Lecture 24
Bootstrap confidence intervals

Confidence Intervals

θ: an unknown parameter of interest
We want to find limits $\underline{\theta}$ and $\overline{\theta}$ such that

$$P(\underline{\theta} \leq \theta) = P(\theta \leq \overline{\theta}) = 1 - \alpha$$

$G(t) = P(\sqrt{n}(\hat{\theta} - \theta) \leq t)$
If $G^{-1}(1-\alpha)$ is known, then

$$\underline{\theta} = \hat{\theta} - G^{-1}(1-\alpha)/\sqrt{n}$$

is an exact $100(1-\alpha)\%$ lower confidence limit for θ
Traditional asymptotic approach:

$$G(t) \to \Phi(t/\sigma), \quad n \to \infty$$

Φ: the standard normal cdf
σ: an unknown scale parameter
$\hat{\sigma}$: a consistent estimator of σ, $G(\hat{\sigma}t) \to \Phi(t)$
Normal approximation (NA) $100(1-\alpha)\%$ confidence limits are
$\underline{\theta}_N = \hat{\theta} - \hat{\sigma} z_{1-\alpha}/\sqrt{n}$, $\overline{\theta}_N = \hat{\theta} + \hat{\sigma} z_{1-\alpha}/\sqrt{n}$
$z_a = \Phi^{-1}(a)$

Bootstrap Confidence Intervals

(1) The hybrid bootstrap (HB)
A bootstrap estimator of $G(t) = P(\sqrt{n}(\hat{\theta} - \theta) \leq t)$ is

$$\hat{G}(t) = P_*(\sqrt{n}(\hat{\theta}^* - \hat{\theta}) \leq t)$$

$G^{-1}(1-\alpha)$ can be estimated by $\hat{G}^{-1}(1-\alpha)$
HB lower and upper confidence limits:
$\underline{\theta}_{HB} = \hat{\theta} - \hat{G}^{-1}(1-\alpha)/\sqrt{n}$
$\overline{\theta}_{HB} = \hat{\theta} - \hat{G}^{-1}(\alpha)/\sqrt{n}$
If $G(t)$ is nearly symmetric, then $-\hat{G}^{-1}(\alpha)$ can be replaced by $\hat{G}^{-1}(1-\alpha)$
Hybrid: Use bootstrap estimate in the construction of confidence interval based on the sampling distribution of $\hat{\theta} - \theta$.

(2) Bootstrap-t (BT)
$G(t) = P(\sqrt{n}(\hat{\theta} - \theta) \leq t) \to \Phi(t/\sigma)$
$\hat{\sigma}$: a consistent estimator of σ
Studentized statistic: $\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma}$

$$H(t) = P\left(\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma} \leq t\right) \to \Phi(t)$$

$\hat{\sigma}^*$: a bootstrap analog of $\hat{\sigma}$
A bootstrap estimator of $H(t)$:

$$\hat{H}(t) = P_*\left(\sqrt{n}(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}^* \leq t\right)$$

BT lower and upper confidence limits:
$\underline{\theta}_{BT} = \hat{\theta} - \hat{\sigma}\hat{H}^{-1}(1-\alpha)/\sqrt{n}$
$\overline{\theta}_{BT} = \hat{\theta} + \hat{\sigma}\hat{H}^{-1}(1-\alpha)/\sqrt{n}$
This is also a hybrid method
It is called BT because it is based on the studentized variable
$\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma}$

(3) The bootstrap percentile (BP)
Bootstrap distribution (histogram)

$$K(t) = P_*(\hat{\theta}^* \leq t) \approx \frac{1}{B}\left(\# \text{ of times } \hat{\theta}^{*b} \leq t\right)$$

$100(1-\alpha)\%$ BP lower confidence limit

$$\underline{\theta}_{BP} = K^{-1}(\alpha) = \inf\{t : K(t) \geq \alpha\}$$

$\underline{\theta}_{BP} \approx \alpha B$th ordered value of $\hat{\theta}^{*1}, ..., \hat{\theta}^{*B}$
$100(1-\alpha)\%$ BP upper confidence limit

$$\overline{\theta}_{BP} = K^{-1}(1-\alpha) = \inf\{t : K(t) \geq 1-\alpha\}$$

$\overline{\theta}_{BP} \approx (1-\alpha)B$th ordered value of $\hat{\theta}^{*1}, ..., \hat{\theta}^{*B}$

There are two improvements over the BP

Accuracy: Justification

Properties of the bootstrap percentile (BP)

Lower confidence bound (level $1 - \alpha$)

$$\underline{\theta}_{BP} = K^{-1}(\alpha), \quad K(t) = P_*(\hat{\theta}^* \leq t)$$

Assumption A:

There is a monotone transformation ϕ such that

$$P(\hat{\phi} - \phi \leq t) = \Psi(t) \quad \text{for all } t \text{ and all } P$$

$\phi = \phi(\theta)$, $\hat{\phi} = \phi(\hat{\theta})$

Ψ: a continuous, increasing cdf symmetric about 0

Result 1. If ϕ is known, then

$$\underline{\theta}_E = \phi^{-1}(\hat{\phi} + z_\alpha), \qquad z_\alpha = \Psi^{-1}(\alpha),$$

is an exact $100(1 - \alpha)\%$ lower confidence bound, i.e.,

$$P(\underline{\theta}_E \leq \theta) = 1 - \alpha$$

Proof

$$
\begin{aligned}
P(\underline{\theta}_E \leq \theta) &= P\left(\phi(\underline{\theta}_E) \leq \phi(\theta)\right) \\
&= P\left(\phi(\phi^{-1}(\hat{\phi} + z_\alpha)) \leq \phi\right) \\
&= P\left(\hat{\phi} + z_\alpha \leq \phi\right) \\
&= P\left(\hat{\phi} - \phi \leq -z_\alpha\right) \\
&= \Psi(-z_\alpha) \\
&= 1 - \Psi(z_\alpha) \\
&= 1 - \alpha
\end{aligned}
$$

Result 2. Under Assumption A,

$$\underline{\theta}_{BP} = \underline{\theta}_E.$$

- This means that the BP interval is exactly correct under Assumption A
- To use the BP, we don't need to know ϕ
- The derivation of ϕ is replaced by computation
- If Assumption A holds approximately (when n is large), then the BP intervals are approximately correct

Proof

Let $w = \phi(\underline{\theta}_{BP}) - \hat{\phi}$ and $\hat{\phi}^* = \phi(\hat{\theta}^*)$
Since Assumption A holds for all P (including P_*),

$$\begin{aligned}
\Psi(w) &= P_*(\hat{\phi}^* - \hat{\phi} \leq w) \\
&= P_*(\hat{\phi}^* - \hat{\phi} \leq \phi(\underline{\theta}_{BP}) - \hat{\phi}) \\
&= P_*(\phi(\hat{\theta}^*) \leq \phi(\underline{\theta}_{BP})) \\
&= P_*(\hat{\theta}^* \leq \underline{\theta}_{BP}) \\
&= K(\underline{\theta}_{BP}) \\
&= \alpha
\end{aligned}$$

By the uniqueness of the quantile,

$$w = \Psi^{-1}(\alpha) = z_\alpha$$

Then

$$\phi(\underline{\theta}_{BP}) = \hat{\phi} + z_\alpha$$

Hence

$$\underline{\theta}_{BP} = \phi^{-1}(\hat{\phi} + z_\alpha) = \underline{\theta}_E$$

The bootstrap bias-corrected percentile (BC)

Is Assumption A too strong?

Assumption B:

There is a monotone transformation ϕ and a constant z_0 such that

$$P(\hat{\phi} - \phi + z_0 \leq t) = \Psi(t) \quad \text{for all } t \text{ and all } P$$

Ψ: a continuous, increasing cdf symmetric about 0

If $z_0 = 0$, then Assumption B reduces to Assumption A.

z_0 is a bias correction.

Since Assumption B holds for all P (including P_*),

$$\begin{aligned} K(\hat{\theta}) &= P_*(\hat{\theta}^* \leq \hat{\theta}) \\ &= P_*(\hat{\phi}^* \leq \hat{\phi}) \\ &= P_*(\hat{\phi}^* - \hat{\phi} + z_0 \leq z_0) \\ &= \Psi(z_0) \end{aligned}$$

i.e., $z_0 = \Psi^{-1}(K(\hat{\theta}))$

Result 3. $\underline{\theta}_E = \phi^{-1}(\hat{\phi} + z_\alpha + z_0)$ is an exact $100(1-\alpha)\%$ lower confidence limit for θ under Assumption B
Proof: Similar to the proof of result 1.

We need to know ϕ and Ψ in order to use $\underline{\theta}_E$
The BC confidence limits are

$$
\begin{aligned}
\underline{\theta}_{BC} &= K^{-1}(\Psi(z_\alpha + 2z_0)) \\
&= K^{-1}(\Psi(z_\alpha + 2\Psi^{-1}(K(\hat{\theta})))) \\
\overline{\theta}_{BC} &= K^{-1}(\Psi(z_{1-\alpha} + 2\Psi^{-1}(K(\hat{\theta}))))
\end{aligned}
$$

When bootstrap Monte Carlo of size B is applied,
$\underline{\theta}_{BC} \approx B\Psi(z_\alpha + 2z_0))$th ordered value of $\hat{\theta}^{*1}, ..., \hat{\theta}^{*B}$
$\overline{\theta}_{BC} \approx B\Psi(z_{1-\alpha} + 2z_0))$th ordered value of $\hat{\theta}^{*1}, ..., \hat{\theta}^{*B}$
BC = BP if $K(\hat{\theta}) = 0.5$, since $\Psi^{-1}(0.5) = 0$
i.e., $\hat{\theta}$ is the median of $K(t)$ (bootstrap histogram)
Thus, BC is a bias-corrected BP
It shifts the BP bound towards left or right according to z_0
BC does not need the explicit form of ϕ
But BC requires the form of Ψ (typically, $\Psi = \Phi$)

Result 4. Under Assumption B,

$$\underline{\theta}_{BC} = \underline{\theta}_E$$

BC is exactly correct if Assumption B holds. If Assumption B holds approximately, then BC is approximately correct

Proof: For any α,

$$\begin{aligned} 1-\alpha &= \Psi(z_{1-\alpha}) \\ &= P_*(\hat{\phi}^* - \hat{\phi} + z_0 \leq z_{1-\alpha}) \\ &= P_*(\hat{\phi}^* \leq \hat{\phi} + \Psi^{-1}(1-\alpha) - z_0) \\ &= P_*(\hat{\theta}^* \leq \phi^{-1}(\hat{\phi} + \Psi^{-1}(1-\alpha) - z_0)) \\ &= K(\phi^{-1}(\hat{\phi} + \Psi^{-1}(1-\alpha) - z_0)) \end{aligned}$$

Hence, for all t,

$$K^{-1}(t) = \phi^{-1}(\hat{\phi} + \Psi^{-1}(t) - z_0)$$

Let $t = \Psi(z_\alpha + 2z_0)$,

$$\begin{aligned} K^{-1}(\Psi(z_\alpha + 2z_0)) &= \phi^{-1}(\hat{\phi} + z_\alpha + 2z_0 - z_0) \\ &= \phi^{-1}(\hat{\phi} + z_\alpha + z_0) \\ &= \underline{\theta}_E \end{aligned}$$

The bootstrap accelerated bias-corrected percentile (BC$_a$)

Assumption B is weaker than Assumption A

Assumption B is still too strong in some cases

Weaker assumptions?

Assumption C:

There is a monotone transformation ϕ and constants z_0 and a (acceleration constant) such that

$$P\left(\frac{\hat{\phi}-\phi}{1+a\phi}+z_0 \leq t\right)=\Psi(t) \quad \text{for all } t \text{ and all } P$$

Ψ: a continuous, increasing cdf symmetric about 0

If $a=0$, Assumption C reduces to Assumption B.

a corrects skewness

Result 5. Under Assumption C,

$$\underline{\theta}_E = \phi^{-1}\left(\hat{\phi} + \frac{(z_\alpha + z_0)(1 + a\hat{\phi})}{1 - a(z_\alpha + z_0)}\right)$$

is an exact $100(1-\alpha)\%$ lower confidence limit for θ

Proof:

$$\begin{aligned}
1 - \alpha &= \Psi(-z_\alpha) \\
&= P\left(\frac{\hat{\phi} - \phi}{1 + a\phi} + z_0 \leq -z_\alpha\right) \\
&= P\left(\hat{\phi} \leq \phi - (z_\alpha + z_0)(1 + a\phi)\right) \\
&= P\left(\frac{\hat{\phi} + z_\alpha + z_0}{1 - a(z_\alpha + z_0)} \leq \phi\right) \\
&= P\left(\hat{\phi} + \frac{(z_\alpha + z_0)(1 + a\hat{\phi})}{1 - a(z_\alpha + z_0)} \leq \phi\right) \\
&= P\left(\underline{\theta}_E \leq \theta\right)
\end{aligned}$$

We need to know a, Ψ, and ϕ in order to use $\underline{\theta}_E$
Under Assumption C, $z_0 = \Psi^{-1}(K(\hat{\theta}))$
First, assume a is known
The BC$_a$ lower confidence bound is

$$\underline{\theta}_{BC_a} = K^{-1}\left(\Psi\left(z_0 + \frac{z_\alpha + z_0}{1 - a(z_\alpha + z_0)}\right)\right)$$

The BC$_a$ upper confidence bound is

$$\overline{\theta}_{BC_a} = K^{-1}\left(\Psi\left(z_0 + \frac{z_{1-\alpha} + z_0}{1 - a(z_{1-\alpha} + z_0)}\right)\right)$$

When bootstrap Monte Carlo of size B is applied,
$\underline{\theta}_{BC_a} \approx B\Psi\left(z_0 + \frac{z_\alpha + z_0}{1-a(z_\alpha + z_0)}\right)$th ordered value of $\hat{\theta}^{*1}, ..., \hat{\theta}^{*B}$
$\overline{\theta}_{BC_a} \approx B\Psi\left(z_0 + \frac{z_{1-\alpha} + z_0}{1-a(z_{1-\alpha} + z_0)}\right)$th ordered value of $\hat{\theta}^{*1}, ..., \hat{\theta}^{*B}$
We don't need the explicit form of ϕ
But we need to know a and Ψ
If $a = 0$, BC$_a$ reduces to BC

Result 6. If a is known, then $\underline{\theta}_{BC_a} = \underline{\theta}_E$.

Hence, BC_a is exactly correct if Assumption C holds. If Assumption C holds approximately, then BC_a is approximately correct

Proof.

$$\begin{aligned}
1-\alpha &= \Psi(z_{1-\alpha}) \\
&= P_*\left(\frac{\hat{\phi}^* - \hat{\phi}}{1+a\hat{\phi}} + z_0 \le z_{1-\alpha}\right) \\
&= P_*\left(\hat{\phi}^* \le \hat{\phi} + (z_{1-\alpha} - z_0)(1+a\hat{\phi})\right) \\
&= P_*\left(\hat{\theta}^* \le \phi^{-1}(\hat{\phi} + (z_{1-\alpha} - z_0)(1+a\hat{\phi}))\right) \\
&= K\left(\phi^{-1}(\hat{\phi} + (z_{1-\alpha} - z_0)(1+a\hat{\phi}))\right) \\
&= K\left(\phi^{-1}(\hat{\phi} + (\Psi^{-1}(1-\alpha) - z_0)(1+a\hat{\phi}))\right)
\end{aligned}$$

This shows that

$$K^{-1}(t) = \phi^{-1}(\hat{\phi} + (\Psi^{-1}(t) - z_0)(1+a\hat{\phi}))$$

Let $t = \Psi\left(z_0 + \frac{z_\alpha + z_0}{1 - a(z_\alpha + z_0)}\right)$,

$$\begin{aligned} \underline{\theta}_E &= \phi^{-1}\left(\hat{\phi} + \frac{(z_\alpha + z_0)(1 + a\hat{\phi})}{1 - a(z_\alpha + z_0)}\right) \\ &= K^{-1}\left(\Psi\left(z_0 + \frac{z_\alpha + z_0}{1 - a(z_\alpha + z_0)}\right)\right) \\ &= \underline{\theta}_{BC_a} \end{aligned}$$

The constant a is unknown and has to be estimated
$\hat{a}$: an estimator of a
$\underline{\theta}_{BC_{\hat{a}}}$ and $\overline{\theta}_{BC_{\hat{a}}}$ are $\text{BC}_{\hat{a}}$ confidence limits
They are also called ABC confidence limits

Asymptotic Accuracy

A confidence set C is first order accurate if

$$P(\theta \in C) = 1 - \alpha + O(n^{-1/2})$$

and second order accurate if

$$P(\theta \in C) = 1 - \alpha + O(n^{-1})$$

For the case of $\hat{\theta}$ is a smooth function of sample means, we have shown the following summary:

1. The BT and bootstrap BC$_a$ one-sided confidence intervals are second order accurate.
2. The the BP, BC, HB, and NA one-sided confidence intervals are in general first order accurate.
3. The equal-tail two-sided confidence intervals produced by all five bootstrap methods and the normal approximation are second order accurate (errors cancel each other).