1. (Invariance Principle) If \( \hat{\theta} \) is the MLE’s of parameter \( \theta \) then the MLE of \( h(\theta) \) is \( h(\hat{\theta}) \) for some function \( h \).

Proof (partial). Consider likelihood function \( L(\theta) \). \( \hat{\theta} \) satisfies \( \frac{dL(\theta)}{d\theta} = 0 \). Let \( \phi = h(\theta) \). Then the likelihood function for \( \phi = h(\theta) \) is given by \( L(h^{-1}(\phi)) \). Differentiating the likelihood with respect to \( \phi \), we have

\[
\frac{dL(h^{-1}(\phi))}{d\phi} = \frac{dL(\theta)}{d\theta} \cdot \frac{1}{h'(\theta)} = 0.
\]

2. Loglikelihood. Maximizing \( L(\theta) \) is equivalent to maximizing \( \ln L(\theta) \) since \( \ln \) is an increasing function.

Example. This technique is best illustrated by finding the MLE of parameters in \( N(\mu, \sigma^2) \).

\[
\hat{\mu} = \bar{X}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2
\]

are the MLE of \( \mu \) and \( \sigma^2 \) respectively. Note that \( \hat{\sigma}^2 \) is not un unbiased estimator of \( \sigma^2 \).

3. Asymptotic unbiasness. When the sample size is large, the maximum likelihood estimator of \( \theta \) is approximately unbiased. The MLE of \( \theta \) is approximately the MVUE of \( \theta \). This is why it is the most widely used estimation technique in statistics. For the previous example,

\[
E\hat{\sigma}^2 = E\frac{n-1}{n} S^2 = \frac{n-1}{n} \sigma^2 \rightarrow \sigma^2 \text{ as } n \rightarrow \infty
\]

4. If explicit density function is not available, you can not apply MLE. In this case apply the method of moment matching.

5. Additional problem (previous midterm). Let \( X_1, X_2 \) be a random sample from \( N(0, 1/\theta) \). Note that the sample size is 2 and the density function for \( X_i \) is

\[
f(x_i) = \frac{\sqrt{\theta}}{\sqrt{2\pi}} \exp\left(-\theta x_i^2/2\right).
\]

Find the likelihood function and use it to obtain the maximum likelihood estimator of \( \theta \).

Solution. The likelihood function is

\[
L(\theta) = \theta/(2\pi) \exp\left(-\theta(x_1^2 + x_2^2)/2\right).
\]

Now get log-likelihood function

\[
\ln L(\theta) = const + \ln \theta - \theta(x_1^2 + x_2^2)/2.
\]

Differentiate with respect to \( \theta \) we get

\[
\frac{d\ln L(\theta)}{d\theta} = \frac{1}{\theta} - \frac{1}{2}(x_1^2 + x_2^2) = 0.
\]

Solving the equation, we get

\[
\hat{\theta} = 2/(x_1^2 + x_2^2).
\]
