Final Exam: Stat 426.

Moulinath Banerjee

University of Michigan

February 18, 2003

INFORMATION: The exam carries 80 points but the maximum you can score is 70 points. Half of what you score contributes towards your total grade.

(1) (a) Suppose that the diameter of a circle is a random variable having the following probability density function.

\[ f(x) = \frac{1}{8} (3x + 1), \quad 0 < x < 2 \]

and 0 otherwise. Determine the probability density function of the area of the circle.

(b) Suppose that \( X \) and \( Y \) are i.i.d. random variables and each has the p.d.f.

\[ f(x) = e^{-x}, \quad x > 0 \]

and

\[ f(x) = 0 \text{ otherwise}. \]

Let

\[ U = X/(X + Y) \quad \text{and} \quad V = X + Y. \]

Determine the joint density of \( U \) and \( V \). Identify the distribution of \( V \). Are \( U \) and \( V \) independent? (6 + 10 = 16 points)

(2) (i) In a city 30% of people are conservatives, 50% are liberals and 20% are independents. In a particular election, 60% of conservatives voted, 80% of liberals voted and 50% of independents voted. If a person is selected at random from the population and it is learnt that they did not vote in the last election, what is the chance that the randomly chosen person is a liberal ?

(ii) Let \( X_1, X_2, \ldots, X_n \) be i.i.d \( U(0, \theta) \) and let \( X(n) \) as usual, denote the maximum of the \( X_i \)'s. Show that the distribution function of \( X(n)/\theta \) is given by,

\[ G(u) = u^n, \quad 0 \leq u \leq 1. \]
Thus \( X_{(n)}/\theta \) is a pivot. Use this result to show that a level \( 1 - \alpha \) confidence interval for \( \theta \) is given by,

\[
\left[ \frac{X_{(n)}}{(1 - \alpha/2)^{1/n}}, \frac{X_{(n)}}{((\alpha/2)^{1/n}} \right].
\]

\( (6 + 7 = 13 \text{ points}) \)

(3) Let \( X_1, X_2, \ldots, X_n \) be i.i.d. \( \text{Exp}(\theta) \).

(i) Show that \( \hat{\theta} = 1/X \) is a minimal sufficient statistic for \( \theta \). (Hint: What is the MLE of \( \theta \) ?)

(ii) Show that \( \sqrt{n} (\hat{\theta} - \theta) \) converges to \( N(0, \theta^2) \). (You can use any standard results on limit distributions of MLEs.)

(iii) Find an appropriate variance stabilising transformation for \( \theta \) - i.e. find an appropriate function \( g \) such that

\[
\sqrt{n} \left( g(\hat{\theta}) - g(\theta) \right) \rightarrow_d N(0, 1)
\]

and use this result to find an approximate level \( 1 - \alpha \) asymptotic confidence interval for \( \theta \).

\( (4 + 4 + 6 = 14 \text{ points}) \)

(4) (i) Consider i.i.d. observations \( X_1, X_2, \ldots, X_n \) from a \( N(0, \sigma^2) \) population. Is the MLE of \( \sigma^2 \) minimal sufficient in this model?

(ii) Let \( X \) follow \( \text{Binomial}(n, \theta) \) with \( 0 < \theta < 1 \). Thus,

\[
p(X, \theta) = \binom{n}{X} \theta^X (1 - \theta)^{n-X}.
\]

Define

\[
T(X) = 1 \text{ if } X = n,
\]

and

\[
T(X) = 0, \text{ otherwise}.
\]

Thus \( T(X) \) is a Bernoulli random variable. Let,

\[
\Psi(\theta) = E_\theta(T(X)) = P_\theta(T(X) = 1).
\]

Then clearly \( T(X) \) is an unbiased estimator of \( \Psi(\theta) \).

(a) Show that the information bound obtained from the Cramer-Rao inequality for unbiased estimators of \( \Psi(\theta) \) is \( n \theta^{2n-1} (1 - \theta) \).

(b) Show that the variance of \( T(X) \) is simply \( \theta^n (1 - \theta^n) \) and that this is strictly larger than the bound obtained from the Cramer-Rao inequality for \( 0 < \theta < 1 - i.e.

\[
\theta^n (1 - \theta^n) > n \theta^{2n-1} (1 - \theta).
\]

\textbf{Note:} It can be shown that \( T(X) \) is the best unbiased estimator of \( \Psi(\theta) \). This shows that the bound obtained in the Cramer-Rao inequality is not sharp in the sense that best unbiased estimators may not achieve this bound.

\( (6 + 5 + 7 = 18 \text{ points}) \)
(5) Consider $X_1, X_2, \ldots$, be an i.i.d. sequence from the density,

$$f(x, \theta) = (\theta + 1) x^\theta, \quad 0 < x < 1.$$ 

Let $\mu_m = E(X^n_m)$ and consider the first $n$ observations, $X_1, X_2, \ldots, X_n$, from this sequence.
(i) Show that $\mu_m = (\theta + 1)/(\theta + m + 1)$ and use this result to show that the MOM estimate of $\theta$ is

$$\hat{\theta}_{MOM} = \frac{(m + 1) \mu_m - 1}{1 - \mu_m},$$

where

$$\hat{\mu}_m = \frac{1}{n} \sum_{i=1}^{n} X^n_i.$$ 

(ii) Show that for fixed $m$, $\hat{\theta}_{MOM}$ converges in probability to $\theta$.
(iii) Show that for a fixed data set $X_1, X_2, \ldots, X_n$,

$$\lim_{m \to \infty} \hat{\theta}_{MOM} = -1.$$ 

This shows that using a very high moment to estimate $\theta$ will give you a pretty bad estimate.

(Hint: Note that each $X_i$ is $< 1$ and therefore so is their maximum $X_{(n)}$ and this is larger than $(X^n_1 + X^n_2 + \ldots + X^n_n)/n$. As $m \to \infty$ what happens to $X^n_{(n)}$? What happens to $n X^n_{(n)}$?)

(iv) We know that the MLE of $\theta$ in this model is

$$\hat{\theta}_{MLE} = -\frac{n}{\sum_{i=1}^{n} \log X_i} - 1.$$ 

Show that $\hat{\theta}_{MLE}$ converges in probability to $\theta$.

(Hint: Let $Y_i = -\log X_i$. Show that $Y_i$ is an exponential random variable and use the law of large numbers to deduce that $\overline{Y}_n = -\sum_{i=1}^{n} \log X_i/n$ converges in probability to $1/(\theta + 1)$. Proceed from there. (5 + 4 + 4 + 6 = 19 points)

1 Appendix

Some useful results are stated here.

(1) (Can use for Problem 3) Let $X_1, X_2, \ldots, X_n$ be i.i.d. observations from an underlying density $f(x, \theta)$. Then the joint density of $X = (X_1, X_2, \ldots, X_n)$ is simply,

$$p(X, \theta) = \prod_{i=1}^{n} f(X_i, \theta) = L(X, \theta),$$

where $L$ is the likelihood function. The MLE, $\hat{\theta}_n$ is that value of $\theta$ (for a fixed $X$) that maximizes the likelihood function. We have,

$$I_n(\theta) = E^{\theta} \left( \frac{\partial}{\partial \theta} \log p(X, \theta) \right)^2 = n I_1(\theta),$$

3
where
\[ I_1(\theta) = E_\theta \left( \frac{\partial}{\partial \theta} \log f(X_1, \theta) \right)^2. \]

Note that the information can also be characterized in terms of the second derivative of the log-likelihood. Thus,
\[ I_n(\theta) = -E_\theta \left( \frac{\partial^2}{\partial \theta^2} \log p(X, \theta) \right) \]
and
\[ I_1(\theta) = -E_\theta \left( \frac{\partial^2}{\partial \theta^2} \log f(X_1, \theta) \right). \]

Under appropriate regularity conditions, the MLE \( \hat{\theta}_n \) satisfies,
\[ \sqrt{n}(\hat{\theta}_n - \theta) \to_d N(0, I_1(\theta)^{-1}), \]
as \( n \to \infty. \)

(2) (Can use for Problem 5) If \( W_1, W_2, W_3, \ldots \) is a sequence of i.i.d. random variables, each with mean \( \mu \) and variance \( \sigma^2 \), then the sequence
\[ \overline{W}_n = \frac{W_1 + W_2 + \ldots + W_n}{n} \]
converges in probability to \( \mu \). In other words,
\[ P(|\overline{W}_n - \mu| > \epsilon) \to 0, \]
as \( n \to \infty \) for any pre-fixed \( \epsilon > 0 \). If \( g \) is a continuous function, then \( g(\overline{W}_n) \) converges in probability to \( g(\mu) \).