PH 712 Probability and Statistical Inference

Part X: Confidence Set/Interval

Zhiyang Zhou (zhou67@uwm.edu, zhiyanggeezhou.github.io)

2024/12/09 10:46:05

Confidence set (CB Sec 9.2.1 & 9.3.1)

- Called a confidence interval (CI) If the set is an interval
- True (but unknown) value of parameter θ, say θ_0
- $(1-\alpha)\times 100\%$ confidence set, say $C(X_1,\ldots,X_n)$: $C(X_1,\ldots,X_n)$ covers θ_0 with probability AT LEAST $(1-\alpha)\times 100\%$, i.e., $\Pr\{\theta_0 \in C(X_1,\ldots,X_n)\}$
  - $C(X_1,\ldots,X_n)$ is a set defined on sample $X_1,\ldots,X_n$ and hence is randomized, while θ_0 is fixed
  - $(1-\alpha)\times 100\%$ is called coverage probability

Construction of a confidence set by inverting a level α test

- (CB Thm 9.2.2) Implementation
  1. For each $\theta^* \in \Theta$, find the rejection region, say $R(\theta^*)$, of a level α test of hypotheses $H_0: \theta = \theta^*$ vs. $H_1: \theta \neq \theta^*$
  2. $C(x_1,\ldots,x_n) = \{\theta : (x_1,\ldots,x_n) \in \text{supp}(X_1,\ldots,X_n)/R(\theta)\}$,
     - $\text{supp}(X_1,\ldots,X_n)/R(\theta)$: the complementary set of $R(\theta)$.
- $(1-\alpha)\times 100\%$ confidence set $C(X_1,\ldots,X_n)$ does not cover θ_0 $\Leftrightarrow$ reject $H_0: \theta = \theta_0$ (vs. $H_1: \theta \neq \theta_0$) at level α
- Special cases:
  - $(1-\alpha)\times 100\%$ (asymptotic) LRT confidence set for θ: $\{\theta : -2(\ell(\theta) - \ell(\hat\theta_{\text{ML}}))\} < \chi^2_{1,1-\alpha}\}$
  - $(1-\alpha)\times 100\%$ Wald confidence set for θ: $\{\theta : |\hat\theta_{\text{ML}} - \theta|/\sqrt{\widehat{\text{var}}(\hat\theta_{\text{ML}})} < \Phi^{-1}_{1-\alpha/2}\}$
  - $(1-\alpha)\times 100\%$ score confidence set for θ: $\{\theta : |\ell'(\theta)|/\sqrt{I_n(\theta)} < \Phi^{-1}_{1-\alpha/2}\}$

CB Examples 10.4.2, 10.4.3 & 10.4.5

- $X_1,\ldots,X_n \overset{\text{iid}}{\sim} \text{Bernlli}(p)$, construct $(1-\alpha)\times 100\%$ confidence set for p.

Bootstrap method

- Implementation
  1. For b in $1:B$, do steps 2–3.
  2. Draw the bth resample $\boldsymbol{x}_b^*$ of size n from the empirical CDF (nonparametric bootstrap) OR a fitted parametric model (parametric bootstrap).
  3. Let $\hat\theta_b^* = \hat\theta(\boldsymbol{x}_b^*)$.

4. $(1-\alpha)$ bootstrap confidence interval for θ is $(q_{\alpha/2}, q_{1-\alpha/2})$, where $q_{\alpha/2}$ and $q_{1-\alpha/2}$ are $\alpha/2$ and $1-\alpha/2$ sample quantiles of $\{\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*\}$, respectively.

CB Examples 10.4.2, 10.4.3 & 10.4.5

- $X_1, \ldots, X_n \overset{\text{iid}}{\sim}$ Bernlli(p), construct $(1-\alpha)$ confidence set for p.

```
options(digits = 4)
set.seed(1)
B = 1e4L
n = 1e3L
alpha = .05
x = rbinom(n, 1, prob = .6)
theta_ml = mean(x)
theta_star_np = numeric(B)
theta_star_p = numeric(B)
# Nonparametric bootstrap
for (b in 1:B){
  x_star = sample(x, size = n, replace = T)
  theta_star_np[b] = mean(x_star)
}
quantile(theta_star_np, probs = c(alpha/2, 1-alpha/2))
# Parametric bootstrap
for (b in 1:B){
  x_star = rbinom(n, size = 1, prob = theta_ml)
  theta_star_p[b] = mean(x_star)
}
quantile(theta_star_p, probs = c(alpha/2, 1-alpha/2))
```