1. A sequence of random variables \( \{X_n\} \) is said to be bounded in probability if for every \( \varepsilon > 0 \), there exists a positive number \( M_\varepsilon \) such that \( P(|X_n| > M_\varepsilon) < \varepsilon \) for all \( n \).

(i) Show that if \( \{X_n\} \) converges in distribution, then \( \{X_n\} \) is bounded in probability.

Let \( F_n \) be the cdf of \( X_n \). Suppose that \( \{X_n\} \) converges in distribution to a cdf \( F \). This means that \( F_n(x) \to F(x) \) for every \( x \in \mathbb{R} \) such that \( F \) is continuous at \( x \). Now for any \( \varepsilon > 0 \) we can choose \( M_\varepsilon > 0 \) such that \( F(-M_\varepsilon) < \varepsilon \) and \( F(M_\varepsilon) > 1-\varepsilon \). Moreover, we can choose \( M_\varepsilon \) such that \( F \) is continuous at \( M_\varepsilon \) and \(-M_\varepsilon \), and hence \( F_n(-M_\varepsilon) \to F(-M_\varepsilon) \) and \( F_n(M_\varepsilon) \to F(M_\varepsilon) \). Consequently, there exists \( N_\varepsilon \) such that \( |F_n(-M_\varepsilon) - F(-M_\varepsilon)| < \varepsilon \) and \( |F_n(M_\varepsilon) - F(M_\varepsilon)| < \varepsilon \) for any \( n > N_\varepsilon \). It follows that \( |F_n(x)| < 2\varepsilon \) and \( F_n(x) \to 1 - 2\varepsilon \) for any \( n > N_\varepsilon \). It follows that \( P(|X_n| > M_\varepsilon) < 4\varepsilon \) for any \( n > N_\varepsilon \). This proves that \( \{X_n\} \) is bounded in probability (why?).

(ii) Show that if there exist positive constants \( r \) and \( C \) such that \( \mathbb{E}[|X_n|^r] \leq C \) for all \( n \), then \( \{X_n\} \) is bounded in probability.

For a given \( \varepsilon > 0 \) take \( M_\varepsilon = (C/\varepsilon)^{1/r} \). Then, by Chebyshev’s inequality, we have

\[
P(|X_n| > M_\varepsilon) = P(|X_n|^r > M_\varepsilon^r) \leq M_\varepsilon^{-r}\mathbb{E}[|X_n|^r] \leq M_\varepsilon^{-r}C = \varepsilon.
\]

(iii) Suppose that \( Y_n \overset{p}\to 0 \) and \( \{X_n\} \) is bounded in probability. Show that \( X_nY_n \overset{p}\to 0 \).

(This can be written as follows: if \( Y_n = o_p(1) \) and \( X_n = O_p(1) \), then \( X_nY_n = o_p(1) \).)

For \( \varepsilon > 0 \) and \( M > 0 \) we can write

\[
P(|X_nY_n| > \varepsilon) = P(|X_nY_n| > \varepsilon, |X_n| > M) + P(|X_nY_n| > \varepsilon, |X_n| \leq M)
\leq P(|X_n| > M) + P(|Y_n| > \varepsilon/M).
\]

Since \( \{X_n\} \) is bounded in probability, we can choose \( M \) large enough such that \( P(|X_n| > M) \) is arbitrary small for all \( n \). Also since \( Y_n \overset{p}\to 0 \), we have that \( P(|Y_n| > \varepsilon/M) \) tends to zero, and hence \( P(|Y_n| > \varepsilon/M) \) is arbitrary small for all \( n \) large enough. It follows that \( P(|X_nY_n| > \varepsilon) \) is arbitrary small for all \( n \) large enough, and hence \( X_nY_n \overset{p}\to 0 \).

2. Suppose that \( X_1, \ldots, X_n \) are iid random variables with pdf:

\[
f(x, \theta) = \begin{cases} 
\frac{1}{2\theta}, & \text{if } x \in [-\theta, \theta], \\
0, & \text{otherwise}, 
\end{cases}
\]

where \( \theta > 0 \) is the unknown parameter. Find the maximum likelihood estimator of \( \theta \) and show that it is consistent. Find a sufficient statistic for \( \theta \).
The likelihood function is \( L(\theta) = 1/(2\theta)^n \) if \( x_i \in [-\theta, \theta] \), \( i = 1, \ldots, n \), and \( L(\theta) = 0 \) otherwise. Consequently, the maximum of \( L(\theta) \) over \( \theta > 0 \) is attained at a larger of the numbers \(-x(1)\) and \( x(n)\), and hence the ML estimator

\[
\hat{\theta} = \max \{-X(1), X(n)\} = \max \{|X_1|, \ldots, |X_n|\}.
\]

The density function of \( X_1, \ldots, X_n \) is

\[
f(x_1, \ldots, x_n, \theta) = \frac{1}{(2\theta)^n} \prod_{i=1}^{n} I(|x_i| \leq \theta) = \frac{I(\max \{|X_1|, \ldots, |X_n|\} \leq \theta)}{(2\theta)^n}.
\]

By the Factorization criterion we then have that \( \max \{|X_1|, \ldots, |X_n|\} \) is a (minimal) sufficient statistic for \( \theta \).

3. Suppose that \((X_1, Y_1), \ldots, (X_n, Y_n)\) are iid pairs of random variables where \( X_i \) and \( Y_i \) are independent each having normal distribution \( N(\mu_i, \sigma^2) \). Find the ML estimators of \( \mu_1, \ldots, \mu_n \) and \( \sigma^2 \). Is the ML estimator of \( \sigma^2 \) consistent?

The log-likelihood function is

\[
\log L(\mu_1, \ldots, \mu_n, \sigma^2) = -n \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} [(x_i - \mu_i)^2 + (y_i - \mu_i)^2] + \text{constant}.
\]

The ML estimators \( \hat{\mu}_i = (X_i + Y_i)/2 \) and \( \hat{\sigma}^2 = \frac{1}{4n} \sum_{i=1}^{n} (X_i - Y_i)^2 \). We have that

\[
E \left[ \frac{1}{4n} \sum_{i=1}^{n} (X_i - Y_i)^2 \right] = \frac{1}{4n} \sum_{i=1}^{n} E[(X_i - Y_i)^2] = \sigma^2/2.
\]

It follows by the LLN that \( \hat{\sigma}^2 \) tends in probability to \( \sigma^2/2 \), as \( n \to \infty \), and hence is not consistent. Note that the situation here is not standard since the number of parameters tends to infinity as \( n \to \infty \).

4. Let \( X_1, \ldots, X_n \) be iid Bernoulli random variables with parameter \( \theta \in (0, 1) \), i.e., \( P(X_i = 1) = \theta \) and \( P(X_i = 0) = 1 - \theta \). (i) Show that \( X_1 + \ldots + X_n \) is a sufficient and complete statistic for \( \theta \). (ii) Show that \( I(X_1 = 1, X_2 = 0) \) is an unbiased estimator of \( \theta(1 - \theta) \).

We have that

\[
E[I(X_1 = 1, X_2 = 0)] = P(X_1 = 1, X_2 = 0) = P(X_1 = 1)P(X_2 = 0) = \theta(1 - \theta),
\]

and hence \( I(X_1 = 1, X_2 = 0) \) is an unbiased estimator of \( \theta(1 - \theta) \).

(iii) Find the UMVU estimator of \( \theta(1 - \theta) \).

By Lehmann-Scheffe theorem we need to calculate \( E[I(X_1 = 1, X_2 = 0)|T] \), where \( T = X_1 + \ldots + X_n \). We have, for \( t = 0, 1, \ldots, n \),

\[
E[I(X_1 = 1, X_2 = 0)|T = t] = \frac{P(X_1=1,X_2=0,\sum_{i=3}^{n} X_i=t)}{P(T=t)} = \frac{P(X_1=1)P(X_2=0)P(\sum_{i=3}^{n} X_i=t-1)}{P(T=t)} = \frac{\binom{n-2}{t-1}(1-\theta)^{t-1}(1-\theta)^{n-t-1}}{n(n-1)} \frac{t(n-t)}{n(n-1)}.
\]

Consequently the the UMVU estimator of \( \theta(1 - \theta) \) is \( \frac{T(n-T)}{n(n-1)} \).
5. Let $X_1, \ldots, X_n$ be iid random variables with pdf $f(x, \theta) = \theta x^{\theta-1}, 0 < x < 1, \theta > 0$. Is there a function of $\theta$, say $g(\theta)$, for which there exists an unbiased estimator whose variance attains the Cramér-Rao lower bound?

We know that there exists an unbiased estimator $W(X_1, \ldots, X_n)$ whose variance attains the Cramér-Rao lower bound iff

$$\prod_{i=1}^{n} f(x_i, \theta) = \exp\{A(\theta)W(x_1, \ldots, x_n) + B(\theta)\} h(x_1, \ldots, x_n). \quad (1)$$

Since

$$\log \prod_{i=1}^{n} f(x_i, \theta) = n \log \theta + (\theta - 1) \sum_{i=1}^{n} \log x_i,$$

it follows that (1) holds if

$$W(x_1, \ldots, x_n) = n^{-1} \sum_{i=1}^{n} \log x_i.$$

Now

$$\mathbb{E}[\log X_i] = \int_{0}^{1} (\log x) \theta x^{\theta-1} dx = \int_{0}^{1} (\log x) dx^\theta = (\log x) x^{\theta} \bigg|_{0}^{1} - \int_{0}^{1} x^\theta d \log x$$

$$= - \int_{0}^{1} x^{\theta-1} dx = -1/\theta.$$

It follows that $W(x_1, \ldots, x_n)$ is an unbiased estimator of $g(\theta) = -1/\theta$ for which the Cramér-Rao lower bound is attained.

Again by direct calculations $\mathbb{E}[(\log X_i)^2] = 2/\theta^2$, and hence $\text{Var}[\log X_i] = 1/\theta^2$. It follows that $\text{Var}[W] = 1/(n\theta^2)$. On the other hand the Cramér-Rao lower bound here is

$$\frac{[g'(\theta)]^2}{n \mathbb{E}[\log X_i]} = \frac{[1/\theta]^2}{n/\theta^2} = \frac{1}{n\theta^2}.$$

We obtain that, indeed, $\text{Var}[W]$ attains its Cramér-Rao lower bound.