# Statistical methods and when to use them

## Systematic uncertainties

Experimental uncertainties that cannot be measured by repetition of the measurements are known as *systematic uncertainties*. These are either due an imperfection of the measurement device (say, a slow stopwatch) or an inability to measure more precisely. The first case can be corrected by standardizing or calibrating the instrument to known behaviors but most of the time will be smaller than the precision of the instrument.

For the latter case, it depends on the style of the instrument. If it is an *analog* device, then the uncertainty due to its precision is given as half the increment of the markings. If the device is *digital*, then the uncertainty is the smallest digit since we cannot see indications between the final digit.

A measurement is then reported as [units].

## Mean and associated uncertainty (standard deviation of the mean)

When taking many points *of the same thing*, we can find the *random* uncertainty associated to conducting the experiment. We can take the mean and associated uncertainty to summarize all the trials conducted. We expect that the uncertainty decreases with an increased number of repeated measurements. You would then report the measurement as [units].

;

## Propagation of uncertainty

Whenever we carry out a calculation of multiple variables with uncertainties, we need to make sure to propagate our uncertainty. If we have a function *f* that is made up of many variables (, etc.) with associated uncertainties , then the uncertainty in the function is given by

where is the derivative of the function *f* with respect to variable .

## Statistics vs Propagation of uncertainty: which to use?

The roles of statistical analysis and uncertainty propagations are complimentary. Most experiments should be able to be analyzed using both methods. If both are possible, you ought to do so both ways to check what they give. The expectation is that both give similar results, but in the cases that one is larger than the other, you would have to choose which one is a better description for your results.

## t-score

We can use the t-score to compare two values and if and only if we also know the associated uncertainties and , respectively.

The resulting t-score can then tell us if and are distinguishable or indistinguishable. See Table 1 for the exact scaling. One thing to note is that if your result is inconclusive, then we are obligated to take more data until you can make a conclusion. By definition, we cannot make a conclusion on something that is inconclusive.

**Table 1**: Results of the t-score and the conclusions that can be made. Also given is the appropriate follow up to the conclusion

|  |  |  |
| --- | --- | --- |
| **t-score** | **Conclusion** | **Follow up** |
| |t| < 1 | Indistinguishable | Collect more data or decrease uncertainty to determine distinguishability |
| 1 < |t| < 3 | Inconclusive | Collect more data or decrease uncertainty to determine (in)distinguishability |
| |t| > 3 | Distinguishable | Re-evaluate or revise model |

## Fitting data to extract a slope

If we have a set of data and we know that the two variables (and ) are related by a linear relationship, we can extract the proportionality constant, . The relationship can be written as:

.

To do this, we can take the “instantaneous” slope [] for each data point and then do a weighted average

The uncertainty associated with this weighted average is

## Determining “goodness” of fit

To determine whether data agrees with a fit, we will use a “goodness-of-fit” parameter, . To calculate , we need to know the uncertainty in our variable. The fitted function will give us the comparison between the data and the prediction. The “goodness-of-fit” parameter is calculated as:

where d is the number of parameters in our fitted function (or degrees of freedom). We can use this in the same way as the t-score such that the results give the conclusions in Table 2. Note that the value used for the comparison is , not .

**Table 2**: Results of and the conclusions that can be made. Also given is the appropriate follow up to the conclusion

|  |  |  |
| --- | --- | --- |
| **t-score** | **Conclusion** | **Follow up** |
| || < 1 | Indistinguishable | If you have multiple models indistinguishable from the data, you should choose the simplest one. Collect more data or decrease uncertainty to determine distinguishability. |
| 1 <||< 3 | Inconclusive | Collect more data or decrease uncertainty to determine (in)distinguishability. |
| |> 3 | Distinguishable | The fit does not agree with the data. Re-evaluate or revise mode. |

## Last and most definitely least, Percent difference

Whenever we need to compare two values and *but we are not able to get the uncertainty in at least one of the values*, then we can use percent difference. This should be the last resort. Percent difference is typically calculated as

The difficulty with percent difference is that we don’t know what an (un)acceptable value should be. We need to make sure to set a standard or contextualize our results for percent differences.