In the percentile method, we’ve assumed that there exists a transformation of \( \theta, \phi(\theta) \), such that

\[
\phi(\hat{\theta}) - \phi(\theta) \sim N(0, 1)
\]

The transformation assumes that neither \( \theta \) nor \( \phi \) are biased, and it assumes that the variance is constant for all values of the parameter. That is, in the percentage intervals, we assume the normalizing transformation normalizes the sampling distribution (we no longer assume unbiased or constant variance).

Let’s consider an interesting probability question:

\[
\phi(\hat{\theta}) - \phi(\theta) \sim N(-z_0\sigma_{\phi}, \sigma_{\phi}) \quad \sigma_{\phi} = 1 + a\phi
\]

That is, there must exist a monotone transformation \( \phi \) such that \( \phi(\hat{\theta}) \sim N \) where

\[
E(\phi(\hat{\theta})) = \phi(\theta) - z_0[1 + a\phi(\theta)] \quad SE(\phi(\hat{\theta})) = 1 + a\phi(\theta)
\]

(Note: in the expected value and SE we’ve assumed that \( c = 1 \). If \( c \neq 1 \), then we can always choose a different transformation, \( \phi' \) so that \( c = 1 \).) Then

\[
P(z_{\alpha/2} \leq \frac{\phi(\hat{\theta}) - \phi(\theta)}{1 + a\phi(\theta)} + z_0 \leq z_{1-\alpha/2}) = 1 - \alpha
\]

A \((1 - \alpha)\)100\% CI for \( \phi(\theta) \) is

\[
\left[ \frac{\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0)}{1 + a(z_{1-\alpha/2} - z_0)}, \frac{\phi(\hat{\theta}) - (z_{\alpha/2} - z_0)}{1 + a(z_{\alpha/2} - z_0)} \right]
\]

Let’s consider an interesting probability question:

\[
P(\frac{\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0)}{1 + a(z_{1-\alpha/2} - z_0)}) = ?
\]

\[
= P(\frac{\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0) - \phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0)\phi(\hat{\theta})a(z_{1-\alpha/2} - z_0)}{(1 + a(z_{1-\alpha/2} - z_0))(1 + a\phi(\hat{\theta}))})
\]

\[
= P(\frac{\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0) - \phi(\hat{\theta})a(z_{1-\alpha/2} - z_0)}{(1 + a(z_{1-\alpha/2} - z_0))(1 + a\phi(\hat{\theta}))})
\]

\[
= P(\frac{\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0)}{1 + a\phi(\hat{\theta})}) \leq \frac{\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0)}{1 + a\phi(\hat{\theta})}
\]

\[
= P(\frac{\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0)}{1 + a\phi(\hat{\theta})}) \leq \frac{\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0)}{1 + a\phi(\hat{\theta})}
\]

\[
= P(\phi(\hat{\theta}) - (z_{1-\alpha/2} - z_0) + z_0 \leq \frac{(z_{\alpha/2} + z_0)}{(1 - a(z_{\alpha/2} + z_0))} + z_0)
\]

\[
= P(Z \leq \frac{(z_{\alpha/2} + z_0)}{(1 - a(z_{\alpha/2} + z_0))} + z_0) = \gamma_1
\]

where \( \gamma_1 = \Phi(\frac{(z_{\alpha/2} + z_0)}{(1 - a(z_{\alpha/2} + z_0))} + z_0)
\]

\[
= \text{pnorm}(\frac{(z_{\alpha/2} + z_0)}{(1 - a(z_{\alpha/2} + z_0))} + z_0)
\]
What we’ve shown is that the $\gamma_1$ quantile of the $\phi(\hat{\theta}^*)$ sampling distribution will be a good estimate for the lower bound of the confidence interval for $\phi(\theta)$. Using the same argument on the upper bound, we find a $(1 - \alpha)100\%$ confidence interval for $\phi(\theta)$ to be:

$$[\phi(\hat{\theta}^*)_{\gamma_1}, \phi(\hat{\theta}^*)_{\gamma_2}]$$

where

$$\gamma_1 = \Phi\left(\frac{z_{\alpha/2} + z_0}{(1 - a(z_{\alpha/2} + z_0)) + z_0}\right)$$

$$\gamma_2 = \Phi\left(\frac{z_{1-\alpha/2} + z_0}{(1 - a(z_{1-\alpha/2} + z_0)) + z_0}\right)$$

Using the transformation respecting property of percentile intervals, we know that a $(1 - \alpha)100\%$ confidence interval for $\theta$ is:

$$[\hat{\theta}^*_{\gamma_1}, \hat{\theta}^*_{\gamma_2}]$$

**How do we estimate $a$ and $z_0$?**

**bias**: $z_0$ is a measure of the bias. Recall:

$$bias = E(\hat{\theta}) - \theta$$

$$\hat{bias} = \hat{\theta}^*(\cdot) - \hat{\theta}$$

But remember that $z_0$ represents the bias for $\phi(\hat{\theta})$, not for $\hat{\theta}$ (and we don’t know $\phi$!). So, we use $\theta$ to see what proportion of $\theta$ values are too low, and we can map it back to the $\phi$ space using the normal distribution:

$$z_0 = \Phi^{-1}\left(\frac{\#\hat{\theta}^* < \hat{\theta}}{B}\right)$$

That is, if $\hat{\theta}^*$ underestimates $\hat{\theta}$, then $\hat{\theta}$ likely underestimates $\theta$; $z_0 > 0$. We think of $z_0$ and the normal quantile associated with the proportion of BS replicates less than $\hat{\theta}$.

**skew**: $a$ is a measure of skew.

$$bias = E(\hat{\theta} - \theta)$$

$$var = E(\hat{\theta} - \theta)^2 = \sigma^2$$

$$skew = E(\hat{\theta} - \theta)^3/\sigma^3$$

We can think of the skew as the rate of chance of the standard error on a normalized scale. If there is no skew, we will estimate $a = 0$. Our estimate of $a$ comes from a procedure known as the jackknife.

$$\hat{a} = \frac{\sum_{i=1}^{n}(\hat{\theta}(\cdot) - \hat{\theta}_{(i)})^3}{6[\sum_{i=1}^{n}(\hat{\theta}(\cdot) - \hat{\theta}_{(i)})^2]^{3/2}}$$