Variance Stabilizing Transformations

Main Idea:
Suppose we wish to construct a CI for an unknown population parameter $\theta$ on the basis of a random sample $(X_1, \ldots, X_n)$, and $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ is an estimator for $\theta$. If $\text{Var}(\hat{\theta}_n)$ is a function of the unknown parameter $\theta$, the goal is to find a transformation $g$, such that $\text{Var}(g(\hat{\theta}_n))$ does not depend on $\theta$. Then one can often construct a CI for $g(\theta)$, and then convert it into a CI for $\theta$ itself.

Motivation:
Ex: Let $X_1, X_2, \ldots, X_n$ be iid $\text{Bernoulli}(\theta)$.
The goal is to construct a CI with the confidence level $1 - \alpha$ for an unknown parameter $\theta$.
The population mean is equal to $\theta$, the population variance $\sigma^2 = \theta(1 - \theta)$. A natural choice of $\hat{\theta}_n$ is $\hat{\theta}_n = \bar{X}_n$. Then

$$E(\hat{\theta}_n) = E(\bar{X}_n) = \theta, \quad \text{Var}(\hat{\theta}_n) = \text{Var}(\bar{X}_n) = \frac{\sigma^2}{n} = \frac{\theta(1 - \theta)}{n}.$$ 

By the CLT,

$$\sqrt{n}(\bar{X}_n - \theta) \xrightarrow{d} N(0, \theta(1 - \theta)), \quad \text{as } n \to \infty,$$

or, equivalently,

$$\frac{\sqrt{n}(\bar{X}_n - \theta)}{\sqrt{\theta(1 - \theta)}} \xrightarrow{d} N(0, 1), \quad \text{as } n \to \infty.$$ 

Thus, for a large sample size $n$,

$$P \left( -\frac{z_{\alpha/2}}{\sqrt{n(1 - \theta)}} \leq \frac{\sqrt{n}(\bar{X}_n - \theta)}{\sqrt{\theta(1 - \theta)}} \leq \frac{z_{\alpha/2}}{\sqrt{n(1 - \theta)}} \right) \approx 1 - \alpha,$$

where

$$P(Z > z_{\alpha/2}) = \frac{\alpha}{2} \quad \text{for } Z \sim N(0, 1).$$

In other words,

$$P \left( \theta \in \left[ \frac{\bar{X}_n - \frac{z_{\alpha/2}}{\sqrt{n}\sqrt{\theta(1 - \theta)}}}{\sqrt{n}}, \frac{\bar{X}_n + \frac{z_{\alpha/2}}{\sqrt{n}\sqrt{\theta(1 - \theta)}}}{\sqrt{n}} \right] \right) \approx 1 - \alpha,$$

but the above interval is not good for estimation of $\theta$ since the interval itself depends on the unknown parameter $\theta$.

Now suppose we can find a transformation $g(\bar{X}_n)$ such that $g$ is an invertible and at least twice continuously differentiable function, satisfying

$$\sqrt{n} \left( g(\bar{X}_n) - g(\theta) \right) \xrightarrow{d} N(0, \epsilon^2),$$

(1)
where $c$ is some known constant.
Then $g$ is called a "variance-stabilizing" transformation and the $(1 - \alpha)$ level CI for $g(\theta)$ is given by:
\[
\left[ g(\bar{X}_n) - z_{\alpha/2} \frac{c}{\sqrt{n}}, \ g(\bar{X}_n) + z_{\alpha/2} \frac{c}{\sqrt{n}} \right].
\]
And, therefore, whenever $g$ is a monotone nondecreasing function, the $(1 - \alpha)$ level CI for $\theta$ is given by:
\[
\left[ g^{-1}\left( g(\bar{X}_n) - z_{\alpha/2} \frac{c}{\sqrt{n}} \right), \ g^{-1}\left( g(\bar{X}_n) + z_{\alpha/2} \frac{c}{\sqrt{n}} \right) \right],
\]
and whenever $g$ is a monotone nonincreasing function, the $(1 - \alpha)$ level CI for $\theta$ is given by:
\[
\left[ g^{-1}\left( g(\bar{X}_n) + z_{\alpha/2} \frac{c}{\sqrt{n}} \right), \ g^{-1}\left( g(\bar{X}_n) - z_{\alpha/2} \frac{c}{\sqrt{n}} \right) \right].
\]

Our next step is to identify the variance-stabilizing transformation $g$.
In general, note that
\[
\sqrt{n} \left( g(\hat{\theta}_n) - g(\theta) \right) = \sqrt{n} \left( (\hat{\theta}_n - \theta)g'(\theta) + \frac{1}{2}g''(\theta^*) (\hat{\theta}_n - \theta)^2 \right),
\]
where $\theta^*$ is a certain point between $\hat{\theta}_n$ and $\theta$.
Then
\[
\sqrt{n} \left( \hat{\theta}_n - \theta \right) \overset{d}{\rightarrow} N(0, \sigma^2(\theta)) \tag{2}
\]
implies that for $g'(\theta) > 0$,
\[
\sqrt{n} \left( \hat{\theta}_n - \theta \right) g'(\theta) \overset{d}{\rightarrow} N(0, \sigma^2(\theta) \left[ g'(\theta) \right]^2), \text{ as } n \rightarrow \infty.
\]
Also if
\[
\hat{\theta}_n \overset{P}{\rightarrow} \theta, \tag{3}
\]
then
\[
\frac{1}{2}g''(\theta^*) \left[ \sqrt{n}(\hat{\theta}_n - \theta) \right] (\hat{\theta}_n - \theta) \overset{d}{\rightarrow} 0.
\]
Thus, under the conditions (2), (3),
\[
\sqrt{n} \left( g(\hat{\theta}_n) - g(\theta) \right) \overset{d}{\rightarrow} N(0, \sigma^2(\theta) \left[ g'(\theta) \right]^2),
\]
and $g$ is a variance-stabilizing transformation (and (1) is satisfied) if and only if
\[
\sigma^2(\theta) \left[ g'(\theta) \right]^2 = c^2.
\]
Without loss of generality, we can let $c = 1$, i.e.
\[
g'(\theta) = \frac{1}{\sigma(\theta)};
\]
or,
\[
g(\theta) = \int_{t \in \Theta: t \leq \theta} \frac{1}{\sigma(t)} dt. \tag{4}
\]

In our Bernoulli example (with \(\hat{\theta}_n = \overline{X}_n\)), by LLN and CLT, conditions (3) and (2) are satisfied. Now let \(g\) be a twice continuously differentiable invertible function satisfying condition (4), namely, (for \(\theta \in (0, 1)\))
\[
g(\theta) = \int_{t=0}^{\theta} \frac{1}{\sqrt{t(1-t)}} dt = \int_{0}^{\sin^{-1}(\sqrt{\theta})} \frac{2\sin(x) \cos(x)}{\sqrt{\sin^2(x) \cos^2(x)}} dx = 2\sin^{-1}(\sqrt{\theta}),
\]
where we put
\[
t = \sin^2(x), \ i.e. x = \arcsin(\sqrt{t}) = \sin^{-1}(\sqrt{t}).
\]
Then \(g(t) = 2\sin^{-1}(\sqrt{t})\) is a variance-stabilizing transformation and condition (1) holds with \(c = 1\), namely,
\[
\sqrt{n}\left(2\sin^{-1}\left(\sqrt{\overline{X}_n}\right) - 2\sin^{-1}\left(\sqrt{\theta}\right)\right) \xrightarrow{d} N(0, 1).
\]
Thus \((1 - \alpha)\) level CI for \(g(\theta) = 2\sin^{-1}(\sqrt{\theta})\) is given by:
\[
\left[2\sin^{-1}(\sqrt{\overline{X}_n}) - z_{a/2} \frac{1}{\sqrt{n}}, \ 2\sin^{-1}(\sqrt{\overline{X}_n}) + z_{a/2} \frac{1}{\sqrt{n}}\right].
\]
Note that \(g^{-1}(x) = \sin^2\left(\frac{x}{2}\right)\), thus a level \((1 - \alpha)\) CI for \(\theta\) is given by:
\[
\left[\sin^2\left(\sin^{-1}(\sqrt{\overline{X}_n}) - z_{a/2} \frac{1}{2\sqrt{n}}\right), \ \sin^2\left(\sin^{-1}(\sqrt{\overline{X}_n}) + z_{a/2} \frac{1}{2\sqrt{n}}\right)\right].
\]

Another example: \(X_1, \ldots, X_n\) iid Poisson(\(\theta\)), where parameter \(\theta > 0\) is unknown. Then the population mean equals to \(\theta\) and the population variance \(\sigma^2 = \theta\). By (4), let
\[
g(\theta) = \int_{0}^{\theta} \frac{1}{\sqrt{t}} dt = 2\sqrt{\theta},
\]
then \(g\) is a variance stabilizing transformation with \(c = 1\) and by (1),
\[
\sqrt{n}\left(2\sqrt{\overline{X}_n} - 2\sqrt{\theta}\right) \xrightarrow{d} N(0, 1).
\]
Thus, \((1 - \alpha)\) level CI for \(2\sqrt{\theta}\) is:
\[
\left[2\sqrt{\overline{X}_n} - z_{a/2} \frac{1}{\sqrt{n}}, \ 2\sqrt{\overline{X}_n} + z_{a/2} \frac{1}{\sqrt{n}}\right].
\]
And the \((1 - \alpha)\) level CI for \(\theta\) is given by:
\[
\left[\left(\sqrt{\overline{X}_n} - z_{a/2} \frac{1}{2\sqrt{n}}\right)^2, \ \left(\sqrt{\overline{X}_n} + z_{a/2} \frac{1}{2\sqrt{n}}\right)^2\right].
\]

3