Confidence Intervals (CI)

$(X_1, X_2, \ldots, X_n)$ i.i.d., θ parameter of interest.

Instead of estimating θ by a single number :
$\hookrightarrow$ better to provide interval estimates

$$\mathrm{P}\left(\theta \in \left[\widehat{m},\ \widehat{M}\right]\right) = 1 - \alpha, \quad \alpha \in [0, 1]$$

$\widehat{m} = \widehat{m}(X_1, \ldots, X_n)$, $\widehat{M} = \widehat{M}(X_1, \ldots, X_n)$: the interval is random

$\hookrightarrow$ Notation : $\mathrm{CI}_{1-\alpha}(\theta) = \left[\widehat{m},\ \widehat{M}\right]$
If $\alpha = 0.05$, 95% confidence interval. $1 - \alpha$: confidence level

Interpretation : If a large number of independent 95% intervals are constructed then approximately 95% of them will contain θ.

Gaussian model

$(X_1, X_2, \ldots, X_n) \sim \mathcal{N}(\mu, \sigma^2)$

$\hookrightarrow$ estimators of (μ, σ^2) : $\overline{X}$, $S^2 = \dfrac{\sum_i (X_i - \overline{X})^2}{n-1}$

Theorem : in the Gaussian model,

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1), \quad \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

$t(n-1)$: Student law with $(n-1)$ degrees of freedom (df).

$\hookrightarrow$ estimated Standard Error of $\overline{X}$:

$$\mathrm{SE}_{\overline{X}} = \frac{S}{\sqrt{n}}$$

$\hookrightarrow$ *t-statistic* :

$$\frac{\overline{X} - \mu}{\mathrm{SE}_{\overline{X}}} \sim t(n-1)$$

CI for the mean

$$\mathrm{CI}_{1-\alpha}(\mu) = \overline{X} \pm t_{1-\alpha/2,n-1} SE_{\overline{X}}$$

Why ?

Gaussian model : CI for the variance

$$\mathrm{P}\left(\chi^2_{\alpha/2,n-1} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{1-\alpha/2,n-1}\right) = 1-\alpha$$

$\chi^2_{\alpha,n-1}$ quantile of the $\chi^2(n-1)$ distribution.

$\hookrightarrow$

$$\mathrm{CI}_{1-\alpha}(\sigma^2) = \left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}},\ \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}\right]$$

where $S^2 = \dfrac{\sum_i (X_i - \overline{X})^2}{n-1}$.

CI : general case

$(X_1, X_2, \ldots, X_n) \sim F$ where F is an unknown distribution function.
θ parameter of interest, estimated by $\widehat{\theta}$.

To compute a confidence interval for θ we need

- a formula to calculate the standard error of $\widehat{\theta}$,
- information about the sampling distribution of $\widehat{\theta}$.

We have a formula for $\mathrm{SE}(\overline{X})$ and $\overline{X}$ is normally distributed by the central limit theorem (n reasonably large)

$\hookrightarrow$ what about the other estimators ?

- We have a simple method to compute the SE of Maximum Likelihood Estimators (MLE) !
- Moreover we know their approximate law !

Fisher information and SE

θ is one-dimensional, $\log L(\theta)$ is the log-likelihood of the sample, $\widehat{\theta}$ MLE

The *Fisher information* of the sample is defined as

$$I(\theta) = -\mathrm{E}\left[(\log L)''(\theta)\right].$$

The standard error of $\widehat{\theta}$ can be approximated by the inverse square root of the Fisher information $1/\sqrt{I(\theta)}$ and estimated by

$$\mathrm{SE}(\widehat{\theta}) = \frac{1}{\sqrt{\widehat{I}}}$$

where $\widehat{I}$ is a consistent estimator of $I(\theta)$.
For example $\widehat{I} = -(\log L)''(\widehat{\theta})$.

θ is a p-dimensional parameter vector.

The Fisher information is a $p \times p$ square matrix :

$$I(\theta) = -\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right].$$

The standard errors of the $\widehat{\theta}_j$ are the square roots of the diagonal entries of $I^{-1}(\theta$ and they are estimated by

$$\mathrm{SE}(\widehat{\theta}_j) = \frac{1}{\sqrt{\widehat{I}_{jj}}}.$$

- There is an explicit method of calcultating standard errors for MLEs
- The calculation of standard errors of MLEs is programmed in statistical software.

Approximate law of MLEs

Theorem

In "regular parametric models", the MLE of θ is approximately normally distributed

$$\widehat{\theta} \sim \mathcal{N}\left(\theta, I^{-1}(\theta)\right)$$

The approximation is still true if $I^{-1}(\theta)$ is estimated by $\widehat{I}^{-1}$ with $\widehat{I} = -H(\widehat{\theta})$, where H is the hessian of the log-likelihood.

Thus, if θ is a real parameter,

$$\mathrm{CI}_{1-\alpha}(\theta) = \widehat{\theta} \pm q_{1-\alpha/2}\mathrm{SE}_{\widehat{\theta}}$$

is an approximate $(1-\alpha)$ confidence interval for θ, with $\mathrm{SE}_{\widehat{\theta}} = \dfrac{1}{\sqrt{\widehat{I}}}$.

Example : MLE of the Pareto index

Assume that $X_1, \dots, X_n$ are i.i.d. Pareto(a, c).

$$CDF(x) = 1 - \left(\frac{c}{x}\right)^a, \quad x > c.$$

This distribution modelizes the income distribution and the parameter c is the minimum income, which is often known. The tail index a can be estimated by maximum likelihood.

- Find the *MLE* of a.
- Find a confidence interval for a.