Example. Let $X_1, X_2, \ldots$ be a random sample from a population with symmetric density, mean θ, variance equal to 1 and $f_X(\theta) > 0$. Consider $T_n^{(1)} = \bar{X}_n$ and $T_n^{(2)} = \text{median}(X_1, X_2, \ldots, X_n)$. From the above results,

$$\sqrt{n}\left(T_n^{(1)} - \theta\right) = \sqrt{n}\left(\bar{X}_n - \theta\right) \xrightarrow[n\to\infty]{d} N(0,1), \text{ and}$$

$$\sqrt{n}\left(T_n^{(2)} - \theta\right) = \sqrt{n}\left(\text{median}(X_1, X_2, \ldots, X_n) - \theta\right) \xrightarrow[n\to\infty]{d} N\left(0, \frac{1}{4f_0^2(0)}\right), \text{ so}$$

$$\mathrm{e}\left(\theta, T_n^{(1)}, T_n^{(2)}\right) = \frac{1}{4f_\theta^2(\theta)}.$$

Example. Suppose $X \sim N(\theta, 1)$ in the above example. Then $f_\theta(\theta) = \frac{1}{\sqrt{2\pi}} = f_0(0)$, so that $4f_\theta^2(\theta) = \frac{4}{2\pi} = \frac{2}{\pi}$. Therefore $\mathrm{e}\left(\theta, T_n^{(1)}, T_n^{(2)}\right) = \pi/2 \approx 1.57$. If we consider X such that $X - \theta \sim t_\nu$ for various values of ν, we obtain the following table listing $\mathrm{e}\left(\theta, T_n^{(2)}, T_n^{(1)}\right)$.

ν	3	4	5	8	∞
ARE	1.62	1.12	0.96	0.80	$\frac{2}{\pi} \approx 0.64$

Note that the sample median is more efficient than the sample mean when $\nu \leq 4$. In fact, the sample mean does not provide a consistent estimator when $\nu = 1$ since t_1 which is the same as Cauchy which does not have a mean. For $\nu = 2$ the variance is not finite.

Information bound and asymptotically efficient estimators

Suppose $T_n(X_1, \ldots, X_n)$ is an estimator of $q(\theta)$. Then the Information Inequality says

$$Var_\theta(T_n) \geq \frac{\left[\frac{\partial}{\partial\theta} E_\theta(T_n)\right]^2}{I_n(\theta)},$$

where $I_n(\theta) = nI_1(\theta)$ if X_i are i.i.d P_θ. Suppose further that T_n is asymptotically normal, in the sense,

$$\sqrt{n}\left(T_n - q(\theta)\right) \xrightarrow[n\to\infty]{d} N(0, \sigma^2(\theta)) \text{ and}$$

$$\sqrt{n}\left(E_\theta(T_n) - q(\theta)\right) \xrightarrow[n\to\infty]{} 0.$$

Then $E_\theta(T_n) = q(\theta) + o\left(\frac{1}{\sqrt{n}}\right)$ so that

$$\frac{\partial}{\partial\theta} E_\theta(T_n) = q'(\theta) + o\left(\frac{1}{\sqrt{n}}\right) \xrightarrow[n\to\infty]{} q'(\theta).$$

Then the Information bound for $\sigma^2(\theta)$ reduces to $(q'(\theta))^2/I_1(\theta)$.

Definition. $T_n = T_n(X_1, \ldots, X_n)$ is said to be asymptotically efficient for estimating $q(\theta)$ if its asymptotic variance is

$$\sigma^2(\theta) = \frac{(q'(\theta))^2}{I_1(\theta)}.$$

Example. Let $X_1, X_2, \ldots$ be i.i.d $N(\theta, 1)$. Then $I_1(\theta) = 1$. Consider $T_n^{(1)} = \bar{X}_n$. Then $\sqrt{n}\left(T_n^{(1)} - \theta\right) \xrightarrow[n\to\infty]{d} N(0,1)$. Note that

$$1 = \sigma^2(\theta) = \frac{1^2}{1} = \frac{1}{I_1(\theta)},$$

so that $T_n^{(1)} = \bar{X}_n$ is asymptotically efficient. Now consider
$T_n^{(2)} = \text{median}(X_1, X_2, \ldots, X_n)$. Then $\sqrt{n}\left(T_n^{(2)} - \theta\right) \xrightarrow[n\to\infty]{d} N(0, \frac{\pi}{2})$. Note that, now,

$$\sigma^2(\theta) = \frac{\pi}{2} > \frac{1}{I_1(\theta)} = 1.$$

Thus $T_n^{(2)}$ is not asymptotically efficient.

Result. Under regularity conditions on the model density, MLE is consistent. i.e., $\hat{\theta}(X_1, X_2, \ldots, X_n) \xrightarrow[n\to\infty]{P} \theta$.

Result. Under regularity conditions on the model density, MLE is asymptotically normal and asymptotically efficient. i.e.,

$$\sqrt{n}\left(\hat{\theta}(X_1, X_2, \ldots, X_n) - \theta\right) \xrightarrow[n\to\infty]{d} N\left(0, \frac{1}{I_1(\theta)}\right).$$

However, the conditions are typically hard to verify. It is easier to prove the result on a case by case basis using other standard limit theorems.

Result. Under regularity conditions on the model and the prior, the Bayes estimate, $\tilde{\theta} = E^{\pi}(\theta|\mathbf{X})$ is asymptotically normal and asymptotically efficient, satisfying

$$\sqrt{n}\left(\tilde{\theta}(X_1, X_2, \ldots, X_n) - \theta\right) \xrightarrow[n\to\infty]{d} N\left(0, \frac{1}{I_1(\theta)}\right).$$

Sketch of asymptotic normality for MLE assuming regularity conditions:

Let $X_1, X_2, \ldots$ be an i.i.d sequence from P_θ having density $f(x|\theta)$, $\theta \in \Theta$ which satisfies the regularity conditions to be specified in the proof below. We have

$$L(\theta, \mathbf{X}) = \prod_{i=1}^{n} f(X_i|\theta),$$

$$\mathcal{L}(\theta, \mathbf{X}) = \log L(\theta, \mathbf{X}) = \sum_{i=1}^{n} \log f(X_i|\theta),$$

Let $\hat{\theta}_n$ be the MLE of θ such that

$$\mathcal{L}'(\hat{\theta}_n(\mathbf{X}), \mathbf{X}) = 0.$$

Condition 1: $\mathcal{L}$ is differentiable three times and $\mathcal{L}'(\hat{\theta}_n) = 0$.

Now we obtain,

$$0 \;=\; \mathcal{L}'(\hat{\theta}_n) = \mathcal{L}'(\theta) + (\hat{\theta}_n - \theta)\mathcal{L}''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta)^2 \mathcal{L}'''(\theta_n^*),$$

where θ_n^* lies between $\hat{\theta}_n$ and θ. Therefore,

$$(\hat{\theta}_n - \theta)\left[\mathcal{L}''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta)\mathcal{L}'''(\theta_n^*)\right] = -\mathcal{L}'(\theta).$$

Hence,

$$\begin{aligned}
\sqrt{n}\left(\hat{\theta}_n - \theta\right) &= \frac{-\sqrt{n}\mathcal{L}'(\theta)}{\mathcal{L}''(\theta) + \frac{1}{2}(\hat{\theta}_n - \theta)\mathcal{L}'''(\theta_n^*)} \\
&= \frac{\frac{1}{\sqrt{n}}\mathcal{L}'(\theta)}{-\frac{1}{n}\mathcal{L}''(\theta) - \frac{1}{2n}(\hat{\theta}_n - \theta)\mathcal{L}'''(\theta_n^*)}.
\end{aligned}$$

Assuming that $\frac{1}{\sqrt{n}}\mathcal{L}'''(\theta_n^*)$ is bounded (Condition 2), and that $\hat{\theta}_n \xrightarrow[n\to\infty]{P} \theta$ (Condition 3, consistency of MLE), we obtain,

$$-\frac{1}{2n}(\hat{\theta}_n - \theta)\mathcal{L}'''(\theta_n^*) \xrightarrow[n\to\infty]{P} 0.$$

Note that, we can also get

$$\begin{aligned}
\frac{1}{\sqrt{n}}\mathcal{L}'(\theta) &= \sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(X_i|\theta), \text{ with} \\
E\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right) &= 0, \quad \text{(Condition 4)} \\
E\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right)^2 &= I_1(\theta). \quad \text{(Condition) 5}
\end{aligned}$$

Then it follows from CLT that

$$\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(X_i|\theta)-0\right]\xrightarrow[n\to\infty]{d}N(0,I_1(\theta)).$$

Therefore,

$$\frac{1}{\sqrt{n}}\mathcal{L}'(\theta)\xrightarrow[n\to\infty]{d}N(0,I_1(\theta)).$$

Note that

$$\begin{aligned}-\frac{1}{n}\mathcal{L}''(\theta) &= -\frac{1}{n}\frac{\partial}{\partial\theta}\mathcal{L}'(\theta)\\ &= -\frac{1}{n}\frac{\partial}{\partial\theta}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(X_i|\theta)=-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta^2}\log f(X_i|\theta).\end{aligned}$$

Now assuming (Condition 6) that

$$-E\left(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right)=I_1(\theta),$$

we have from WLLN,

$$-\frac{1}{n}\mathcal{L}''(\theta)=-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta^2}\log f(X_i|\theta)\xrightarrow[n\to\infty]{P}I_1(\theta).$$

Finally, we obtain

$$\sqrt{n}\left(\hat{\theta}_n-\theta\right)\xrightarrow[n\to\infty]{d}\frac{1}{I_1(\theta)}N(0,I_1(\theta))=N\left(0,\frac{1}{I_1(\theta)}\right).$$

Variance Stabilizing Transformations

Result. Suppose we have an i.i.d sequence $X_1, X_2, \ldots$ with $E(X_i)=\mu(\theta)$ and $Var(X_i)=\sigma^2(\theta)<\infty$. Then, from CLT, we have

$$\sqrt{n}\left(\bar{X}_n-\mu(\theta)\right)\xrightarrow[n\to\infty]{d}N\left(0,\sigma^2(\theta)\right).$$

Note, however, that s.e.$(\hat{\mu})=$ s.e.$(\bar{X}_n)=\sigma(\theta)/\sqrt{n}$ involves the usually unknown $\sigma^2(\theta)$. This may pose difficulties while using this for procedures such as large sample confidence intervals and tests. It is desirable to remove that dependence.

Recall that if h is differentiable and $h'(\mu) \neq 0$, then

$$\sqrt{n}\left(h(\bar{X}_n) - h(\mu)\right) \xrightarrow[n\to\infty]{d} N\left(0, (h'(\mu))^2\sigma^2\right).$$

If $(h'(\mu))^2\sigma^2 = (h'(\mu(\theta)))^2\sigma^2(\theta)$ can be made independent of θ for some h then the difficulty may be alleviated. For example, then, we could use

$$h(\bar{X}_n) \pm z_{1-\alpha/2}\frac{c}{\sqrt{n}}, \qquad c^2 = (h'(\mu))^2\sigma^2$$

as our confidence interval for $h(\mu)$. (Invert it to get a CI for μ.) In other words, $\frac{h(\bar{X}_n)-h(\mu)}{c/\sqrt{n}}$ becomes a pivot. *Variance stabilizing transformation* then is to find h such that $\sigma^2(\theta)(h'(\mu(\theta)))^2 \equiv c^2$.

Example. $S_n \sim \text{Binomial}(n, \theta)$. $\bar{X}_n = S_n/n$. $\mu(\theta) = \theta$.

$$\sqrt{n}\left(\bar{X}_n - \mu(\theta)\right) \xrightarrow[n\to\infty]{d} N\left(0, \sigma^2(\theta) = \theta(1-\theta)\right).$$

We require h such that

$$\begin{aligned}
\sigma^2(\theta)(h'(\mu(\theta)))^2 &\equiv c^2. \text{ i.e.,} \\
(h'(\theta))^2 &= \frac{c^2}{\theta(1-\theta)}, \text{ or} \\
h'(\theta) &= \frac{c}{\sqrt{\theta(1-\theta)}}.
\end{aligned}$$

Solving it, we get $h(\theta) \propto \sin^{-1}(\sqrt{\theta})$. Therefore the required transformation is $h(x) = \sin^{-1}(\sqrt{x})$. (Note, $\sin(h(x)) = \sqrt{x} \implies \cos(h(x))h'(x) = \frac{1}{2\sqrt{x}} \implies h'(x) = \frac{1}{2\sqrt{x}}\frac{1}{\cos(h(x)))} = \frac{1}{2\sqrt{x}}\frac{1}{\sqrt{1-(\sqrt{x})^2}}$.)

Example. $Y_1, Y_2, \ldots$ i.i.d Poisson(λ). Then $\mu(\lambda) = \lambda = \sigma^2(\lambda)$. We have $\sqrt{n}\left(\bar{Y}_n - \lambda\right) \xrightarrow[n\to\infty]{d} N(0, \lambda)$. Find h such that

$$\begin{aligned}
\sigma^2(\lambda)(h'(\lambda))^2 &\equiv c^2. \text{ i.e.,} \\
(h'(\lambda))^2 &= \frac{c^2}{\lambda}, \text{ or} \\
h'(\lambda) &= \frac{c}{\sqrt{\lambda}}.
\end{aligned}$$

Solving it, we get $h(\lambda) \propto \sqrt{\lambda}$. Therefore the required transformation is $h(x) = \sqrt{x}$.