Confidence Interval

$$\hat{\theta} \pm Z_{\alpha/2}\, \hat{\sigma}$$

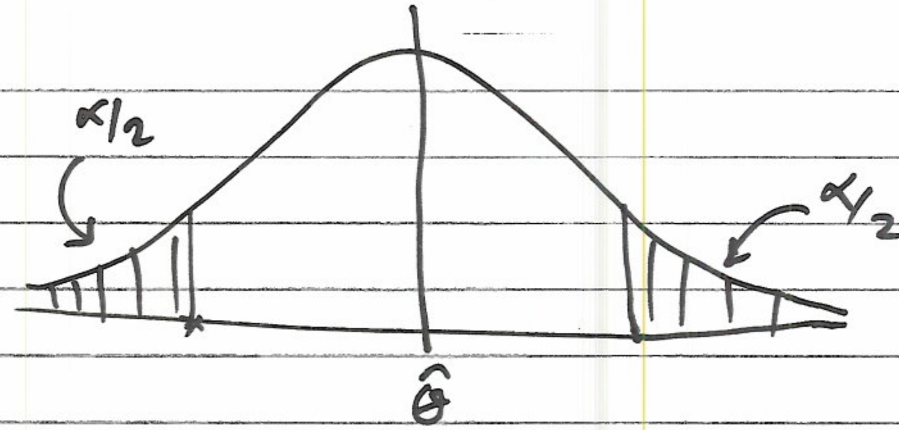
$\hat{\theta}$ is point estimate of parameter θ., $\hat{\sigma}$ is estimate of standard deviation of $\hat{\theta}$.

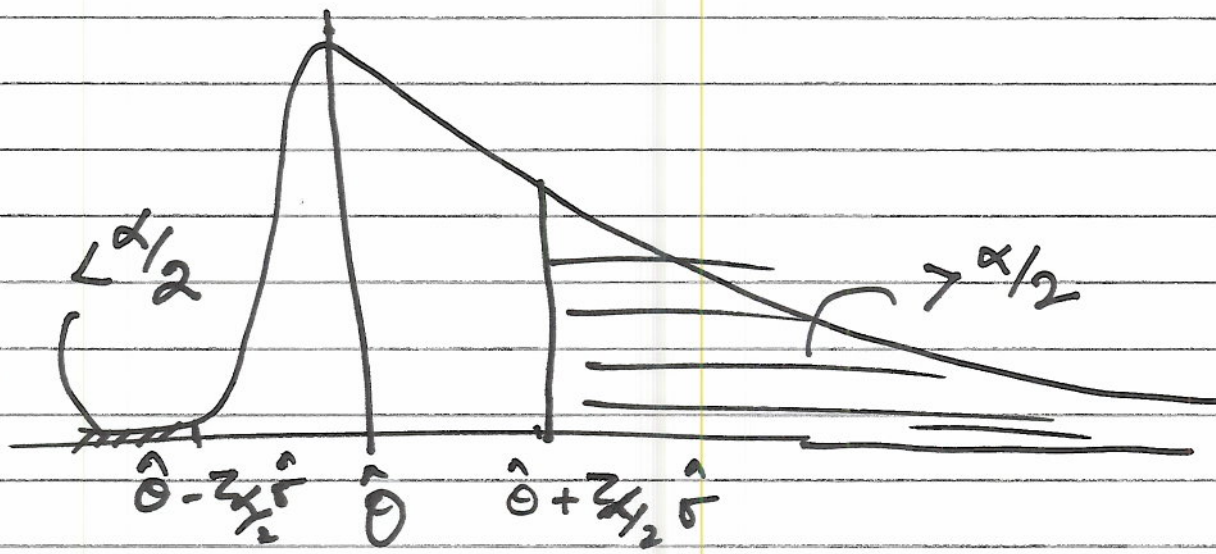
- Easy, automatic

- $\hat{\theta}$ and $\hat{\sigma}$ are typically obtained using maximum Likelihood value.

Problem:- Quite inaccurate in practice

- The coverage probability Can be substatially lower.





Problem with symmetric confidence Interval.

$$P(\theta > \hat{\theta} + z_{\alpha/2}\hat{\sigma}) > \alpha/2.$$ Liberal.

$$P(\theta < \hat{\theta} + z_{\alpha/2}\hat{\sigma}) < \alpha/2.$$ conservati

Using t-distribution (student t-correction)
is not satisfactory either

$$\hat{\theta} \pm t_\alpha \hat{\sigma}.$$

The percentile Method

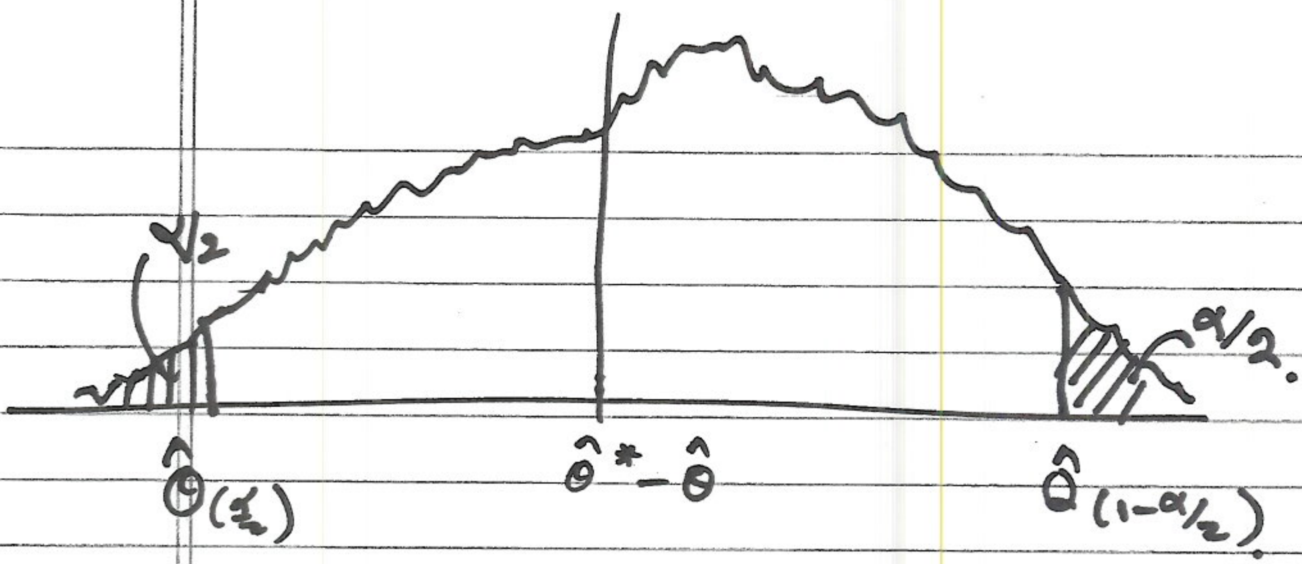
original data sample size N.

Take a random sample (of size N)
with replacement from original sample

$$x_1, \ldots \ldots, x_N$$

Bootstrap 1	x_1^*	x_3^*	x_1^*	x_7^*	$\ldots x_K^*$	$\hat{\theta}^*$
$\vdots$						
Bootstrap B	x_7^*	x_9^*	x_8^*	$\ldots$	x_9^*	$\hat{\theta}^*$

$$\hat{\theta}_i^*, \quad i=1, \cdots, B.$$



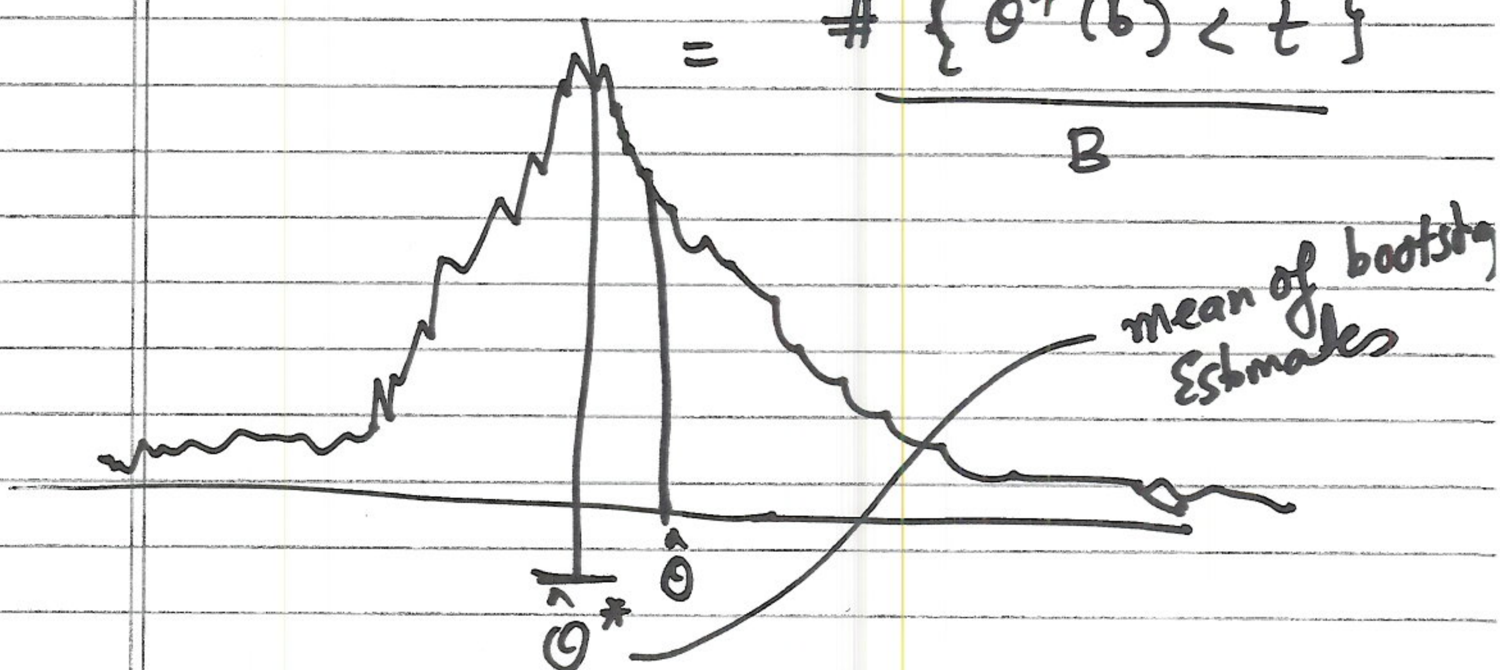
Bootstrap percentile Method C.I.

$$= \left[\hat{\theta}_{\frac{\alpha}{2}} , \hat{\theta}_{1-\frac{\alpha}{2}} \right]$$

Bootstrap probability

Let $\widehat{CDF}(t) = P_{*}(\hat{\theta}^{*} < t)$

$$= \frac{\# \{ \hat{\theta}^{*}(b) < t \}}{B}$$



Bias- Corrected percentile Method.

- Pivotal quantity or pivot is a function of observations and unobservable parameters whose probability distribution does not depend on the unknown parameters

Let $x_1, \dots, x_n$ be i.i.d. $\sim N(\mu, \sigma^2)$

$$g(\underset{\sim}{x}; \mu, \sigma^2) = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

g is called a pivotal quantity
(here it is normal pivotal quantity)

$$g(\underset{\sim}{x}; \mu, \sigma^2) = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$

Easy to construct

$$(X_i, Y_i) \sim \text{Bivariate Normal}\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\begin{pmatrix}\sigma_1^2 & \rho\\ & \sigma_2^2\end{pmatrix}\right)$$

parameter of interest is ρ.

$$r = \frac{\frac{1}{n-1}\sum(X_i-\bar{X})(Y_i-\bar{Y})}{S_X\,S_Y}$$

$$\sqrt{n}\,(r-\rho) \sim N(0, (1-\rho^2)^2)$$

Note: Variance changes with sample.

$$z = \tanh^{-1} r = \frac{1}{2}\ln\frac{1+r}{1-r}, \quad \psi = \tanh^{-1}\rho$$

Fisher's z transformation

$$\sqrt{n}\,(z-\psi) \sim N(0,1)$$

$$\mathrm{Var}(z) = 1/n$$

for small n, z is better approximation to normal distribution than r.

How to obtain such pivot?

Bias corrected percentile method is based on hypothesizing a transformation to a normal pivotal quantity.

θ parameter of interest

$\hat{\theta}$ Estimated parameter

$\hat{\theta}^*$ bootstrap sample based estimated θ.

Suppose there exist a transformation $g(\cdot)$ that is monotonic increasing such that

$$\phi = g(\theta) \quad , \quad \hat{\phi} = g(\hat{\theta}) \quad , \quad \hat{\phi}^* = g(\hat{\theta}^*)$$

Satisfy

$$\hat{\phi} - \phi \sim N(-z_0 \sigma, \sigma^2)$$

and $\hat{\phi}^* - \hat{\phi} \underset{*}{\sim} N(-z_0 \sigma, \sigma^2)$

$$\phi = \tanh^{-1} \rho$$

$$\sigma^2 = \frac{1}{n-3}, \quad z_0 = -\rho\sqrt{n-3} \big/ 2(n-1)$$

How does one obtains such transformation in practice?

- Bias-corrected percentile method requires <u>no</u> knowledge of the transformation $g(\cdot)$ or the constant z_0.

$$\boxed{\text{Recall if } \hat{\theta} \sim N(\theta, \sigma^2) \text{ then C.I. for } \theta = \hat{\theta} \pm Z_{\alpha/2}\sigma}$$

Since $\hat{\phi} - \phi \sim N(-z_0\sigma, \sigma^2)$

$\Leftrightarrow \quad \hat{\phi} + z_0\sigma \sim N(\phi, \sigma^2)$

$\Leftrightarrow \quad \boxed{\text{C.I. for } \phi = \hat{\phi} + z_0\sigma \pm Z_{\alpha/2}\sigma}$

$$\phi = g(\theta) \Leftrightarrow \theta = g^{-1}(\phi)$$

then CI for θ

~~C.I. for ϕ~~ $= \text{C.I. for } g^{-1}(\phi)$

$$= g^{-1}\left[\hat{\phi} + z_0\sigma \pm z_{\alpha/2}\sigma\right]$$

That is ~~[struck]~~ find C.I. for $g(\theta)$.

- No problem because its estimate is normally distributed. (symmetric)
- So, a symmetric C.I. around $g(\theta)$ is correct.
- Then back transform to θ-scale using g^{-1}

(However, note that we don't know what g is!).

Let $\widehat{CDG}$ be bootstrap cumulative distribution function of $\hat{\phi}^*$

and $\widehat{CDF}$ be bootstrap cumulative distribution function of $\hat{\theta}^*$

$$\widehat{CDG}(g(t)) = \text{Prob}_*\{\hat{\phi}^* \le g(t)\}$$
$$= \text{Prob}_*\{g^{-1}(\hat{\phi}^*) \le t\}$$
$$= \text{Prob}\{\hat{\theta}^* \le t\}$$
$$= \widehat{CDF}(t).$$

$$\boxed{\widehat{CDG}(g(t)) = \widehat{CDF}(t)}$$

$$\Leftrightarrow \boxed{g^{-1}\,\widehat{CDG}^{-1}(\cdot) = \widehat{CDF}^{-1}(\cdot)}$$

$$P_*\left\{\hat{\phi}^* \le \hat{\phi}\right\} = P\left\{\frac{\hat{\phi} - \hat{\phi} + z_0\sigma}{\sigma} \le z_0\right\}$$

$$= \Phi(z_0)$$

Recall $\hat{\phi}^* - \hat{\phi} \sim N(-z_0\sigma, \sigma^2\}$

$$P(\hat{\phi}^* \le \hat{\phi}) = P(\hat{\phi}^* - \phi \le 0)$$

$$= P\left((\hat{\phi} - \phi) + z_0\sigma \le z_0\sigma\right)$$

$$= P\left(\frac{(\hat{\phi} - \phi) + z_0\sigma}{\sigma} \le z_0\right)$$

$$= \Phi(z_0).$$

Also

$$P(\hat{\phi}^* \le \hat{\phi}) = \widehat{CDG}(\hat{\phi})$$

$$= \widehat{CDG}(g(\hat{\theta}))$$

$$= \widehat{CDF}(\hat{\theta})$$

$$\widehat{CDF}(\hat{\theta}) = \Phi(z_0)$$

$$z_0 = \Phi^{-1}(\widehat{CDF}(\hat{\theta}))$$

From page 9

C.I. for $\theta = g^{-1}\left\{ \hat{\phi} + z_0\sigma \pm z_{\alpha/2}\sigma \right\}$

$$\hat{CDG}\left(\hat{\phi} + z_0\sigma \pm z_{\alpha/2}\sigma \right)$$

$$= P_*\left(\hat{\phi}^* \le \hat{\phi} + z_0\sigma \pm z_{\alpha/2}\sigma \right)$$

$$= P_*\left(\hat{\phi}^* - \hat{\phi} \le z_0\sigma \pm z_{\alpha/2}\sigma \right)$$

$$= P_*\left(\hat{\phi}^* - \hat{\phi} + z_0\sigma \le 2z_0\sigma \pm z_{\alpha/2}\sigma \right)$$

$$= P_*\left(\frac{\hat{\phi}^* - \hat{\phi} + z_0\sigma}{\sigma} \le 2z_0 \pm z_{\frac{\alpha}{2}} \right)$$

$$= \Phi\left(2z_0 \pm z_{\frac{\alpha}{2}} \right)$$

$$\hat{CDG}\left(\hat{\phi} + z_0\sigma \pm z_{\alpha/2}\sigma \right) = \Phi\left(2z_0 \pm z_{\frac{\alpha}{2}} \right)$$

$$\Rightarrow \quad \hat{\phi} + z_0\sigma \pm z_{\alpha/2}\sigma = CDG^{-1}\left[\Phi\left(2z_0 \pm z_{\alpha/2} \right) \right]$$

C. I. for θ

$$= g^{-1}\left\{ \hat{\phi} + z_0 \sigma \pm z_{\alpha/2} \sigma \right\}$$

$$= g^{-1}\left\{ CDG^{-1}\left[\Phi(2 z_0 \pm z_{\alpha/2}) \right] \right\}$$

$$= CDF^{-1}\left\{ \Phi(2 z_0 \pm z_{\alpha/2}) \right\}$$