Chapter 7. Confidence Sets
Lecture 19: Pivotal quantities and confidence sets

Confidence sets

X: a sample from a population $P \in \mathscr{P}$.
$\theta = \theta(P)$: a functional from $\mathscr{P}$ to $\Theta \subset \mathscr{R}^k$ for a fixed integer k.
$C(X)$: a *confidence set* for θ, a set in $\mathscr{B}_\Theta$ (the class of Borel sets on Θ) depending only on X.
The confidence coefficient of $C(X)$:

$$\inf_{P \in \mathscr{P}} P(\theta \in C(X))$$

If the confidence coefficient of $C(X)$ is $\geq 1-\alpha$ for fixed $\alpha \in (0,1)$, then we say that $C(X)$ has confidence level $1-\alpha$ or $C(X)$ is a level $1-\alpha$ confidence set.

We focus on

- methods of constructing confidence sets;
- properties of confidence sets.

Construction of Confidence Sets: Pivotal quantities

The most popular method of constructing confidence sets is the use of pivotal quantities defined as follows.

Definition 7.1

A known Borel function $\Re$ of (X, θ) is called a *pivotal quantity* if and only if the distribution of $\Re(X, \theta)$ does not depend on P.

Remarks

- A pivotal quantity depends on P through $\theta = \theta(P)$.
- A pivotal quantity is usually not a statistic, although its distribution is known.
- With a pivotal quantity $\Re(X, \theta)$, a level $1 - \alpha$ confidence set for any given $\alpha \in (0, 1)$ can be obtained.
- If $\Re(X, \theta)$ has a continuous c.d.f., then we can obtain a confidence set $C(X)$ that has confidence coefficient $1 - \alpha$.

Construction

First, find two constants c_1 and c_2 such that

$$P(c_1 \leq \Re(X,\theta) \leq c_2) \geq 1-\alpha.$$

Next, define

$$C(X) = \{\theta \in \Theta : \ c_1 \leq \Re(X,\theta) \leq c_2\}.$$

Then $C(X)$ is a level $1-\alpha$ confidence set, since

$$\begin{aligned} \inf_{P\in\mathscr{P}} P\big(\theta \in C(X)\big) &= \inf_{P\in\mathscr{P}} P\big(c_1 \leq \Re(X,\theta) \leq c_2\big) \\ &= P\big(c_1 \leq \Re(X,\theta) \leq c_2\big) \\ &\geq 1-\alpha. \end{aligned}$$

The confidence coefficient of $C(X)$ may not be $1-\alpha$.

If $\Re(X,\theta)$ has a continuous c.d.f., then we can choose c_i's such that the equality in the last expression holds and the confidence set $C(X)$ has confidence coefficient $1-\alpha$.

In a given problem, there may not exist any pivotal quantity, or there may be many different pivotal quantities and one has to choose one based on some principles or criteria, which are discussed in §7.2.

Computation

When $\Re(X,\theta)$ and c_i's are chosen, we need to compute the confidence set $C(X) = \{c_1 \le \Re(X,\theta) \le c_2\}$.

This can be done by inverting $c_1 \le \Re(X,\theta) \le c_2$.

For example, if θ is real-valued and $\Re(X,\theta)$ is monotone in θ when X is fixed, then

$$C(X) = \{\theta : \underline{\theta}(X) \le \theta \le \overline{\theta}(X)\}$$

for some $\underline{\theta}(X) < \overline{\theta}(X)$, i.e., $C(X)$ is an interval (finite or infinite). If $\Re(X,\theta)$ is not monotone, then $C(X)$ may be a union of several intervals.

For real-valued θ, a confidence interval rather than a complex set such as a union of several intervals is generally preferred since it is simple and the result is easy to interpret.

When θ is multivariate, inverting $c_1 \le \Re(X,\theta) \le c_2$ may be complicated.

In most cases where explicit forms of $C(X)$ do not exist, $C(X)$ can still be obtained numerically.

Example 7.2

Let $X_1, ..., X_n$ be i.i.d. random variables from the uniform distribution $U(0, \theta)$.

Consider the problem of finding a confidence set for θ.

Note that the family $\mathscr{P}$ in this case is a scale family so that the results in Example 7.1 can be used.

But a better confidence interval can be obtained based on the sufficient and complete statistic $X_{(n)}$ for which $X_{(n)}/\theta$ is a pivotal quantity (Example 7.13).

Note that $X_{(n)}/\theta$ has the Lebesgue p.d.f. $nx^{n-1}I_{(0,1)}(x)$.

Hence c_i's should satisfy $c_2^n - c_1^n = 1 - \alpha$.

The resulting confidence interval for θ is

$$[c_2^{-1}X_{(n)}, c_1^{-1}X_{(n)}].$$

Choices of c_i's are discussed in Example 7.13.

Example 7.3 (Fieller's interval)

Let (X_{i1}, X_{i2}), $i = 1, ..., n$, be i.i.d. bivariate normal with unknown $\mu_j = E(X_{1j})$, $\sigma_j^2 = \mathrm{Var}(X_{1j})$, $j = 1, 2$, and $\sigma_{12} = \mathrm{Cov}(X_{11}, X_{12})$.
Let $\theta = \mu_2/\mu_1$ be the parameter of interest ($\mu_1 \neq 0$).
Define $Y_i(\theta) = X_{i2} - \theta X_{i1}$.
Then $Y_1(\theta), ..., Y_n(\theta)$ are i.i.d. from $N(0, \sigma_2^2 - 2\theta\sigma_{12} + \theta^2\sigma_1^2)$.
Let

$$S^2(\theta) = \frac{1}{n-1}\sum_{i=1}^{n}[Y_i(\theta) - \bar{Y}(\theta)]^2 = S_2^2 - 2\theta S_{12} + \theta^2 S_1^2,$$

where $\bar{Y}(\theta)$ is the average of $Y_i(\theta)$'s and S_i^2 and S_{12} are sample variances and covariance based on X_{ij}'s.
It follows from Examples 1.16 and 2.18 that $\sqrt{n}\bar{Y}(\theta)/S(\theta)$ has the t-distribution t_{n-1} and, therefore, is a pivotal quantity.
Let $t_{n-1,\alpha}$ be the $(1-\alpha)$th quantile of the t-distribution t_{n-1}.
Then

$$C(X) = \{\theta : n[\bar{Y}(\theta)]^2/S^2(\theta) \leq t_{n-1,\alpha/2}^2\}$$

is a confidence set for θ with confidence coefficient $1-\alpha$.

Note that $n[\bar{Y}(\theta)]^2 = t^2_{n-1,\alpha/2} S^2(\theta)$ defines a parabola in θ. Depending on the roots of the parabola, $C(X)$ can be a finite interval, the complement of a finite interval, or the whole real line (exercise).

Proposition 7.1 (Existence of pivotal quantities in parametric problems)

Let $T(X) = (T_1(X), ..., T_s(X))$ and $T_1, ..., T_s$ be independent statistics. Suppose that each T_i has a continuous c.d.f. $F_{T_i,\theta}$ indexed by θ. Then $\Re(X,\theta) = \prod_{i=1}^s F_{T_i,\theta}(T_i(X))$ is a pivotal quantity.

Proof

The result follows from the fact that $F_{T_i,\theta}(T_i)$'s are i.i.d. from the uniform distribution $U(0,1)$.

When θ and T in Proposition 7.1 are real-valued, we can use the following result to construct confidence intervals for θ even when the c.d.f. of T is not continuous.

Theorem 7.1

Suppose that P is in a parametric family indexed by a real-valued θ. Let $T(X)$ be a real-valued statistic with c.d.f. $F_{T,\theta}(t)$ and let α_1 and α_2 be fixed positive constants such that $\alpha_1 + \alpha_2 = \alpha < \frac{1}{2}$.
(i) Suppose that $F_{T,\theta}(t)$ and $F_{T,\theta}(t-)$ are nonincreasing in θ for each fixed t.
Define

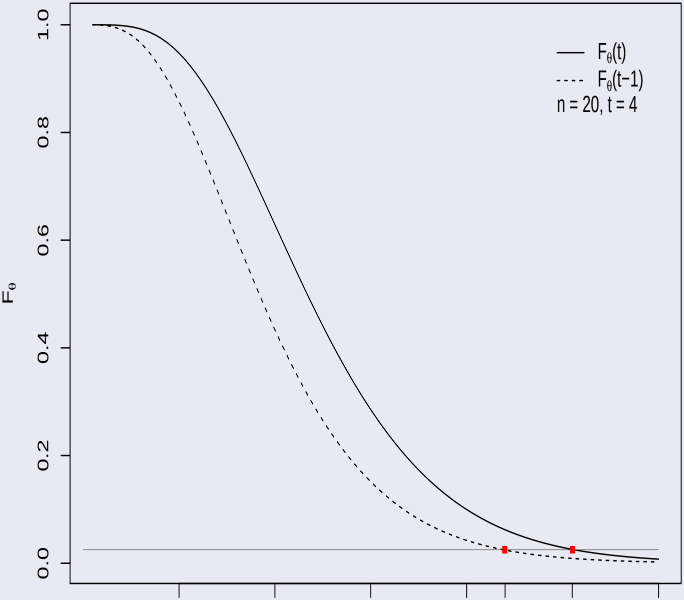
$$\overline{\theta} = \sup\{\theta : F_{T,\theta}(T) \geq \alpha_1\} \qquad \text{and} \qquad \underline{\theta} = \inf\{\theta : F_{T,\theta}(T-) \leq 1 - \alpha_2\}.$$

Then $[\underline{\theta}(T), \overline{\theta}(T)]$ is a level $1 - \alpha$ confidence interval for θ.
(ii) If $F_{T,\theta}(t)$ and $F_{T,\theta}(t-)$ are nondecreasing in θ for each t, then the same result holds with

$$\underline{\theta} = \inf\{\theta : F_{T,\theta}(T) \geq \alpha_1\} \qquad \text{and} \qquad \overline{\theta} = \sup\{\theta : F_{T,\theta}(T-) \leq 1 - \alpha_2\}.$$

(iii) If $F_{T,\theta}$ is a continuous c.d.f. for any θ, then $F_{T,\theta}(T)$ is a pivotal quantity and the confidence interval in (i) or (ii) has confidence coefficient $1 - \alpha$.

Figure: A confidence interval by using $F_{T,\theta}(t)$



Proof of Theorem 7.1.

We only need to prove (i).
Under the given condition, $\theta > \overline{\theta}$ implies $F_{T,\theta}(T) < \alpha_1$ and $\theta < \underline{\theta}$ implies $F_{T,\theta}(T-) > 1 - \alpha_2$.
Hence,

$$P\big(\underline{\theta} \le \theta \le \overline{\theta}\big) \ge 1 - P\big(F_{T,\theta}(T) < \alpha_1\big) - P\big(F_{T,\theta}(T-) > 1 - \alpha_2\big).$$

The result follows from

$$P\big(F_{T,\theta}(T) < \alpha_1\big) \le \alpha_1 \quad \text{and} \quad P\big(F_{T,\theta}(T-) > 1 - \alpha_2\big) \le \alpha_2.$$

The proof of this inequality is left as an exercise.

Discussion

When the parametric family in Theorem 7.1 has monotone likelihood ratio in $T(X)$, it follows from Lemma 6.3 that the condition in Theorem 7.1(i) holds; in fact, it follows from Exercise 2 in §6.6 that $F_{T,\theta}(t)$ is strictly decreasing for any t at which $0 < F_{T,\theta}(t) < 1$.

Discussion

If $F_{T,\theta}(t)$ is also continuous in θ, $\lim_{\theta\to\theta_-} F_{T,\theta}(t) > \alpha_1$, and $\lim_{\theta\to\theta_+} F_{T,\theta}(t) < \alpha_1$, where θ_- and θ_+ are the two ends of the parameter space, then $\overline{\theta}$ is the unique solution of $F_{T,\theta}(t) = \alpha_1$. (Figure)
A similar conclusion can be drawn for $\underline{\theta}$.
Theorem 7.1 can be applied to obtain the confidence interval for θ in Example 7.2 (exercise).
The following example concerns a discrete $F_{T,\theta}$.

Example 7.5

Let $X_1,...,X_n$ be i.i.d. random variables from the Poisson distribution $P(\theta)$ with an unknown $\theta > 0$ and $T(X) = \sum_{i=1}^n X_i$.
Note that T is sufficient and complete for θ and has the Poisson distribution $P(n\theta)$.
Thus

$$F_{T,\theta}(t) = \sum_{j=0}^{t} \frac{e^{-n\theta}(n\theta)^j}{j!}, \qquad t = 0,1,2,....$$

Since the Poisson family has monotone likelihood ratio in T and $0 < F_{T,\theta}(t) < 1$ for any t, $F_{T,\theta}(t)$ is strictly decreasing in θ.
Also, $F_{T,\theta}(t)$ is continuous in θ and $F_{T,\theta}(t)$ tends to 1 and 0 as θ tends to 0 and ∞, respectively.
Thus, Theorem 7.1 applies and $\overline{\theta}$ is the unique solution of $F_{T,\theta}(T) = \alpha_1$.
Since $F_{T,\theta}(t-) = F_{T,\theta}(t-1)$ for $t > 0$, $\underline{\theta}$ is the unique solution of $F_{T,\theta}(t-1) = 1-\alpha_2$ when $T = t > 0$ and $\underline{\theta} = 0$ when $T = 0$.

In fact, in this case explicit forms of $\underline{\theta}$ and $\overline{\theta}$ can be obtained from

$$\frac{1}{\Gamma(t)}\int_{\lambda}^{\infty} x^{t-1}e^{-x}dx = \sum_{j=0}^{t-1}\frac{e^{-\lambda}\lambda^j}{j!}, \qquad t = 1,2,....$$

Using this equality, it can be shown (exercise) that

$$\overline{\theta} = (2n)^{-1}\chi^2_{2(T+1),\alpha_1} \qquad \text{and} \qquad \underline{\theta} = (2n)^{-1}\chi^2_{2T,1-\alpha_2},$$

where $\chi^2_{r,\alpha}$ is the $(1-\alpha)$th quantile of the chi-square distribution χ^2_r and $\chi^2_{0,a}$ is defined to be 0.

So far we have considered examples for parametric problems.
In a nonparametric problem, a pivotal quantity may not exist and we have to consider approximate pivotal quantities (§7.3 and §7.4).

Asymptotic criterion

In some problems, especially in nonparametric problems, it is difficult to find a reasonable confidence set with a given confidence coefficient or confidence level $1-\alpha$.
A common approach is to find a confidence set whose confidence coefficient or confidence level is nearly $1-\alpha$ when the sample size n is large.
A confidence set $C(X)$ for θ has asymptotic confidence level $1-\alpha$ if

$$\liminf_n P(\theta \in C(X)) \geq 1-\alpha \quad \text{for any } P \in \mathscr{P}$$

If

$$\lim_{n\to\infty} P(\theta \in C(X)) = 1-\alpha \quad \text{for any } P \in \mathscr{P}$$

then $C(X)$ is a $1-\alpha$ *asymptotically correct* confidence set.
Note that asymptotic correctness is not the same as having limiting confidence coefficient $1-\alpha$ (Definition 2.14).

Asymptotically pivotal quantities

A known Borel function of (X, θ), $\Re_n(X, \theta)$, is said to be *asymptotically pivotal* iff the limiting distribution of $\Re_n(X, \theta)$ does not depend on P.
Like a pivotal quantity in constructing confidence sets (§7.1.1) with a given confidence coefficient or confidence level, an asymptotically pivotal quantity can be used in constructing asymptotically correct confidence sets.

Most asymptotically pivotal quantities are of the form $\widehat{V}_n^{-1/2}(\widehat{\theta}_n - \theta)$, where $\widehat{\theta}_n$ is an estimator of θ that is asymptotically normal, i.e.,

$$V_n^{-1/2}(\widehat{\theta}_n - \theta) \to_d N_k(0, I_k),$$

$\widehat{V}_n$ is a consistent estimator of the asymptotic covariance matrix V_n.
The resulting $1-\alpha$ asymptotically correct confidence sets are

$$C(X) = \{\theta : \|\widehat{V}_n^{-1/2}(\widehat{\theta}_n - \theta)\|^2 \le \chi^2_{k,\alpha}\},$$

where $\chi^2_{k,\alpha}$ is the $(1-\alpha)$th quantile of the chi-square distribution χ^2_k.
If θ is real-valued ($k = 1$), then $C(X)$ is a confidence interval.
When $k > 1$, $C(X)$ is an ellipsoid.

Example 7.20 (Functions of means)

Suppose that $X_1, ..., X_n$ are i.i.d. random vectors having a c.d.f. F on $\mathscr{R}^d$ and that the unknown parameter of interest is $\theta = g(\mu)$, where $\mu = E(X_1)$ and g is a known differentiable function from $\mathscr{R}^d$ to $\mathscr{R}^k$, $k \leq d$.

From the CLT, Theorem 1.12, $\widehat{\theta}_n = g(\bar{X})$ satisfies

$$V_n^{-1/2}(\widehat{\theta}_n - \theta) \to_d N_k(0, I_k)$$

$$V_n = [\nabla g(\mu)]^\tau \mathrm{Var}(X_1) \nabla g(\mu)/n$$

A consistent estimator of the asymptotic covariance matrix V_n is

$$\widehat{V}_n = [\nabla g(\bar{X})]^\tau S^2 \nabla g(\bar{X})/n.$$

Thus,

$$C(X) = \{\theta : \|\widehat{V}_n^{-1/2}(\widehat{\theta}_n - \theta)\|^2 \leq \chi^2_{k,\alpha}\},$$

is a $1-\alpha$ asymptotically correct confidence set for θ.

Example 7.22 (Linear models)

Consider linear model $X = Z\beta + \varepsilon$, where ε has i.i.d. components with mean 0 and variance σ^2.
Assume that Z is of full rank and that the conditions in Theorem 3.12 hold.
It follows from Theorem 1.9(iii) and Theorem 3.12 that for the LSE $\widehat{\beta}$,

$$V_n^{-1/2}(\widehat{\beta} - \beta) \to_d N_p(0, I_p)$$

$$V_n = \sigma^2 (Z^\tau Z)^{-1}$$

A consistent estimator for V_n is $\widehat{V}_n = (n-p)^{-1} SSR (Z^\tau Z)^{-1}$ (see §5.5.1).
Thus, a $1-\alpha$ asymptotically correct confidence set for β is

$$C(X) = \{\beta : (\widehat{\beta} - \beta)^\tau (Z^\tau Z)(\widehat{\beta} - \beta) \le \chi^2_{p,\alpha} SSR/(n-p)\}.$$

Note that this confidence set is different from the one in Example 7.9 derived under the normality assumption on ε.