MATH2033 Mathematical Statistics
Assignment 6 Suggested Solutions

1. The asymptotic variance of $\hat{\sigma}^2$ is $\frac{1}{nI(\sigma)} = \frac{\sigma^2}{n}$

$$\log f(x|\sigma) = -\log 2 - \log \sigma + \left(-\frac{|x|}{\sigma}\right)$$
$$\frac{\partial \log f(x|\sigma)}{\partial \sigma} = -\frac{1}{\sigma} + \frac{|x|}{\sigma^2}, \frac{\partial^2 \log f(x|\sigma)}{\partial \sigma^2} \qquad = \frac{1}{\sigma^2} - 2\frac{|x|}{\sigma^3}.$$

Since

$$\begin{aligned} E(|X|) &= \int_{-\infty}^{0} \frac{-x}{2\sigma} \exp\left(-\frac{-x}{\sigma}\right) \, \mathrm{d}x + \int_{0}^{\infty} \frac{x}{2\sigma} \exp\left(-\frac{x}{\sigma}\right) \, \mathrm{d}x \\ &= \int_{\infty}^{0} \frac{y}{2\sigma} \exp\left(-\frac{y}{\sigma}\right) (-\,\mathrm{d}y) + \int_{0}^{\infty} \frac{x}{2\sigma} \exp\left(-\frac{x}{\sigma}\right) \, \mathrm{d}x \\ &= 2\int_{0}^{\infty} \frac{x}{2\sigma} \exp\left(-\frac{x}{\sigma}\right) \, \mathrm{d}x = \sigma \end{aligned}$$

Clearly,

$$\begin{aligned} I(\sigma) &= -E\left[\frac{1}{\sigma^2} - 2\frac{|X|}{\sigma^3}\right] \\ &= -\frac{1}{\sigma^2} + 2\frac{E(|X|)}{\sigma^3} = -\frac{1}{\sigma^2} + 2\frac{\sigma}{\sigma^3} = \frac{1}{\sigma^2}, \end{aligned}$$

The asymptotic variance of the mle is $1/[nI(\sigma)]$, which is $\frac{\sigma^2}{n}$.

2. Since we don't know the true value of the population variance as well, we will use the fact that

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

then a 90% confidence interval for μ is $\left[\bar{x} \pm t_{n-1,0.05} \cdot \frac{s}{\sqrt{n}}\right]$, i.e.

$$\left(10 \pm 1.761 \cdot \frac{5}{\sqrt{15}}\right) \quad \to \quad (7.727, \quad 12.273)$$

Similarly, we use the fact that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

then a 90% confidence interval for σ^2 is $\left[\frac{(n-1)s^2}{\chi^2_{0.05,n-1}}, \ \frac{(n-1)s^2}{\chi^2_{0.95,n-1}}\right]$, i.e.

$$\left(\frac{14 \cdot 25}{23.685}, \ \frac{14 \cdot 25}{6.571}\right) \quad \to \quad (14.778, \quad 53.264).$$

3. A binomial random variable X can be viewed as the sum of an i.i.d. random sample $\{Y_1, \cdots, Y_n\}$ from a Bernoulli distribution with probability p.

$$f(y_i) = p^{y_i}(1-p)^{1-y_i}, \quad X = \sum_{i=1}^{n} Y_i$$

(a) Taking logarithm on the density function is

$$\log f(y|p) = y \log p + (1-y)\log(1-p)$$

Then the log-likelihood function is

$$l(p) = \sum_{i=1}^{n} \log f(Y_i|p) = \sum_{i=1}^{n} Y_i \log p + (1-Y_i)\log(1-p)$$

Then

$$\frac{\partial}{\partial p} l(p) = \frac{\sum_{i=1}^{n} Y_i}{p} - \frac{n - \sum_{i=1}^{n} Y_i}{1-p} = 0 \quad \rightarrow \quad \hat{p} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{X}{n}$$

(b) The mle is unbiased since $E(\hat{p}) = E(X/n) = \frac{1}{n}E(X) = p$. Then

$$MSE(\hat{p}) = \text{Var}(\hat{p}) = \text{Var}\left(\frac{X}{n}\right) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}$$

Then consider the Cramer-Rao's lower bound

$$I(p) = -E\left(\frac{\partial^2}{\partial p^2} \log f(Y|p)\right) = -E\left(-\frac{Y}{p^2} - \frac{1-Y}{(1-p)^2}\right) = \frac{1}{p(1-p)}$$

Thus,

$$MSE(\hat{p}) = \frac{p(1-p)}{n} = \frac{1}{nI(p)}$$

and it reaches the Cramer-Rao's lower bounded. No unbiased estimator can possibly be more precise.

(c) The statistics $n\hat{p} = X \sim \text{Bin}(n, p)$. But notice that the distribution relies on the unknown parameter p, it is impossible to use this exact one to construct a confidence interval for p. In fact, $\hat{p} = X/n$ is the sample mean of an i.i.d. sample from Bernoulli(p), with $E(\hat{p}) = p$ and $\text{Var}(\hat{p}) = p(1-p)/n$. When n is large, it follows from CLT immediately that $Z = (\hat{p} - p)/\sqrt{p(1-p)/n}$ approximately follows $N(0,1)$. After replacing the unknown $p(1-p)$ with its estimate $\hat{p}(1-\hat{p})$, an approximate 90% confidence interval for p is given by

$$\left(\hat{p} \pm z_{0.05} \cdot \sqrt{\hat{p}(1-\hat{p})/n}\right)$$

Remark: [*Normal Approximation to the Binomial Distribution*]
Since a binomial random variable is the sum of independent Bernoulli random variables, its distribution can be approximated by a normal distribution when n is large. A frequently used rule of thumb is that the approximation is reasonable when $np > 5$ and $n(1-p) > 5$. The approximation is especially useful for large values of n, for which tables are not readily available.

4. (a) The log-density function is

$$\log f(X|\theta) = -\log(\sqrt{2\pi\theta}) - \frac{X^2}{2\theta}$$

Then the score function is

$$\frac{\partial}{\partial\theta}\log f(X|\theta) = -\frac{1}{2\theta} + \frac{X^2}{2\theta^2} \quad \text{and} \quad \frac{\partial^2}{\partial\theta^2}\log f(X|\theta) = \frac{1}{2\theta^2} - \frac{X^2}{\theta^3}$$

$$I(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right) = -\frac{1}{2\theta^2} + \frac{E(X^2)}{\theta^3} = -\frac{1}{2\theta^2} + \frac{\text{Var}(X) + E(X)^2}{\theta^3} = \frac{1}{2\theta^2}$$

(b) Find the asymptotic variance of the mle $\hat{\theta}_{MLE}$ is $1/nI(\theta)$, that is, $2\theta^2/n$.

(c) The asymptotic distribution of $\sqrt{n}(\hat{\theta}_{MLE} - \theta)$ is approximately $N(0, I(\theta))$, namely, $N(0, 2\theta^2)$.

Remark: The mle of θ is unbiased since

$$\hat{\theta}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \quad \text{with mean} \quad E(\hat{\theta}_{MLE}) = \frac{n\theta}{n} = \theta$$

The exact distribution of $\dfrac{n\hat{\theta}_{MLE}}{\theta}$ is $\chi^2(n)$. An exact $(1-\alpha)100\%$ confidence interval can be constructed from the statistics $n\hat{\theta}_{MLE}/\theta$.

5. (a) As in HW4-Q1, S^2 is an unbiased estimator of σ^2, while $E(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$.

(b) Since $\frac{(n-1)S^2}{\sigma^2}$ follows a χ^2_{n-1} distribution,

$$\text{Var}\left[\frac{(n-1)S^2}{\sigma^2}\right] = \frac{(n-1)^2}{\sigma^4}\text{Var}(S^2) = 2(n-1) \quad \to \quad MSE(S^2) = \text{Var}(S^2) = \frac{2\sigma^4}{(n-1)}$$

On the other hand, $n\hat{\sigma}^2 = (n-1)S^2$, so $\frac{n\hat{\sigma}^2}{\sigma^2}$ also follows a χ^2_{n-1} distribution,

$$MSE(\hat{\sigma}^2) = \text{Var}(\hat{\sigma}^2) + \left[E(\hat{\sigma}^2) - \sigma^2\right]^2 = \frac{2(n-1)\sigma^4}{n^2} + \left[-\frac{1}{n}\sigma^2\right]^2 = \frac{(2n-1)\sigma^4}{n^2}$$

By comparing the ratio, $\hat{\sigma}^2$ has smaller MSE,

$$\frac{MSE(\hat{\sigma}^2)}{MSE(S^2)} = \frac{2(n-\frac{1}{2})(n-1)}{2n^2} < 1.$$

(c) Similarly, denoting $k\sum_{i=1}^{n}(X_i - \overline{X})^2$ as S_k^2, then $\frac{S_k^2}{k\sigma^2}$ follows a χ^2_{n-1} distribution with

$$E\left[S_k^2\right] = k(n-1)\sigma^2 \quad \text{and} \quad \text{Var}\left[S_k^2\right] = 2k^2(n-1)\sigma^4$$

Then

$$MSE(S_k^2) = 2k^2(n-1)\sigma^4 + (k(n-1)-1)^2\sigma^4 = \left[k^2(n^2-1) - 2k(n-1) + 1\right]\sigma^4$$

To minimize the MSE, let

$$\frac{\partial}{\partial k}\left[k^2(n^2-1) - 2k(n-1) + 1\right] = 2(n-1)(k(n+1)-1) = 0 \quad \to \quad k = \frac{1}{n+1}$$

Since $MSE(S_k^2)$ is a parabola with an upward opening, then $k = 1/(n+1)$ is the point where it achieves its minimum $\frac{2\sigma^4}{n+1}$.

6. Let X be the number of success within 10 independent throws, and

$$H_0: \ p = \frac{1}{2} \quad \text{v.s.} \quad H_1: \ p \neq \frac{1}{2}$$

Decision Rule: Reject H_0 if $X = 0$ or $X = 10$, then

$$P(\text{Reject } H_0 | \, H_0 \text{ is true}) = P(X=0) + P(X=10) = \binom{10}{0} 0.5^0 (1-0.5)^{10} + \binom{10}{10} 0.5^{10} (1-0.5)^0 = \frac{1}{512}$$

The level of significance is $\frac{1}{512}$.

7. The test statistic is $Z = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}}$ which follows a standard normal distribution

(a) Reject H_0 if the observation is too small, namely, $z \leq -z_{0.08}$

$$z = \frac{\bar{y} - 120}{18/\sqrt{25}} = -1.61 \leq -z_{0.08} = -1.405 \quad \rightarrow \text{Reject } H_0.$$

(b) Reject H_0 if the observation is either too large or too small, namely, $z \leq -z_{0.005}$ or $z \geq z_{0.005}$

$$z = \frac{\bar{y} - 42.9}{3.2/\sqrt{16}} = 2.75 \geq z_{0.005} = 2.575 \quad \rightarrow \text{Reject } H_0.$$

(c) Reject H_0 if the observation is too large, namely, $z \geq z_{0.13}$

$$z = \frac{\bar{y} - 14.2}{4.1/\sqrt{9}} = 1.17 \geq z_{0.13} = 1.126 \quad \rightarrow \text{Reject } H_0.$$