## Uncertainty, confidence intervals and significant figures

EE-209 Eléments de Statistiques pour les Data Sciences

## Significant figures

In data science, continuous quantities are not known to arbitrary precision. On the contrary any measurement has a level of uncertainty. This means that in the decimal expansion of a number, digits beyond the precision, should be dropped and the number should be rounded to the last significant figure.

For example, if the uncertainty is  $0.000\underline{1}$  then  $x=0.0\underline{452}71298 \rightarrow x=0.0453$  For example, if the uncertainty is  $\underline{1}\cdot 10^2$  then  $x=\underbrace{5032}_{\text{significant digits}} \rightarrow x=5.0\cdot 10^3$ .

Note that in the examples above the uncertainty is specified with a **single** significant digit, which determines the level of precision.

So we removed digits beyond the precision level and rounded to the closest number at the same precision.

## Margin of error and reporting numbers at an appropriate precision level

In statistical estimation the level of uncertainty u is determined rigorously by confidence intervals: for symmetric CIs, u is the margin of error.

If a CI is of the form  $[\hat{\theta}-u,\hat{\theta}+u]$  then it is usually considered that

- u should be reported with k=1 or 2 (or at most 3) significant digits<sup>1</sup>.
- ② the precision of reported values for  $\hat{\theta}, \ \hat{\theta} u$  and  $\hat{\theta} + u$  should be the same as for u.
- lacktriangle the rounding should be done "outwards", i.e. down for  $\hat{\theta}+u$ , and up for  $\hat{\theta}+u$ .

For example, if  $\hat{\theta}=1.069871$  and u=0.002415,we can calculate  $\hat{\theta}-u=1.067456$  and  $\hat{\theta}+u=1.072286;$ if we keep 2 significant digits, then we report

$$u = 0.0024, \quad \hat{\theta} = 1.0699 \quad \text{and} \quad [\hat{\theta} - u, \hat{\theta} + u] = [1.0674, 1.0723].$$

If the first significant digit in u is a 1,2 or 3 we usually keep more significant digits than if it is a larger number, because the relative error produced by rounding is larger.

<sup>&</sup>lt;sup>1</sup>The Guide to the Expression of Uncertainty in Measurement (GUM) published by the Joint Committee for Guides in Metrology (JCGM) recommends to report uncertainty "to at most two significant digits" (7.2.6).