Solutions to Exam #3

1. Define the following terms:

   (a) Likelihood ratio statistic

   Answer: In general, suppose we want to test the hypothesis

   $$\mathrm{H}_o: \quad \theta \in \Omega_o$$

   versus the alternative

   $$\mathrm{H}_1: \quad \theta \in \Omega_1,$$

   based on a random, $X_1, X_2, \ldots, X_n$, sample from a distribution with distribution function $f(x \mid \theta)$. The likelihood ratio statistic is given by

   $$\Lambda(x_1, x_2, \ldots, x_n) = \frac{\sup_{\theta \in \Omega_o} L(\theta \mid x_1, x_2, \ldots, x_n)}{\sup_{\theta \in \Omega} L(\theta \mid x_1, x_2, \ldots, x_n)},$$

   where $\Omega = \Omega_o \cup \Omega_1$, with $\Omega_o \cap \Omega_1 = \emptyset$, and

   $$L(\theta \mid x_1, x_2, \ldots, x_n) = f(x_1 \mid \theta) \cdot f(x_2 \mid \theta) \cdots f(x_n \mid \theta)$$

   is the likelihood function. □

   (b) Fisher information

   Answer: Given a distribution function, $f(x \mid \theta)$, with some parameter θ, the Fisher information of the parameter θ is

   $$I(\theta) = \mathrm{var}\left(\frac{\partial}{\partial \theta} \ln f(x \mid \theta) \right).$$

   □

   (c) Efficient estimator

   Answer: An unbiased estimator, W, of a parameter, θ, is said to be an efficient estimator of θ if

   $$\mathrm{var}(W) = \frac{1}{nI(\theta)},$$

   where $I(\theta)$ is the Fisher information of θ. □

2. Provide concise answers to the following questions:

   (a) State the Neyman–Pearson Lemma

   Answer: Out of all the tests at a fixed significance level, α, of the simple hypothesis $\text{H}_o\colon\ \theta = \theta_o$ versus $\text{H}_1\colon\ \theta = \theta_1$, the LRT yields the largest possible power. □

   (b) Give an example of an estimator which is a maximum likelihood estimator, but it is not unbiased.

   Answer: Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal(μ, σ^2) distribution. The statistic

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_1 - \overline{X}_n)^2$$

   is the MLE of σ^2, and it is not unbiased. □

   (c) State the Crámer–Rao inequality.

   Answer: Let W be an estimator of a parameter, θ, based on a random sample, $X_1, X_2, \ldots, X_n$, from a distribution with distribution function $f(x \mid \theta)$. Put $g(\theta) = E_\theta(W)$. The Crámer–Rao inequality states that

$$\text{var}(W) \geqslant \frac{[g'(\theta)]^2}{nI(\theta)},$$

   where $I(\theta)$ is the Fisher information.
   This inequality is valid provided that the information function, $I(\theta)$, is defined and integration and differentiation can be interchanged as in

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{\infty} h(x)f(x \mid \theta)\ \ \mathrm{d}x = \int_{-\infty}^{\infty} h(x)\frac{\partial}{\partial\theta}f(x \mid \theta)\ \ \mathrm{d}x.$$

   □

3. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Gamma$(3, \theta)$ distribution. Find the MLE for θ. Justify your answer.

   Solution: The distribution function is given by

$$f(x \mid \theta) = \frac{1}{\Gamma(3)\theta^3}\ x^2\ e^{-x/\theta} \qquad \text{for } 0 < x < \infty$$

and zero elsewhere, where $\Gamma(3) = \Gamma(2+1) = 2! = 2$. Thus, the likelihood function in this case is

$$L(\theta \mid x_1, x_2, \ldots, x_n) = \frac{1}{2^n \theta^{3n}} \, (x_1 \cdot x_2 \cdots x_n)^2 \, e^{-y/\theta},$$

where $y = \displaystyle\sum_{i=1}^{n} x_i$.

In order to find an MLE for θ, we need to maximize the function

$$\begin{aligned} \ell(\theta) &= \ln L(\theta \mid x_1, x_2, \ldots, x_n) \\ &= -3n \ln \theta - \frac{y}{\theta} + \ln\left(\frac{(x_1 \cdot x_2 \cdots x_n)^2}{2^n}\right), \end{aligned}$$

whose derivatives are

$$\ell'(\theta) = -\frac{3n}{\theta} + \frac{y}{\theta^2},$$

and

$$\ell''(\theta) = \frac{3n}{\theta^2} - \frac{2y}{\theta^3}.$$

Thus, $\widehat{\theta} = \dfrac{y}{3n}$ is a critical point of ℓ with

$$\ell'(\widehat{\theta}) = \frac{3n}{\widehat{\theta}^2} - \frac{6n\widehat{\theta}}{\widehat{\theta}^3} = -\frac{3n}{\widehat{\theta}^2} < 0.$$

Hence,

$$\widehat{\theta} = \frac{1}{3n} \sum_{i=1}^{n} X_i$$

is the MLE for θ. □

4. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a uniform distribution over the interval $[0, \theta]$ for some parameter $\theta > 0$ and let $W = 2\overline{X}_n$, where $\overline{X}_n$ denotes the sample mean.

   Compute the following:

   (a) $\text{bias}_\theta(W)$,

   (b) $\text{MSE}_\theta(W)$.

Solution:

(a) Compute

$$E(W) = E(2\overline{X}_n)2 = 2E(\overline{X}_n) = 2E(X_1) = 2\cdot\frac{\theta}{2} = \theta.$$

Thus,

$$\text{bias}_\theta(W) = E(\theta) - \theta = 0.$$

(b) Compute

$$\begin{aligned}
\text{MSE}_\theta(W) &= \text{var}(W) + [\text{bias}_\theta(W)]^2 \\
&= \text{var}(2\overline{X}_n) \\
&= 4\cdot\text{var}(\overline{X}_n) \\
&= 4\cdot\frac{\text{var}(X_1)}{n} \\
&= \frac{4}{n}\cdot\frac{\theta^2}{12} \\
&= \frac{\theta^2}{3n}.
\end{aligned}$$

□

5. Let X_1, X_2 denote two independent observations from a Bernoulli(p) distribution with parameter p, with $0 < p < 1$.

   Construct the most powerful test at a significance level $\alpha = 0.04$ to test the simple hypotheses

$$\text{H}_o: \quad p = 0.2 \quad \text{versus} \quad \text{H}_1: \quad p = 0.4.$$

   What is the power of the test?

Solution: The likelihood function is

$$L(p \mid x_2, x_2) = p^y(1-p)^{2-y},$$

where $y = x_1 + x_2$.

According to the Neyman–Pearson Lemma, the most powerful test at a given level α is provided by the LRT; that is, a test with rejection region

$$R\colon \quad \Lambda(x_1, x_2) \leqslant c,$$

for some $c \in (0, 1)$ determined by α, where

$$\Lambda(x_1, x_2) = \frac{L(0.2 \mid x_2, x_2)}{L(0.4 \mid x_2, x_2)} = \frac{16}{9}\left(\frac{3}{8}\right)^y.$$

In order to find the rejection region, R, we express the LRT in terms of the statistic

$$Y = X_1 + X_2$$

as follows:

$$\Lambda(x_1, x_2) \leqslant c$$

if and only if

$$\frac{16}{9}\left(\frac{3}{8}\right)^y \leqslant c,$$

if and only if

$$\left(\frac{3}{8}\right)^y \leqslant \frac{9c}{16}.$$

Taking the natural logarithm on both sides we obtain that

$$y \ln\left(\frac{3}{8}\right) \leqslant \ln\left(\frac{9c}{16}\right).$$

Thus, solving for y,

$$y \geqslant \frac{\ln\left(\frac{9c}{16}\right)}{\ln\left(\frac{3}{8}\right)},$$

since $\ln\left(\frac{3}{8}\right) < 0$. We then have that the rejection region for the LRT is

$$R\colon \quad Y \geqslant b,$$

for some $b > 0$, where $Y = X_1 + X_2 \sim \text{binomial}(2, p)$.
To determine the value of b that yields a significance level $\alpha = 0.04$, solve for b in the expression

$$\text{P}(Y \geqslant b) = 0.04, \quad \text{for } Y \sim \text{binomial}(2, 0.2).$$

	y		
	0	1	2
$p_Y(y \mid 0.2)$	0.64	0.32	0.04
$p_Y(y \mid 0.4)$	0.36	0.48	0.16

Table 1: Binomial$(2, p)$ Probabilities

The values of the probabilities for Y, $p_Y(y \mid p)$, under the two hypotheses are given in Table 1.

We see int he table that to get a significance level of $\alpha = 0.04$ we must have $b = 2$. Thus, the most powerful test at level $\alpha = 0.04$ rejects H_o if

$$Y \geqslant 2.$$

The power of this test is the probability that the test will reject H_o if H_1 is true; that is, if $p = 0.4$. We see from the entry in the last row and last column in Table 1 this probability is $\gamma(0.4) = 0.16$. □