We assume that the $x_i^r$ RVs come from a known density that has unknown parameter $\theta$. Consider a set of observations forming a vector $x = [x_1, \ldots, x_N, x_1, \ldots, x_N]$. Two cases: inference of unknown quantities.

Parameter Estimation

1. Quantity is fixed but unknown parameter estimation
2. Quantity is random and unknown-random variable estimator

Parameter Estimation
For i.i.d. samples

Given $N$ observations, find the ML estimate of $\eta$.

Example: Let $x \sim \mathcal{N}(\eta, \sigma^2)$.

The maximum likelihood estimate of $\theta$ is that which makes the $x$
Sample mean \( \equiv \frac{1}{N} \sum_{i=1}^{N} x_i = \mu \) \( \Leftrightarrow \)

\[
\frac{\partial \mathcal{L}}{\partial \mu} = \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} \mathcal{L}_i = \left( (\mu|x)^{\mu|x} f \right)_{\mu=\mu} \]

Taking the derivative

\[
\frac{\partial^2 \mathcal{L}}{\partial \mu^2} = \sum_{i=1}^{N} - \left( (\mu|x)^{\mu|x} f \right)_{\mu=\mu} \]

Thus, Choose \( \mu \) to maximize any monotonic function of \( (\mu|x)^{\mu|x} f \)

Maximizing \( (\mu|x)^{\mu|x} f \) is equivalent to maximizing any monotonic function of \( (\mu|x)^{\mu|x} f \), for a fixed set of observations, \( \mu \) is set to be the value of \( (\mu|x)^{\mu|x} f \).
Also a RV.

Since $\theta_{\text{ML}}$ is a function of the RV $x_1, x_2, \ldots, x_N$, we see that $\theta_{\text{ML}}$ is a function of $\theta$.

The likelihood function of $\theta$ is:

$$0 = \left[ (\theta | x)^{T} \xi_{f} \right] \theta_{\text{ML}} - \left[ \xi_{f} \right] \theta_{\theta}$$

or

$$0 = \left[ (\theta | x)^{T} \xi_{f} \right] \frac{\theta_{\text{ML}} \theta_{\theta}}{\xi_{f}}$$

The ML estimate of $\theta$ is obtained as the solution to

$$(\theta | x)^{T} \xi_{f} \theta_{\text{ML}} = (x)^{T} \xi_{f}$$

In general, the ML estimate of $\theta$ is $\theta_{\text{ML}}$. 

ELCE-636: Statistical Signal Processing
\[ \frac{L}{I} = \frac{\mathcal{L} \bigcap_{i=1}^{N} \frac{N}{I}}{I} = \text{ML} \varnothing \]

\[ 0 = \frac{\mathcal{L} \bigcap_i \varnothing}{N} = \]

\[ \left[ \frac{\mathcal{L} \bigcap_i \varnothing - (\varnothing) \cap \text{in } N}{\varnothing} \right \frac{\varnothing}{\varnothing} = \left[ (\mathcal{L} \bigcap_i \bigcap_i an) \bigcap_i \varnothing \right \frac{\varnothing}{\varnothing} \]

To find the ML estimate:

\[ \mathcal{L} \bigcap_i \varnothing = \prod_{i=1}^{N} \varnothing = (\mathcal{L} \bigcap_i \bigcap_i an) \bigcap_i \varnothing \]

arrival rate \( \varnothing \).

We measure \( N \mathcal{L} \bigcap_i \bigcap_i an \) and wish to estimate the \( N \mathcal{L} \bigcap_i \bigcap_i an \) and customer arrival times at the store. Ballon is a RV.

Example: The time between customer arrivals at the store. Ballon is a RV.
A consistent estimate, $\hat{\theta}$, is unbiased and efficient with respect to $\{N\theta\}$ for all $N$, then $\hat{\theta}$ is

\[
\text{var}(\hat{\theta}) \geq (\text{var}(\theta))_{\text{other}}
\]

3. $\hat{\theta}$ is efficient in comparison to other estimates if

\[
\text{for arbitrary } \varepsilon \quad I = \{ \text{ } \varepsilon > |\theta - \hat{\theta}| \} \quad \text{for } \text{all } N
\]

\[
\lim_{N \to \infty} \text{pr} \left( I \right) = 0
\]

2. $\hat{\theta}$ is consistent if

\[
\theta - \{ \text{bias} \} \mathcal{E} = \{ \{N\theta\}\mathcal{E} = \{N\theta\} = \theta
\]

1. An estimate $\hat{\theta}$ is unbiased if $\{N\theta\}$ is unbiased.

Since estimates are themselves RVs, we can state the following properties:

Properties of Estimates
\[
\{ \epsilon(r_l - \frac{1}{N} \sum_{l=1}^{N} \frac{N}{l}) \} \mathcal{E} = (N \epsilon r_l) \mathcal{E}
\]

Thus \( N \epsilon \) is unbiased.

\[
r_l = (N \epsilon r_N) \frac{N}{l} = \left\{ \epsilon x \sum_{N}^{1} \frac{N}{l} \right\} \mathcal{E} = \{N \epsilon l\} \mathcal{E}
\]

Example: Is the sample mean consistent?

\[
I = \{ \epsilon < |\theta - N \theta| \}_{\text{lim}} \mathcal{P}^{\leftarrow N} \quad \text{If var}(\epsilon) \quad \text{gives}
\]

\[
\frac{\epsilon^2}{(N \epsilon r_N) \text{var}} \geq \{ \epsilon < |\theta - N \theta| \}_{\text{lim}} \mathcal{P}^{+}
\]

The last result follows from the Chebyshev inequality.
\[
\left\{ \left[ \begin{array}{c} \theta \\ \mathbf{x} \end{array} \right] \right\} \in \mathcal{F} \leq (\theta) \text{var}
\]

Theorem: If \( \theta \) is an unbiased estimate of \( \theta \), then \( \text{var}(\theta) \leq \frac{\text{var}(\theta)}{\text{var}(\theta)} \).

Consider the variance of an estimate of \( \theta \).

**The Cramér-Rao Bound**

Consistent:

Thus, unbiased and \( \text{var}(\theta) \) is

\[
\frac{N}{\mathcal{F}} = \mathcal{O}
\]

\[
(\mathcal{O} \mathcal{N}) \frac{z}{\mathcal{I}} = \mathcal{O}
\]

\[
\left\{ \frac{z}{\mathcal{I}} \left( \mathcal{N} \mathcal{I} \right) \right\} \mathcal{F} \sum_{\mathcal{N}} \frac{z}{\mathcal{I}} = \mathcal{O}
\]

---

**ELEC-636: Statistical Signal Processing**
\[ 0 = \mathcal{X} p(\theta|X) \theta|X f(\theta - \theta) \int_{-\infty}^{\infty} \frac{\theta \Theta}{\Theta} \] 

Taking the derivative:

\[ 0 = \mathcal{X} p(\theta|X) \theta|X f(\theta - \theta) \int_{-\infty}^{\infty} = \{ \theta - \theta \} \mathcal{H} \] 

**Proof:** Since \( \theta \) is unbiased (minimum variance estimate), if any estimate satisfies the bound with equality, it is an efficient estimate.

\[ \mathcal{H} \] and \( \mathcal{H} \)

where it is assumed

\[ \left( \left\{ \left[ \left( \mathcal{H} \right) \mathcal{H} \right] \mathcal{H} \right\} \mathcal{H} \right) \leq \mathcal{H} \]

for equivalently

ELC-636: Statistical Signal Processing
\[ I = \left( x_p \left( (\theta - \theta)(\theta|x) \theta | x f \right) \right) \left( \left( \frac{\theta e}{[(\theta|x) \theta | x f] \cup \mathcal{E}} \right) \in - \int \right) \]

This can be equivalently expressed as:

\[ I = x_p(\theta - \theta)(\theta|x) \theta | x f \frac{\theta e}{[(\theta|x) \theta | x f] \cup \mathcal{E}} \in - \int \]

Using this in the above:

\[ \frac{\theta e}{(\theta|x) \theta | x f \mathcal{E}} = (\theta|x) \theta | x f \frac{\theta e}{[(\theta|x) \theta | x f] \cup \mathcal{E}} \]

Note the following equality:

\[ 0 = x_p(\theta - \theta) \frac{\theta e}{(\theta|x) \theta | x f \mathcal{E}} \in - \int + x_p(\theta|x) \theta | x f \in - \int - \]
\[ (\theta - \tilde{\theta}) \gamma = \begin{cases} \text{MIL criteria} \\
\text{In the above, let} \ \tilde{\theta} = \theta \\
(\theta - \tilde{\theta}) \gamma = (\theta | x) \frac{\theta \theta}{\theta} \left\{ \frac{\theta \theta}{\theta} \right\} \end{cases} \]

Thus

\[ \begin{cases} \text{var} \\
(\theta | x) \frac{\theta \theta}{\theta} \left\{ \frac{\theta \theta}{\theta} \right\} \end{cases} \]

We can now use the continuous version of Schwartz's inequality
is an efficient estimator that doesn't exist, then we don't know how good \( \theta_{\text{ML}} \). exists, then it is the ML estimate. Hence if an efficient estimate (one that satisfies the bound with equality)

\[ \theta_{\text{ML}} = \theta \]

Therefore