2020-02-10

Estimators that produce a single value are point estimators

We turn now to interval estimators - where the estimated quantities are endpoints of an interval.

We would like to be able to pick out a region that will contain the true value of ϑ with high probability.

That is, we seek ϑ_l and ϑ_u st

$$P(\vartheta_l \leq \vartheta \leq \vartheta_u) = 1-\alpha \quad \text{for some small } \alpha.$$

If this is the case, we call $[\vartheta_l, \vartheta_u]$ a two-sided confidence interval with confidence coefficient $1-\alpha$.

It also makes sense to talk about one-sided confidence intervals: lower $[\vartheta_l, \infty)$ and upper $(-\infty, \vartheta_u]$.

In higher generality we could speak of a confidence region C to be any subset of the parameter space with $P(\vartheta \in C) = 1-\alpha$.

Disconnected confidence regions are beyond our scope.

The pivotal method

Suppose $f(\vartheta, X_1, \ldots, X_N) \sim \mathcal{D}$ is a function of ϑ and the data that has a known distribution. Suppose also that $\mathcal{D}$ does not depend on ϑ.

Then we can use $F_{\mathcal{D}}^{-1}(\alpha/2)$ and $F_{\mathcal{D}}^{-1}(1-\frac{\alpha}{2})$ to create a confidence interval!

We call the quantity $f(\vartheta, X_1, \ldots, X_N)$ a pivotal quantity.

Example

$Y \sim \text{Uniform}(0, \vartheta)$

Then $Y/\vartheta \sim \text{Uniform}(0,1)$ is pivotal.

A 90 % confidence interval can be created by using $F^{-1}(0.05) = 0.05$ and $F^{-1}(0.95) = 0.95$:

We want

$\mathbb{P}(\vartheta_L \leq \vartheta \leq \vartheta_u) = 0.9$

We know

$\mathbb{P}(0.05 \leq Y/\vartheta \leq 0.95) = 0.9$

Solve both inequalities for ϑ gives us

$\mathbb{P}(Y/0.95 \leq \vartheta \leq Y/0.05) = 0.9.$

Example $Y_1, \dots, Y_n \sim \text{Uniform}(0, \vartheta)$ iid.

$Y_{(n)} = \max(Y_i)$

$U = \frac{1}{\vartheta} Y_{(n)}$

It turns out that $F_U = \begin{cases} 0 & u < 0 \\ u^n & 0 \le u \le 1 \\ 1 & u > 1 \end{cases}$

So U is pivotal!

$\mathbb{P}(F_U^{-1}(0.05) \le U \le F_U^{-1}(0.95)) = 0.9$

Same algebra as last example gives us that

$\left[Y_{(n)} / \sqrt[n]{0.95}, \; Y_{(n)} / \sqrt[n]{0.05} \right]$ is a 90% confidence interval

Large-sample C.I.s

All of $\bar{X}$, $\bar{X}-\bar{Y}$, $\hat{p}$, $\hat{p}_1-\hat{p}_2$ have asymptotically normal sampling distributions.
Hence, for large enough samples,

$$\hat{\vartheta} \sim N(\vartheta, \sigma_{\hat{\vartheta}}) \quad \text{So} \quad \hat{\vartheta}-\vartheta \sim N(0, \sigma_{\hat{\vartheta}})$$

and so

$$Z = \frac{\hat{\vartheta}-\vartheta}{\sigma_{\hat{\vartheta}}} \sim N(0,1) \text{ is a pivotal quantity!}$$

We usually use this by picking

$$z_{\alpha/2} = F^{-1}_{N(0,1)}(\alpha/2) \qquad z_{1-\alpha/2} = F^{-1}_{N(0,1)}\left(1-\frac{\alpha}{2}\right)$$

Notice that since $N(0,1)$ is symmetric around 0,
$z_{\alpha/2} = -z_{1-\alpha/2}$. Call $z_{1-\alpha/2} =: z$. Then:

$$\mathbb{P}(-z \le Z \le z) = 1-\alpha$$

So

$$\mathbb{P}\left(-z \le \frac{\hat{\vartheta}-\vartheta}{\sigma_{\hat{\vartheta}}} \le z\right) = 1-\alpha$$

Solving for ϑ gets us $\vartheta \le \hat{\vartheta} + \sigma_{\hat{\vartheta}} \cdot z$
and $\hat{\vartheta} - \sigma_{\hat{\vartheta}} z \le \vartheta$

(The corresponding statistical test family is called Wald's test)

We should talk about what these probabilities mean

When we claim that

$$\vartheta \in \hat{\vartheta} \pm z_{\alpha/2} \cdot \sigma_{\hat{\vartheta}} \qquad \text{with } \mathbb{P} = 90\% \text{ say,}$$

we are saying that if we were to repeat the entire procedure many times, the true ϑ would be in the C.I. 90% of the repetitions.

Interpreting this probability is the source of a LOT of misused statistics and a core disagreement between frequentist and Bayesian statistics

---

Exercise Write out the C.I. for μ using $\bar{X}$ and this large sample construction.

Experiment design

Consider CIs for μ using $\bar{X}$.

If I have n and α and σ^2 (or approximate using s^2) then I can derive the precision of my confidence interval

$$w = v_u^g - v_\ell^g = \bar{X} + z_{\alpha/2} \cdot \sigma_{\bar{X}} - \bar{X} - z_{\alpha/2} \cdot \sigma_{\bar{X}} = 2 z_{\alpha/2} \cdot \sigma_{\bar{X}}$$

In fact, given any three of n, α, σ^2, w we can calculate the fourth. So we can pick precision in advance and derive required sample size to reach that precision.

n – expensive
α – certainty
σ^2 – noisiness
w – precision