

Computer Exercise: Toy-experiments of fits of a constant

Macro `p0toyf.C`, accessible at

[http : //www.desy.de/~obehnke/stat/school_mar10/p0toyf.C](http://www.desy.de/~obehnke/stat/school_mar10/p0toyf.C)

generates toy experiments for the fit of a constant (“p0-fit”) from several measurements of the same resolution. A physics example for this would be the determination of the (average) vertical position of a horizontal flying track in track-detectors. Additional Problem: the detectors are noisy, for each detector sometimes instead of a good signal hit a flatly distributed noise hit is observed.

- Steering parameters in the macro:
 - *Ntra* = Number of repeated toy experiments (Default is 1000)
 - *Ndet* = Number of measurements (Default is 10)
 - *frac_noise* Average fraction of noise hits (Default is 0.)
- Output
 - Histo *p0d* - Distribution of residual: fitted constant - true value
 - Histo *chi2d* - χ^2 distribution from the *p0*-fits
 - Histo *pchi2d* - χ^2 -probability distribution from the *p0*-fits
 - Histo *a* - shows the *p0*-fit for the last toy experiment

Tasks:

- Take a deep breath and have a few minutes look in the code and try to understand what’s going on
- a) Run the macro as it is with `.x p0toyf.C` and plot the histograms by `p0d → Draw(); chi2d → Draw(); pchi2d → Draw(); a → Draw();` Fill the RMS value of the *p0d* histogram in the table below (first row) and also the mean values of the *chi2d* and *pchi2d* distributions.

- b) Now edit the macro and set $frac_noise=0.1$;
- run the macro again and fill the obtained values in the table (second row).
 - How much has the RMS of the residuals increased?
 - Can you identify the bad track-fits in the $pchi2d$ distribution?
- c) Rejection of tracks with bad χ^2 -probability: Book another residual histogram $p0d_rej$ in the macro:
- fill it only for the case that the χ^2 -probability is not too small (see also hints in the code)
 - fill in the third row of the table how many tracks survive this selection and the resulting RMS of the residuals.
- d) Hit-outlier-rejection: Advanced method for super experts!
Book another three histograms $p0d_iter$, $chi2d_iter$ and $pchi2d_iter$ in the macro and
- try hit outlier rejection and repeating the track-fit (see detailed instructions in the code);
 - fill the resulting RMS of $p0d_iter$ into the fourth row of the table.
 - How does the RMS results of this method compare to the other cases a)-c)?
- Repeat the above exercises b)-d) for increased $frac_noise=0.5$ and fill the results in the table. How much are the RMS results worse in this case?
 - If time permits repeat the above exercises a)-d) for different number of detectors $ndet = 5$ and study how the results change qualitatively.

| Task | N_{det} | $frac_{noise}$ | Outlier rejection | Resid.-distr. | #tracks | Resid. RMS | chi2d mean | pchi2d mean |
|------|-----------|----------------|--|---------------|---------|------------|------------|-------------|
| a) | 10 | 0.0 | No | p0d | 1000 | 0.0032 | 8.8 | 0.51 |
| b) | 10 | 0.1 | No | p0d | 1000 | 0.0064 | 15.9 | 0.245 |
| c) | 10 | 0.1 | Reject tracks with bad pchi2 | p0d_rej | 596 | 0.0036 | – | – |
| d) | 10 | 0.1 | Outlier hit-rejection and repeated track-fit | p0d_iter | 1000 | 0.00396 | – | – |
| b) | 10 | 0.5 | No | p0d | 1000 | 0.0133 | 34 | 0.0053 |
| c) | 10 | 0.5 | Reject tracks with bad pchi2 | p0d_rej | 23 | 0.004 | – | – |
| d) | 10 | 0.5 | Outlier hit-rejection and repeated track-fit | p0d_iter | 1000 | 0.0099 | – | – |