Theorem. Let X have one-parameter exponential family density

$$f(x|\theta) = \exp(c(\theta)T(x) + d(\theta) + S(x))I_A(x).$$

Consider the mean-value parametrization, $\delta(\theta) = E_\theta(T)$. Then

$$I(\delta) = \frac{1}{Var(T)}.$$

Poof. The natual parametrization, $\eta = c(\theta)$ gives

$$\begin{aligned} f^*(x|\eta) &= \exp(\eta T(x) + d_0(\eta) + S(x))I_A(x), \\ \log f^*(x|\eta) &= \eta T(x) + d_0(\eta) + S(x)), \\ \frac{\partial}{\partial \eta} \log f^*(x|\eta) &= T(x) + d_0'(\eta) = T(x) - E_\eta(T). \end{aligned}$$

Therefore,

$$I^*(\eta) = E_\eta \left(T - E_\eta(T)\right)^2 = Var_\eta(T).$$

Now, $\delta = E(T) = -d_0'(\eta) = h(\eta)$, so

$$\frac{d\eta}{d\delta} = \left(\frac{d\delta}{d\eta}\right)^{-1} = (-d_0''(\eta))^{-1} = \frac{1}{Var(T)}.$$

Therefore,

$$\begin{aligned} I(\delta) &= I^*(\eta)\left(\frac{d\eta}{d\delta}\right)^2 \Big|_{\eta=h^{-1}(\delta)} \\ &= Var(T)\frac{1}{(Var(T))^2}. \end{aligned}$$

Information Inequality (Cramer-Rao). Suppose the conditions (A) and (B) hold, and $0 < I(\theta) < \infty$. Let $T(X)$ be any statistic with $Var(T) < \infty$ and such that the derivative w.r.t. θ of

$$E_\theta(T) = \int T(x)f(x|\theta)\,dx$$

exists and can be obtained by differentiating under the integral sign. Then

$$Var_\theta(T(X)) \geq \frac{\left[\frac{d}{d\theta}E_\theta(T)\right]^2}{I(\theta)}.$$

Note. This is called the C-R lower bound on the variance of a statistic.

Proof. Note that

$$\begin{aligned}
\frac{d}{d\theta}E_\theta(T) &= \frac{d}{d\theta}\int_A T(x)f(x|\theta)\,dx = \int_A T(x)\frac{\partial}{\partial\theta}f(x|\theta)\,dx \\
&= \int_A T(x)\left[\frac{\frac{\partial}{\partial\theta}f(x|\theta)}{f(x|\theta)}\right]f(x|\theta)\,dx = \int T(x)S(x)f(x|\theta)\,dx \\
&= E_\theta\left(T(X)S(X)\right),
\end{aligned}$$

where $S(x) = \frac{\partial}{\partial\theta}\log f(x|\theta)$. Further,

$$\begin{aligned}
E_\theta S(X) &= \int_A S(x)f(x|\theta)\,dx = \int_A \frac{\partial}{\partial\theta}\log f(x|\theta)f(x|\theta)\,dx \\
&= \int_A \left(\frac{\frac{\partial}{\partial\theta}f(x|\theta)}{f(x|\theta)}\right)f(x|\theta)\,dx = \int \frac{\partial}{\partial\theta}f(x|\theta)\,dx \\
&= \frac{d}{d\theta}\int f(x|\theta)\,dx = 0.
\end{aligned}$$

Therefore,

$$\frac{d}{d\theta}E_\theta(T) = Cov_\theta(T(X), S(X)).$$

Since $|Cov(T(X), S(X))| \leq \sqrt{Var(T(X))Var(S(X))}$, and $Var_\theta(S(X)) = Var_\theta\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right) = I(\theta)$, we obtain

$$|\frac{d}{d\theta}E_\theta(T)| \leq \sqrt{Var_\theta(T)}\sqrt{I(\theta)}.$$

Therefore,

$$Var_\theta(T) \geq \frac{\left[\frac{d}{d\theta}E_\theta(T)\right]^2}{I(\theta)}.$$

Note. 1. For the class of all unbiased estimators of θ, we have $E_\theta(T) = \theta$ and $\frac{d}{d\theta}E_\theta(T) = 1$, so that

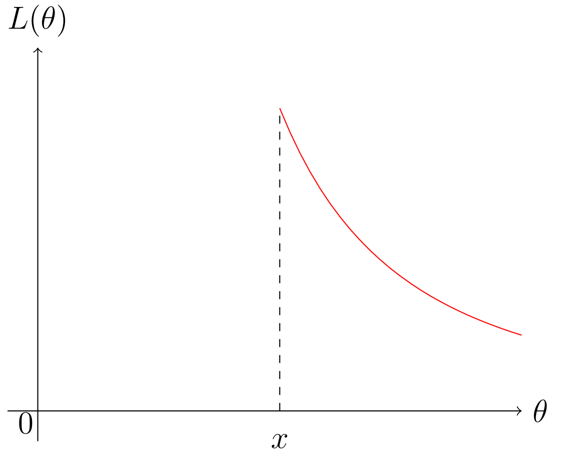
$$Var_\theta(T) \geq \frac{1}{I(\theta)}.$$

This lower bound is independent of any any particular T. Therefore, if there exists an unbiased estimator which attains this lower bound at all θ, it is the UMVUE.

Usually $T = T(X_1, \ldots, X_n) = T_n$, and $E_\theta(T_n) \to \theta$, which is called asymptotic unbiasedness. Then one would like to know what the asymptotic variance is, or whether it is the least it can be. MLE usually has this property.

Here is an example which shows that regularity conditions are indeed required.

Example. Let $X \sim U[0, \theta]$, $\theta > 0$. Then the likelihood function is as follows.



$$
\begin{aligned}
f(x|\theta) &= \frac{1}{\theta} \text{ for } x \leq \theta, \\
\log f(x|\theta) &= -\log\theta \text{ for } x \leq \theta, \\
\frac{\partial}{\partial\theta}\log f(x|\theta) &= \begin{cases} -\frac{1}{\theta} & \text{if } \theta > x; \\ \text{undefined} & \text{if } \theta = x; \end{cases}
\end{aligned}
$$

Since $P_\theta(X = \theta) = 0$,

$$\frac{\partial}{\partial\theta}\log f(X|\theta) = -\frac{1}{\theta}, \quad \text{w.p. 1 under } P_\theta.$$

Therefore,

$$
\begin{aligned}
E_\theta\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\right] &= -\frac{1}{\theta} \neq 0, \\
Var_\theta\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\right] &= 0, \\
E_\theta\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\right]^2 &= \frac{1}{\theta^2}.
\end{aligned}
$$

Then, what is $I(\theta)$? For unbiased estimators T of θ, do we have

$$Var_\theta(T) \geq \frac{1}{I(\theta)} = \theta^2 \text{ or } \infty?$$

Consider $T(X) = 2X$. Since $E(X) = \theta/2$, T is an unbiased estimator of θ. Note that $Var(T) = Var(2X) = 4Var(X) = 4\left(\frac{\theta^2}{12}\right) = \theta^2/3 < \theta^2 < \infty$. Note that conditions (A) and (B) are violated in this model.