, Test #1
(Feb. 11, 2004. Duration: 100 minutes.)

SOLUTIONS

1. Let Ω = S = [0, 1], and let \( L_0(θ \mid s) = e^θ \). Determine (with explanation) whether or not each of the following likelihood functions is equivalent to the likelihood function \( L_0(θ \mid s) \).

(a) \( L_1(θ \mid s) = s^2 + e^θ \).

Solution. Here \( \frac{L_1(θ \mid s)}{L_0(θ \mid s)} = \frac{s^2 + e^θ}{e^θ} = s^2 e^{-θ} + 1 \), which depends on θ. Hence, \( L_1 \) is NOT equivalent to \( L_0 \).

(b) \( L_2(θ \mid s) = e^{s^2 + θ} \).

Solution. Here \( \frac{L_2(θ \mid s)}{L_0(θ \mid s)} = \frac{e^{s^2 + θ}}{e^θ} = e^{s^2} \), which does not depend on θ. Hence, \( L_2 \) IS equivalent to \( L_0 \).

(c) \( L_3(θ \mid s) = e^{s^2 + θ} \).

Solution. Here \( \frac{L_3(θ \mid s)}{L_0(θ \mid s)} = \frac{e^{s^2 + θ}}{e^θ} = e^{(s^2 − 1)θ} \), which depends on θ. Hence, \( L_3 \) is NOT equivalent to \( L_0 \).

2. Let Ω = S = (0, 1). Suppose the likelihood function, given an observation \( s \in S \), is given by \( L(θ \mid s) = θ^{2s} (1 − θ)^{4s} \), for \( θ \in Ω \).

(a) Compute (with explanation) the Score Function for this likelihood.

Solution. Here \( \ell(θ \mid s) = \log L(θ \mid s) = 2s \log(θ) + 4s \log(1 − θ) \), so the Score Function is \( S(θ \mid s) = \frac{∂}{∂θ} \ell(θ \mid s) = \frac{2s}{θ} − \frac{4s}{1−θ} \).

(b) Solve (with explanation) the corresponding Score Equation.

Solution. The Score Equation is \( S(θ \mid s) = 0 \), which is equivalent to \( 2s(1 − θ) − 4s(1 − θ) − 2s = 0 \), or \( θ = 2s/6s = 1/3 \).

(c) Determine (with explanation) the MLE, \( \hat{θ} \), for θ.

Solution. Here the derivative \( S(θ \mid s) \) is well-defined throughout Ω. And, the second derivative \( \left( \frac{∂}{∂θ} \right)^2 \ell(θ \mid s) = −2sθ^2 \) \(−4s(1 − θ)^{-2} < 0 \) for all \( θ \in Ω \) and \( s \in S \). And, on the boundary as \( θ → 0 \) or \( θ → 1 \), the likelihood goes to 0. Hence, the solution to the Score Equation must be a global maximum, so \( \hat{θ} = 1/3 \).
Let $\Omega = (0, \infty)$, $S = [6, \infty)$, and $P_\theta = \text{Uniform}[6, 5\theta + 6]$ for $\theta \in \Omega$. Suppose we observe the observations $x_1, x_2, \ldots, x_n$, with $x_i \geq 6$ for all $i$.

(a) Compute (with full explanation) the MLE, $\hat{\theta}$, for $\theta$.

**Solution.** The density of $P_\theta$ is equal to $1/5\theta$ for $6 \leq x_i \leq 5\theta + 6$, otherwise 0. Hence, the likelihood function $L(\theta \mid x_1, \ldots, x_n)$ is equal to $(1/5\theta)^n$ provided that $6 \leq x_i \leq 5\theta + 6$ for all $i$, i.e. $\max_{1 \leq i \leq n} x_i \leq 5\theta + 6$, otherwise it equals 0. Hence, the likelihood is maximised when $(1/5\theta)^n$ is as large as possible (i.e., $\theta$ is as small as possible), subject to the constraint that $\max_{1 \leq i \leq n} x_i \leq 5\theta + 6$, i.e. $5\theta + 6 \geq \max_{1 \leq i \leq n} x_i$, i.e. $\theta \geq [(\max_{1 \leq i \leq n} x_i) - 6]/5$. The smallest $\theta$ satisfying this constraint is $\hat{\theta} = [\max_{1 \leq i \leq n} (x_i - 6)]$, which is the MLE.

(b) Compute (with explanation) the MLE for $\theta^2$.

**Solution.** Since the mapping $\theta \mapsto \theta^2$ is 1–1 on $S$, we can use the “Plug-In Estimator” as the MLE for $\theta^2$. Thus the MLE for $\theta^2$ is equal to $(\hat{\theta})^2 = \left(\frac{1}{n} \max_{1 \leq i \leq n} (x_i - 6)\right)^2 = \frac{1}{2n} \max_{1 \leq i \leq n} (x_i - 6)^2$.

4. Suppose we observe three observations: $x_1 = 2$, $x_2 = 3$, $x_3 = 7$.

(a) Compute $\bar{x}$ and $S^2$. [Provide actual numbers, not just formulae.]

**Solution.** $\bar{x} = \frac{1}{3}[2 + 3 + 7] = 12/3 = 4$.

$S^2 = \frac{1}{3-1}[(2 - 4)^2 + (3 - 4)^2 + (7 - 4)^2] = \frac{1}{2}[4 + 1 + 9] = 14/2 = 7$.

(b) Suppose the statistical model is a Location-Scale Model, with $\Omega = \mathbb{R} \times (0, \infty)$, and $P_{(\mu, \sigma^2)} = N(\mu, \sigma^2)$ for $(\mu, \sigma^2) \in \Omega$. Compute (with explanation) a 95% confidence interval for $\mu$. [You should provide an explicit numerical formula, but you do not need to simplify arithmetic expressions. You may use the facts that if $T_2 \sim t(2)$, $T_3 \sim t(3)$, and $T_4 \sim t(4)$, then $P[T_2 \leq -2.92] = P[T_3 \leq -2.35] = P[T_4 \leq -2.13] = 0.05$, and $P[T_2 \leq -4.30] = P[T_3 \leq -3.18] = P[T_4 \leq -2.78] = 0.025$.

**Solution.** We know that under $P_{\theta}$, $T \equiv \sqrt{n}/S^2 (X - \mu) \sim t(n - 1)$, i.e. $T \equiv \sqrt{\frac{1}{n-1}} (X - \mu) \sim t(2)$. Hence, $P[-4.30 < T < +4.30] = 1 - P[T \leq -4.30] - P[T \geq +4.30] = 1 - 2 P[T \leq -4.30] = 1 - 2(0.025) = 0.95$. Thus, $P[-4.30 < T < +4.30] = P[X - 4.30 \sqrt{\frac{7}{3}} < \mu < X - 4.30 \sqrt{\frac{7}{3}}]$. Hence, a 95% C.I. is $(\bar{x} - 4.30 \sqrt{\frac{7}{3}}, \bar{x} + 4.30 \sqrt{\frac{7}{3}}) = (4 - 4.30 \sqrt{\frac{7}{3}}, 4 + 4.30 \sqrt{\frac{7}{3}})$. [This equals $(-2.57, 10.57)$, but you don’t need to compute that.]

(c) Suppose the statistical model is a Location Model, with $\Omega = \mathbb{R}$, and $P_\theta = N(\theta, 4)$ for
\( \theta \in \Omega \). Compute (with explanation) a P-value for the null hypothesis \( H_0 : \theta = 6 \) versus the alternative hypothesis \( H_1 : \theta \neq 6 \). [You may leave your answer in terms of the \( \Phi \) function.]

**Solution.** We know that under \( P_0 \), \( Z \equiv \sqrt{n/\sigma^2}(\bar{X} - 6) = \sqrt{3/4}(\bar{X} - 6) \sim N(0,1) \). The observed value of \( Z \) was \( \sqrt{3/4}(4 - 6) = -\sqrt{3} \). The probability (under \( P_0 \)) of observing a value which is at least as surprising, is equal to \( P[|Z| \geq \sqrt{3}] = 2 \Phi(-\sqrt{3}) \). [This equals 0.0833, but you don’t need to compute that.]

5. Let \( \Omega = S = \mathbb{R} \), with \( P_\theta = \text{Uniform}[\theta - 3, \theta + 3] \) for \( \theta \in \Omega \). Suppose we observe \( x_1, x_2, \ldots, x_{100} \), and that \( \bar{x} = 11 \).

(a) Find \( C_1 > 0 \) and \( C_2 \) (which may depend on \( \theta \), but may not depend on \( x_1, \ldots, x_{100} \)) such that if \( Z = C_1(\bar{X} - C_2) \), then under \( P_0 \), \( Z \) has mean 0 and variance 1. [Here \( \bar{X} \) stands for the corresponding random variable, as opposed to the observed value \( \bar{x} \). Also, recall that the Uniform\([a,b]\) distribution has mean \((a+b)/2\), and variance \((b-a)^2/12\).]

**Solution.** Here \( P_0 \) has mean \(((\theta - 3) + (\theta + 3))/2 = \theta \), and variance \(((\theta + 3) - (\theta - 3))^2/12 = 6^2/12 = 3/2 \). Hence, \( \bar{X} \) has mean \( \theta \) and variance \( 3/2n = 3/100 \).

Hence, if \( C_1 = 1/\sqrt{3/100} = 10/\sqrt{3} \) and \( C_2 = \theta \), then \( Z = C_1(\bar{X} - C_2) = 10(\bar{X} - \theta)/\sqrt{3} \) has mean 0 and variance 1 under \( P_0 \).

(b) Compute (with explanation) an approximate 95% confidence interval for \( \theta \). [Hint: Use the C.L.T.]

**Solution.** Since \( n = 100 \) is reasonably large, we can use the C.L.T. approximation to conclude that under \( P_0 \), \( Z \approx N(0,1) \), i.e. \( 10(\bar{X} - \theta)/\sqrt{3} \approx N(0,1) \).

Thus \( 0.95 \approx P[-1.96 < (\bar{X} - \theta)(10/\sqrt{3}) < +1.96] = P[\bar{X} - (\sqrt{3}/10)1.96 < \theta < \bar{X} + (\sqrt{3}/10)1.96] \). Hence, a 95% C.I. is \((\bar{x} - 1.96\sqrt{3}/10, \bar{x} + 1.96\sqrt{3}/10) = (11 - 0.196\sqrt{3}, 11 + 0.196\sqrt{3}) \). [This equals \((10.66, 11.34) \), but you don’t need to compute that.]

6. Suppose \( \Omega = S = \mathbb{R} \), and we observe two observations \( x_1 \) and \( x_2 \), and the likelihood function is given by \( L(\theta | x_1, x_2) = \exp[(x_1 - \theta)^2] \exp[2\theta x_2] \). Let \( T(x_1, x_2) = x_1 - x_2 \).

(a) Is \( T \) a sufficient statistic for \( \theta \)? (Explain your reasoning.)

**Solution.** Yes, \( T \) is sufficient. Indeed, \( L(\theta | x_1, x_2) = \exp[(x_1 - \theta)^2 + 2\theta x_2] = \exp[x_1^2 - 2\theta x_1 + \theta^2 + 2\theta x_2] = \exp[x_1^2 + \theta^2 - 2\theta T(x_1, x_2)] = h(x_1, x_2) g_\theta(T(x_1, x_2)) \), where \( h(x_1, x_2) = \exp[x_1^2] \), and \( g_\theta(t) = \exp[\theta^2 - 2\theta t] \). Hence, by the Factorisation Theorem, \( T \) is sufficient.

(b) Is \( T \) a minimal sufficient statistic for \( \theta \)? (Explain your reasoning.)

**Solution.** Yes, \( T \) is minimal.
Proof #1: Indeed, if \( L(\theta \mid x_1, x_2) = K L(\theta \mid y_1, y_2) \) for all \( \theta \in \Omega \), then

\[
L(1 \mid x_1, x_2)/L(1 \mid y_1, y_2) = L(0 \mid x_1, x_2)/L(0 \mid y_1, y_2).
\]

Hence,

\[
L(1 \mid x_1, x_2)/L(0 \mid x_1, x_2) = L(1 \mid y_1, y_2)/L(0 \mid y_1, y_2),
\]

i.e.

\[
\exp[x_1^2 + 1^2 - 2(1)T(x_1, x_2)]/\exp[x_1^2 + 0^2 - 2(0)T(x_1, x_2)]
= \exp[y_1^2 + 1^2 - 2(1)T(y_1, y_2)]/\exp[y_1^2 + 0^2 - 2(0)T(y_1, y_2)],
\]

i.e. \( \exp[1 - 2T(x_1, x_2)] = \exp[1 - 2T(y_1, y_2)] \). It follows that \( 1 - 2T(x_1, x_2) = 1 - 2T(y_1, y_2) \), and so \( T(x_1, x_2) = T(y_1, y_2) \). Hence, \( T \) is minimal.

Proof #2: If \( L(\theta \mid x_1, x_2) = K L(\theta \mid y_1, y_2) \) for all \( \theta \in \Omega \), then \( S(\theta \mid x_1, x_2) = S(\theta \mid y_1, y_2) \), i.e. \( 2\theta - 2\theta T(x_1, x_2) = 2\theta - 2\theta T(y_1, y_2) \), and so \( T(x_1, x_2) = T(y_1, y_2) \). Hence, \( T \) is minimal.

Proof #3: The solution to the Score Equation is \( \theta = T(x_1, x_2) \). Hence, since equivalent likelihoods have the same Score Equation, they also have the same value of \( T \). Hence, \( T \) is minimal.