Applied Statistics

Measuring the length of a Table...





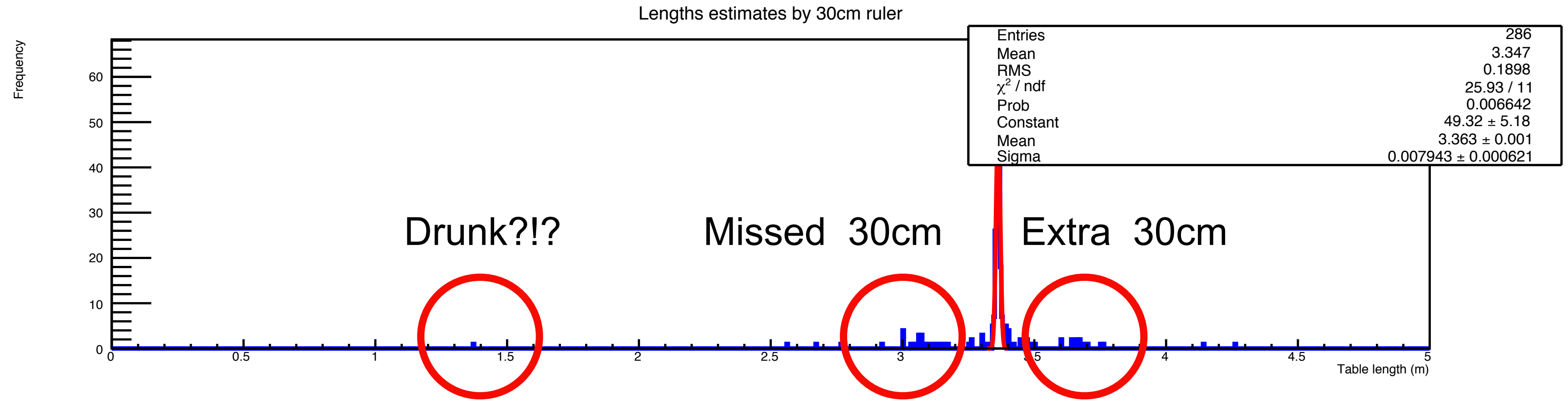

Troels C. Petersen (NBI)

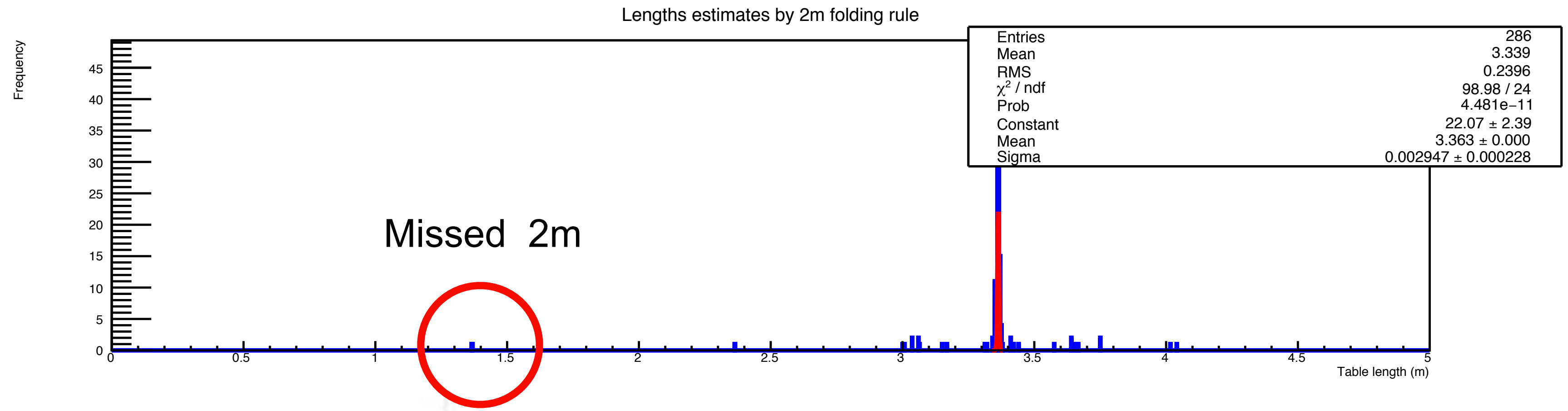
"Statistics is merely a quantisation of common sense"

The table measurement data

The initial dataset contains:

- 30cm measurements: **286** Range: [137.0, 450.0]
- 2m measurements: **284** Range: [137.0, 450.0]





Raw ("Naive") results

30cm:

Mean = 3.3455 ± 0.0112 m

RMS = 0.19 m (N = 286)

2m:

Mean = 3.3388 ± 0.0142 m

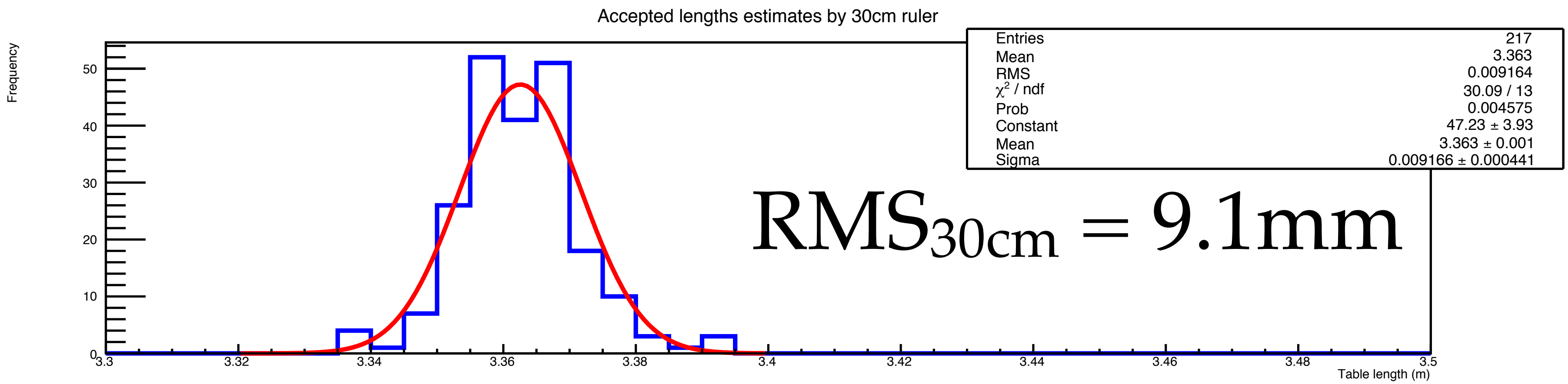
RMS = 0.24 m (N = 284)

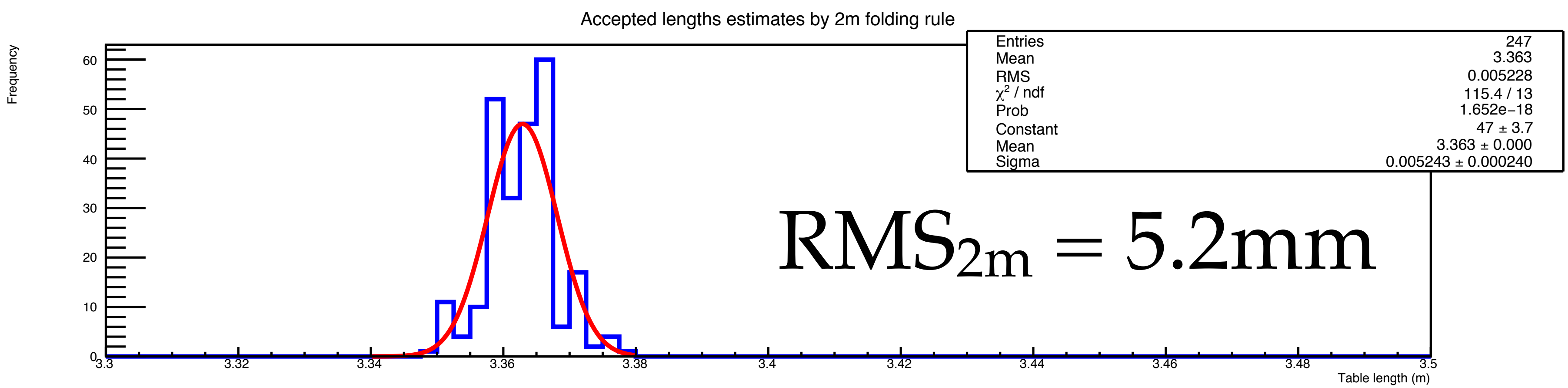
Correspondence between 30cm and 2m measurement:

Diff = 0.0067 ± 0.0181 m (0.37 σ)

Inspecting the data

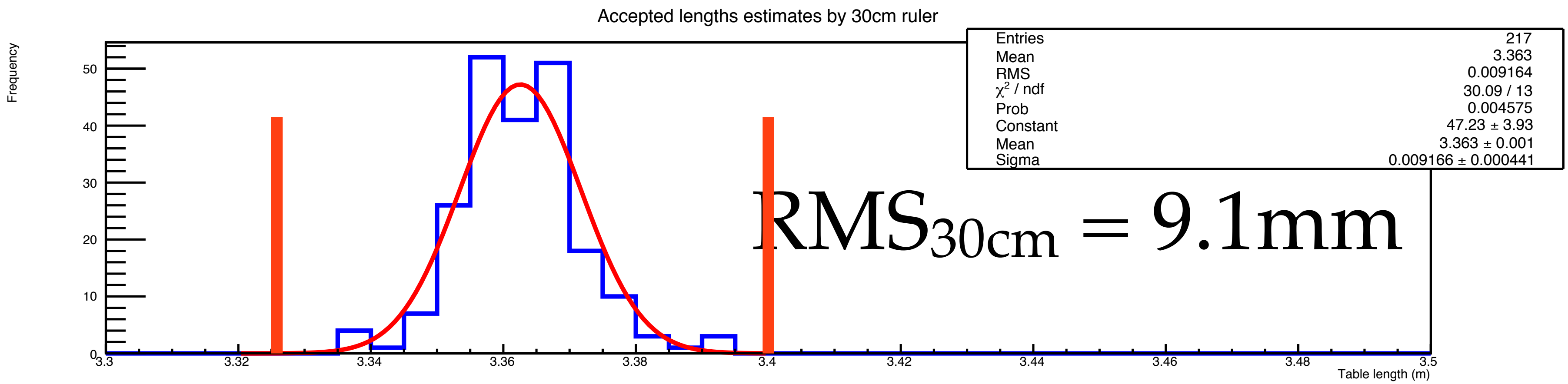
The 30cm peak seems Gaussian (p=5%) with binning 0.005 (smaller gives peaks).
The 2m peak does not seem Gaussian with any binning (here 0.0025), yet “collected”.

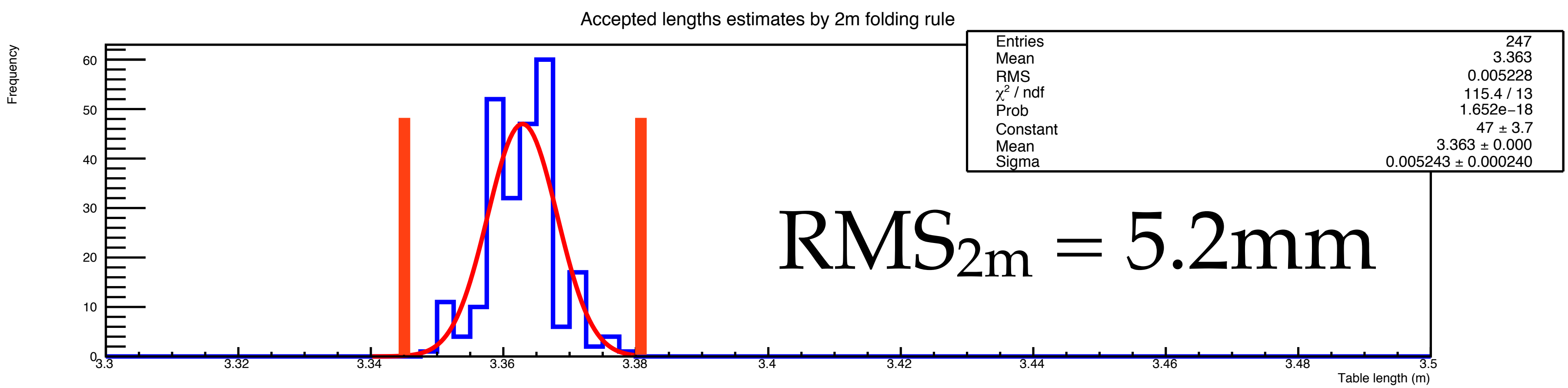




Inspecting the data

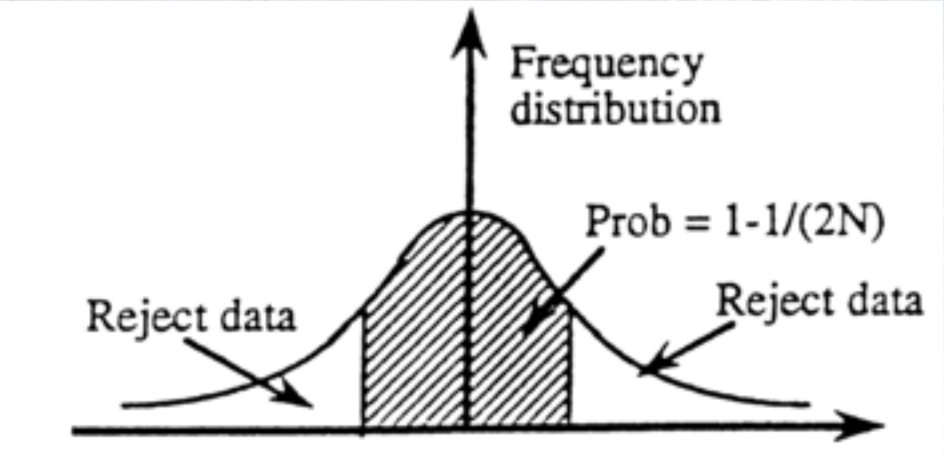
There are clearly some **mis-measurements**, which we would like to **exclude**. Using the measured RMS, and accepting that this only includes the best measurements, I **decide** to include measurements within **4 × RMS**:





Removing data points

Removing improbable data points is formalised in **Chauvenet's Criterion,** though many other methods exists (see Peirce, Grubbs, etc.)



The idea is to assume that the distribution is Gaussian, and ask what the probability of the farthest point is. If it is below some value, which is to be determined ahead of applying the criterion, then the point is removed, and the criterion is reapplied until no more points should be removed.

I choose to say, that if the outermost point in the Gaussian case has less than 10% chance of being this far out (taking the total number of points into account), then I reject it.

However, **ALWAYS keep a record of your original data**, as it may contain more effects than you originally thought.

Unweighted results

<u>30cm:</u>

Mean = 3.36227 ± 0.00061 m

RMS = 0.009 m (N = 217)

<u>2m:</u>

Mean = 3.36273 ± 0.00032 m

RMS = 0.005 m (N = 247)

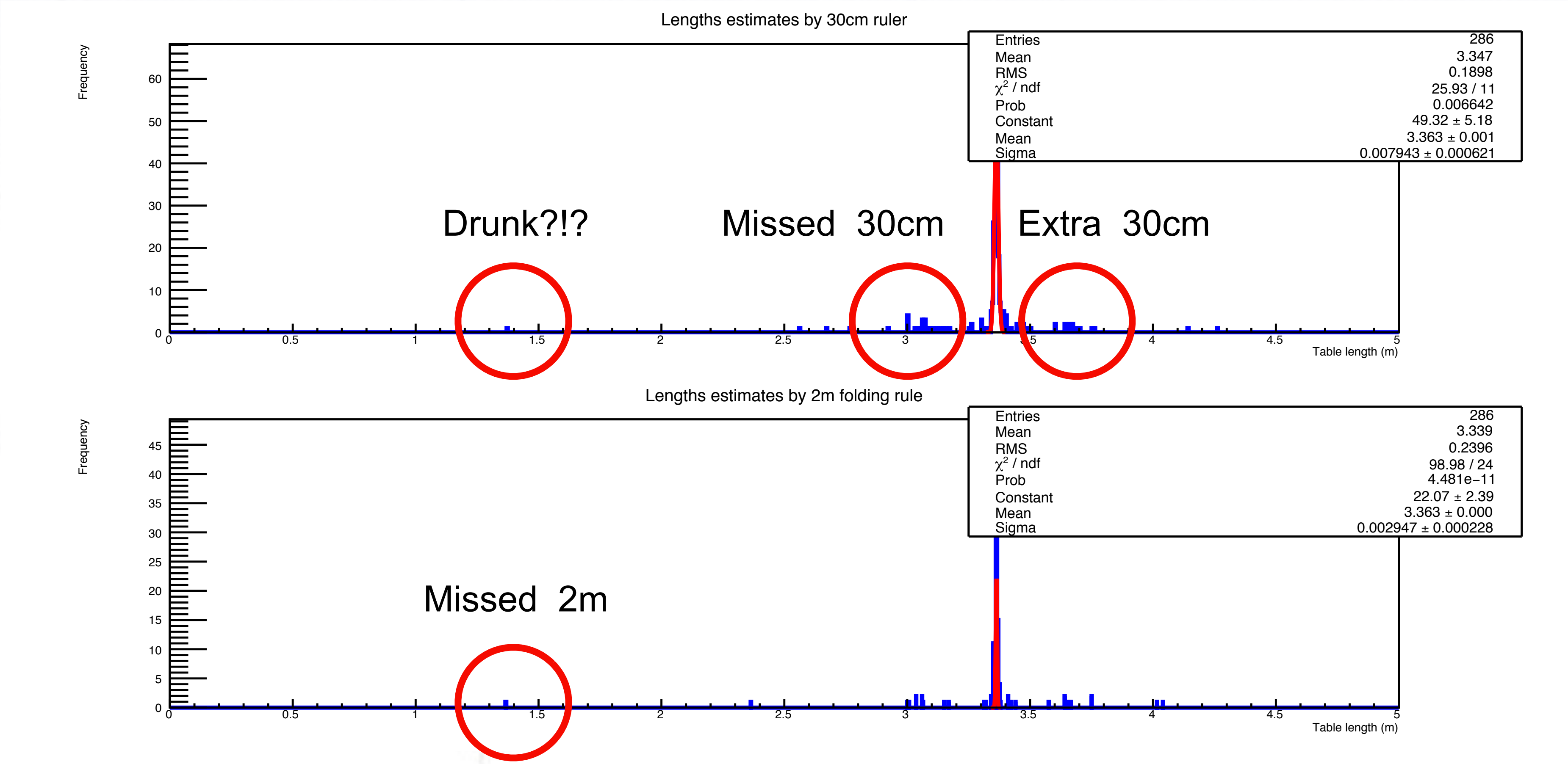
<u>Correspondence between 30cm and 2m measurement:</u>

Diff = -0.00046 ± 0.00086 m (-0.54σ)

Improvement in error from naive to selected measurement 30cm: 18.5
Improvement in error from naive to selected measurement 2m: 44.9

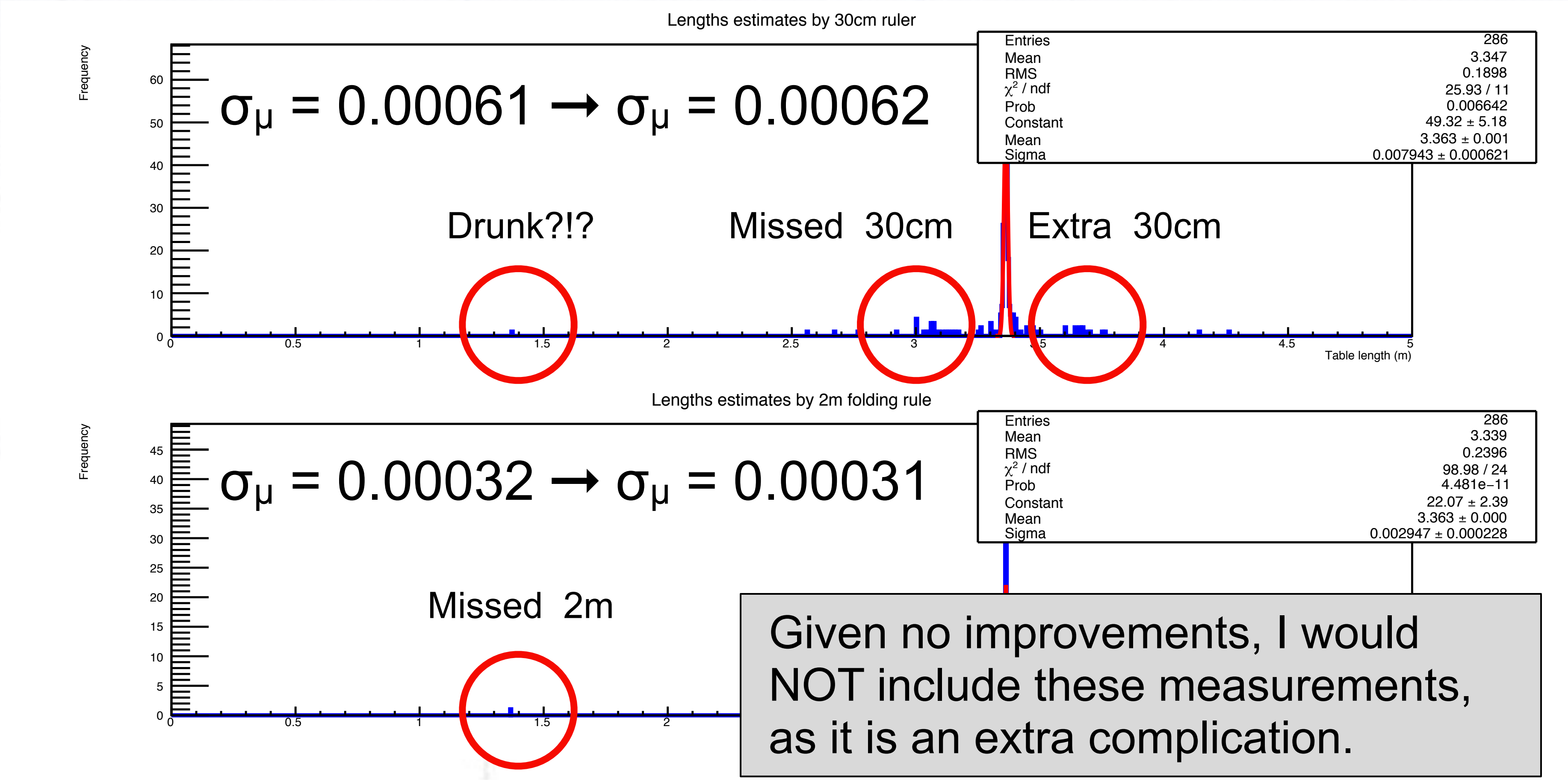
Include offsets?

There are some clear and understandable mis-measurements.
Should one correct and include these?
Depends on resulting improvement, but decide without seeing the final result.



Include offsets?

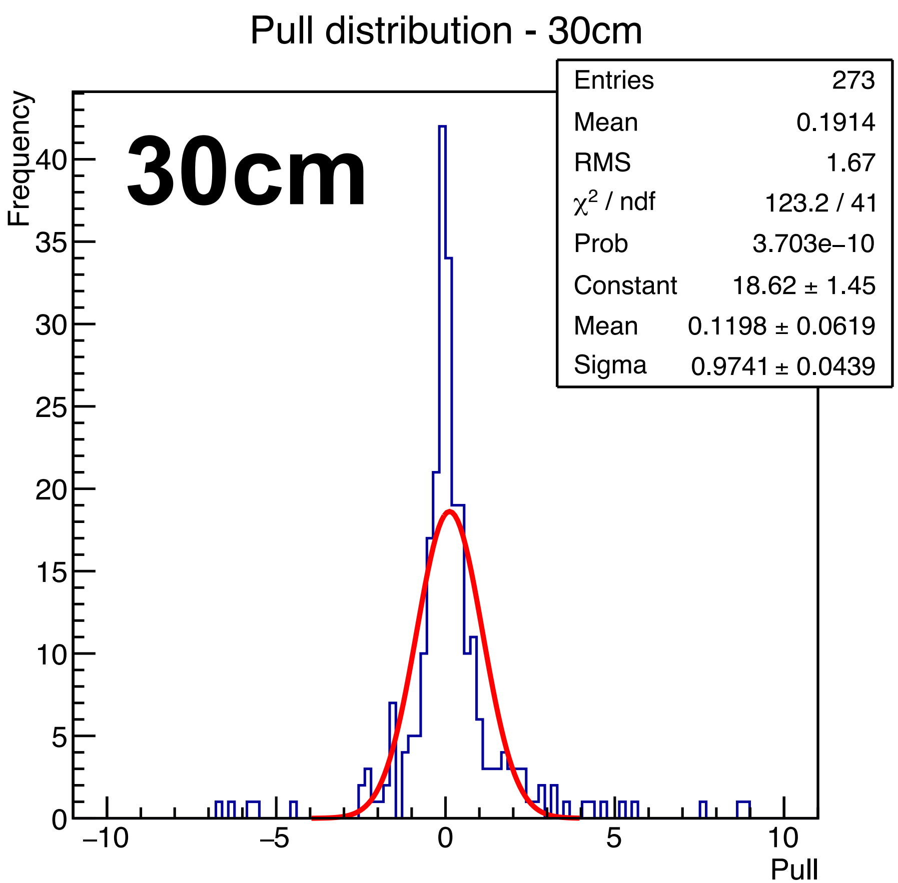
There are some clear and understandable mis-measurements.
Should one correct and include these?
Depends on resulting improvement, but decide without seeing the final result.

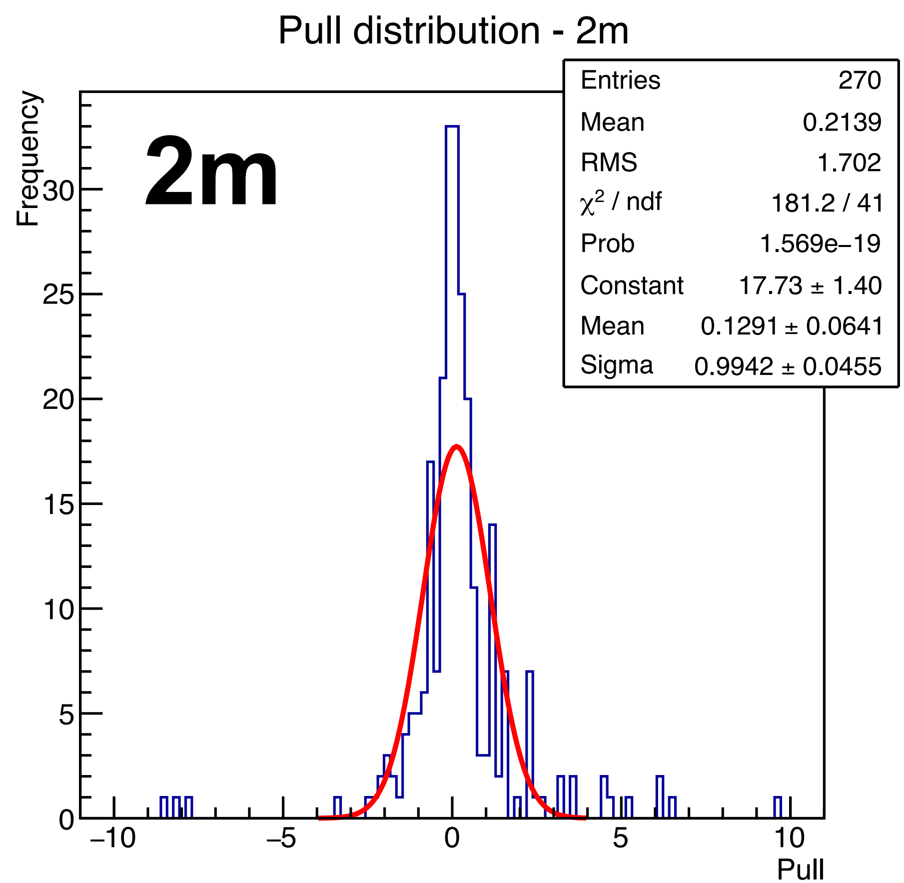


Weighted analysis

Considering the quoted uncertainties, we first need to evaluate their quality.

The plot to consider is a **PULL** plot, i.e. the distribution of: $z = \frac{x_i - \mu}{\sigma_i}$





The pulls should be unit Gaussian - I **decide** to exclude measurements beyond **4σ**.

Weighted results

<u>30cm:</u>

Mean = 3.36371 ± 0.00035 m

RMS = undefined! (N = 230)

<u>2m:</u>

Mean = 3.36366 ± 0.00017 m

RMS = undefined! (N = 239)

<u>Correspondence between 30cm and 2m measurement:</u>

Diff = 0.00005 ± 0.00039 m (0.13 σ)

30cm: $\sigma_\mu = 0.00062 \rightarrow \sigma_\mu = 0.00035$

2m: $\sigma_\mu = 0.00031 \rightarrow \sigma_\mu = 0.00017$

Weighted results

<u>30cm:</u>

Mean = 3.36371 ± 0.00035 m

Chi2 = 222.3, Ndof = 229, Prob = 0.61

<u>2m:</u>

Mean = 3.36366 ± 0.00017 m

Chi2 = 211.9, Ndof = 238, Prob = 0.89

<u>Correspondence between 30cm and 2m measurement:</u>

Diff = 0.00005 ± 0.00039 m (0.13 σ)

30cm: $\sigma_\mu = 0.00062 \rightarrow \sigma_\mu = 0.00035$

2m: $\sigma_\mu = 0.00031 \rightarrow \sigma_\mu = 0.00017$

A problem?

Correspondence: unweighted vs. weighted 30cm meas:

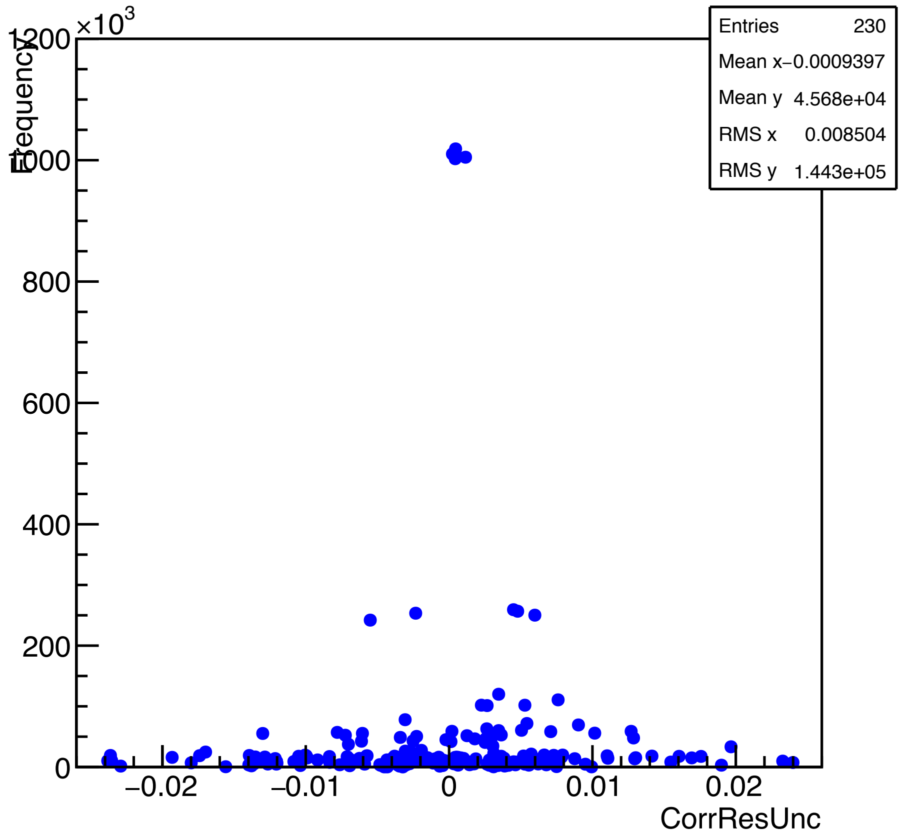
Diff = 0.00145 +- 0.00070 (2.05 sigma)

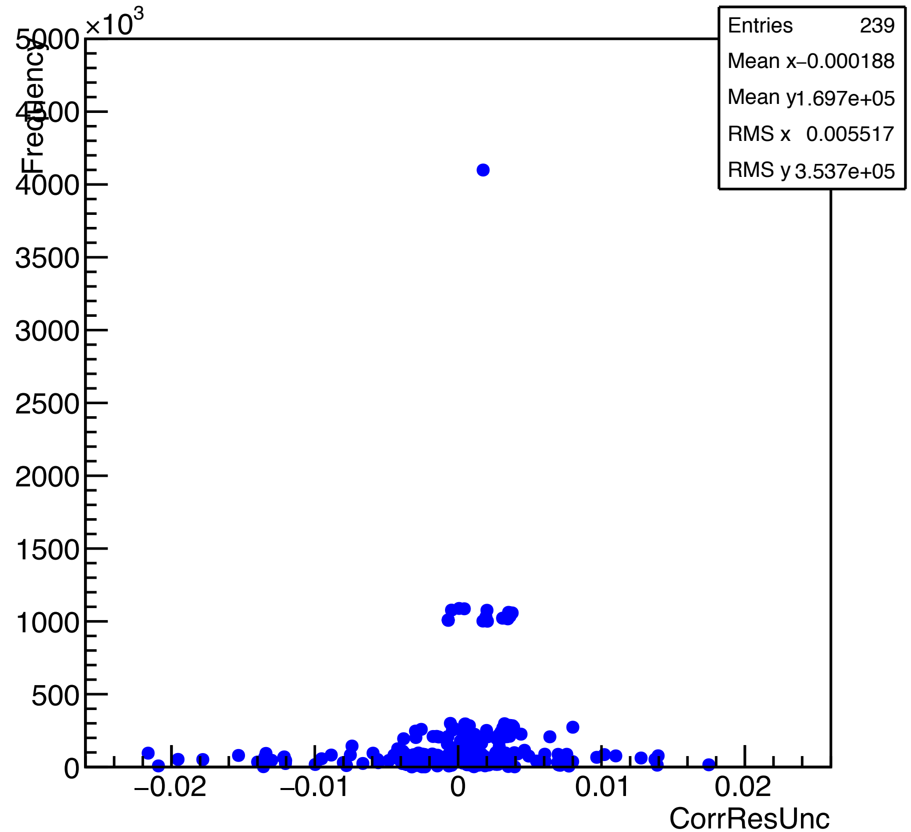
Correspondence: unweighted vs. weighted 2m meas:

Diff = 0.00093 +- 0.00036 (2.57 sigma)

A correlation?

Plotted is the **measurement residuals** (measurement - mean) as a function of the **weight of the measurement** (i.e. 1/uncertainty^2).



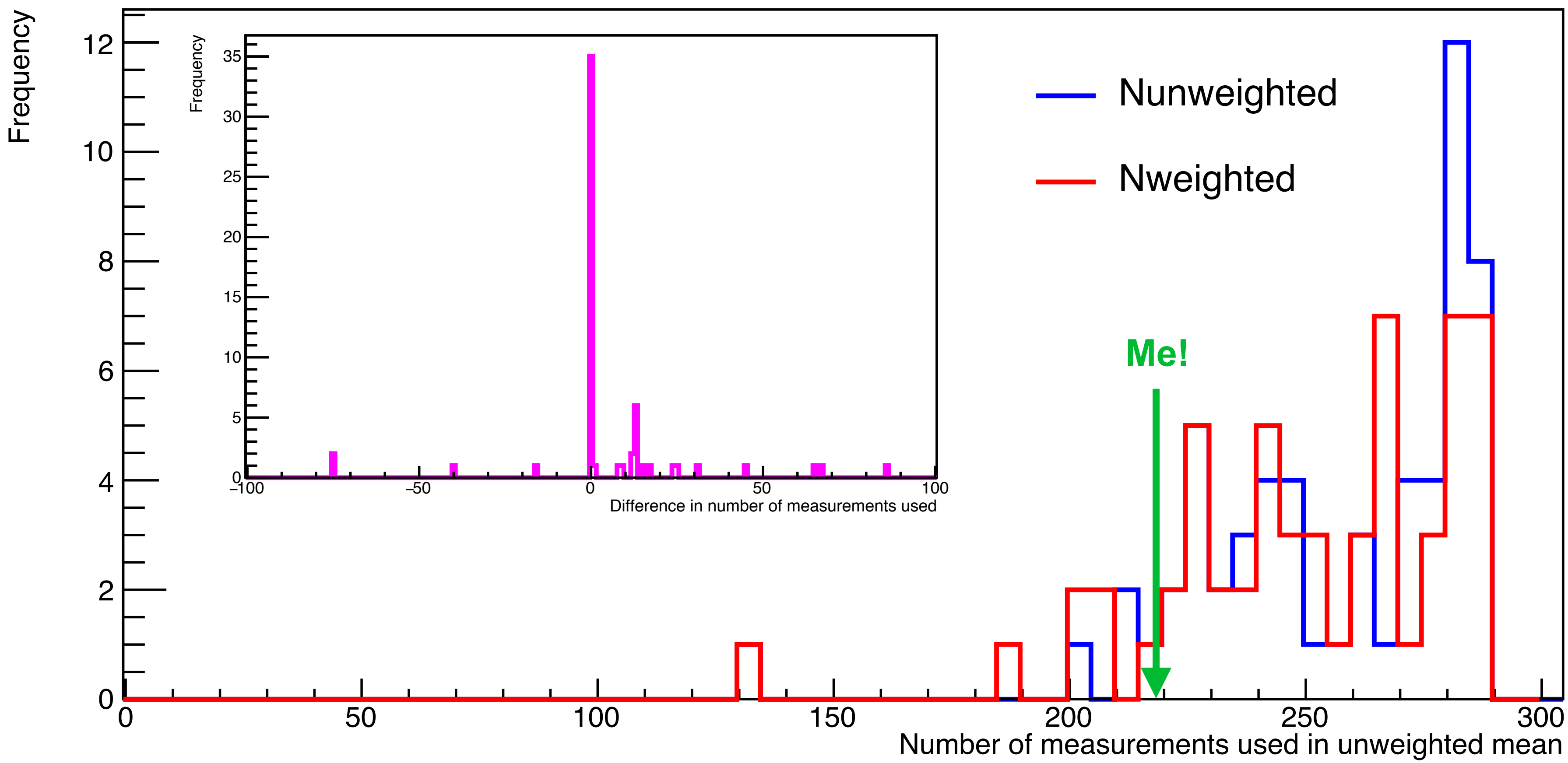


There is a slight correlation between the lengths and uncertainties quoted!

"Longer measurements have smaller errors!". Why? I don't know. But data shows it!

Your measurement results

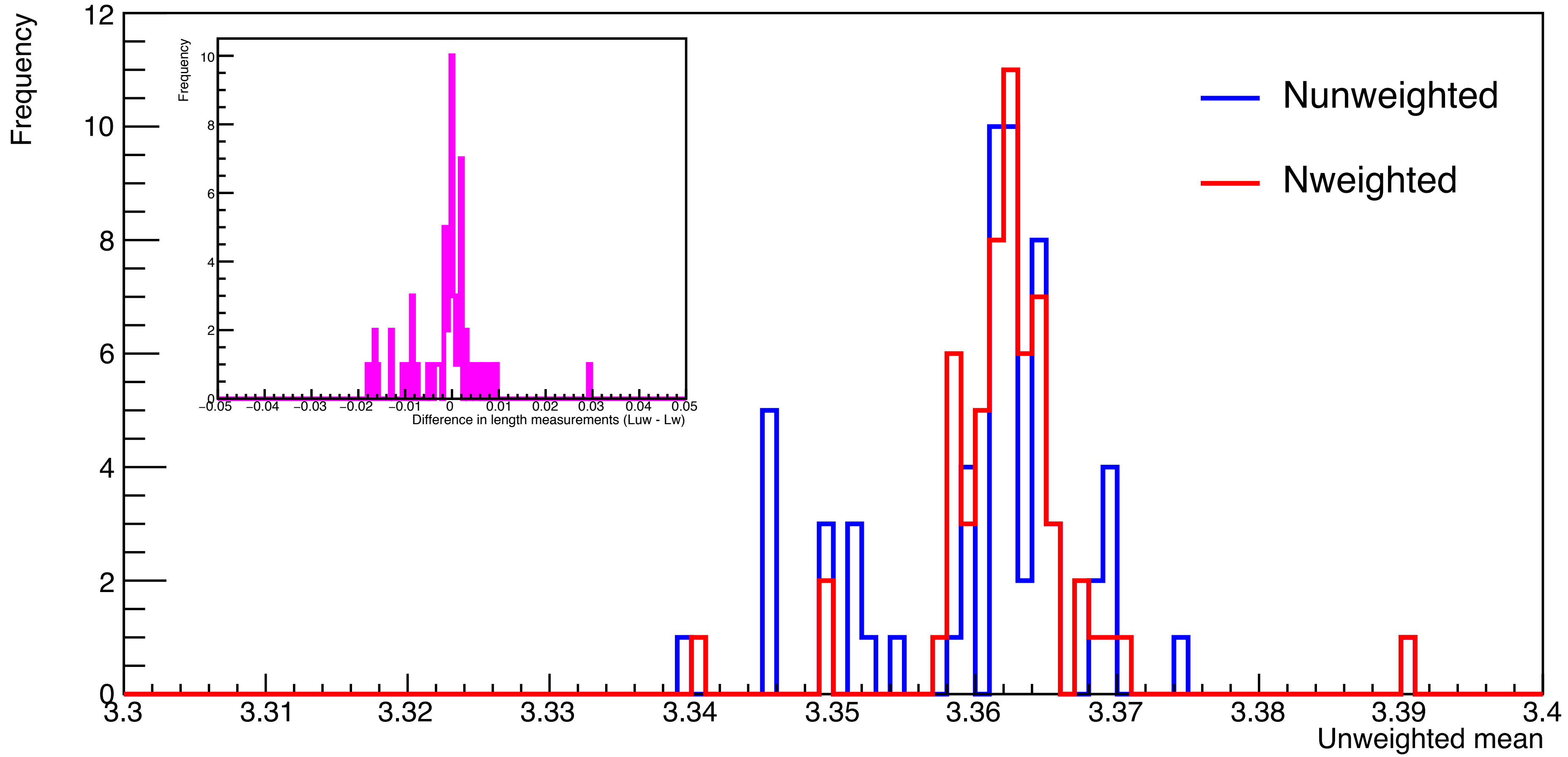
The number of measurements used varied quiet a bit.



But remember that the impact is only sqrt(N), and thus not that important!

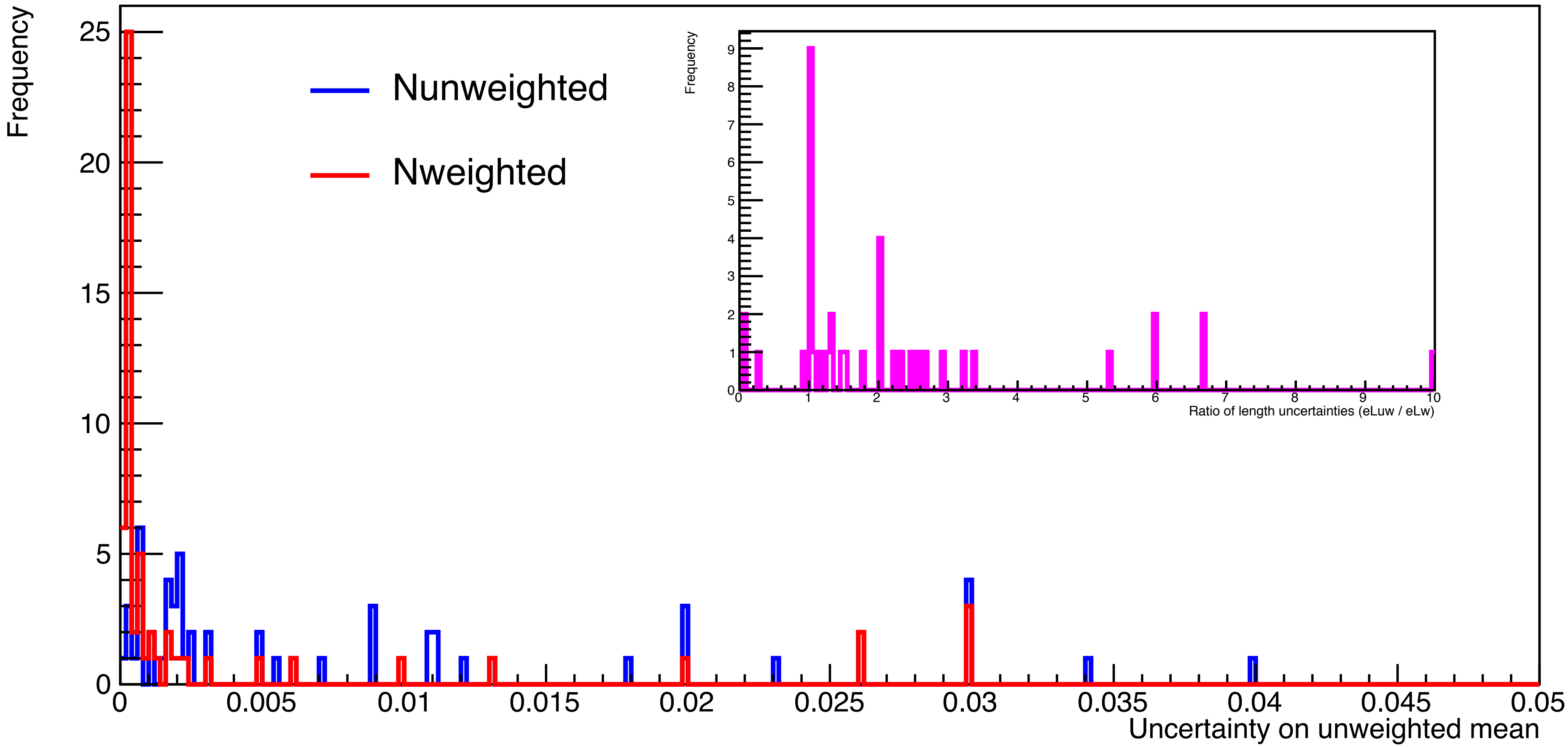
Length results

Results are relatively consistent… 80% of you get a value within 0.1% of “true”



Uncertainties

The uncertainties varied quite a bit - by more than a factor 100! Think about that.



I got: L(unweighted) = 3.36227 ± 0.00061 m, L(weighted) = 3.36371 ± 0.00035 m

Conclusions

<u>Specifically on the analysis:</u>

- Greatest improvement came from simply removing mis-measurements!
- Weighted result was a further improvement, but required good uncertainties.
- The uncertainties are accepted as “reasonable”, as they have good pull distributions, and improve the result.
- The 30cm and 2m results match, giving credibility to the stated precision.

<u>More generally:</u>

- What appears to be a trivial task, turns out to require some thought anyhow.
  (Ask yourself how many fellow students would have been able to get a good result and error?)
- There were several choices to be made in the analysis:
  1. Which measurements to accept.
  2. Which uncertainties to accept.
  3. To correct or discard understood mis-measurements.
- All this can be solved with simple Python/ROOT code.