School Of Statistics

26-30 mai 2014

Autrans

Jérôme Baudot
(baudot@in2p3.fr)

Institut Pluridisciplinaire Hubert Curien Strasbourg

2 applications of basic principles

-Exponential fit with cuts

-Uncertainty on a low or high efficiency

Exponential fit with cuts

- Reminder on simple maximum likelihood fit to exponential lifetime distribution
- Taking into account measurement range
- Crosschecking

- Based on: Reduction of the statistical power per event due to upper lifetime cuts in lifetime measurements by Jonas Rademakers, arXiv:hap-ex/0502042

Presenting the problem

- Measurements (=statistics):
  - N times: t_i, i=1...N
  - t_i = realisation of a random variable T
  - T is expected to follow an exponential law

- Exponential law:
  - Prob. Dens. Function pdf for $t \in [0,\infty]$

$$f(t) = \frac{1}{\tau} \exp^{-\frac{t}{\tau}}$$

- "Ideal case" assumptions:
  - T has same excursion $[0,\infty]$
  - No inefficiency, no finite resolution "perfect measurements"

Solution of the ideal case

- **The maths:**

$$f(t) = \frac{1}{\tau}\,\exp^{-\frac{t}{\tau}} \qquad \begin{cases} \displaystyle\int_0^\infty f(t)\,\mathrm{d}t = 1 \\ \displaystyle <t> = \int_0^\infty t.f(t)\,\mathrm{d}t = \tau \end{cases}$$

$$L = \log\mathcal{L} = \log\left(\prod_{i=1}^{N} f(t)\right) \qquad L = -N\,\log(\tau) - N\sum_{i=1}^{N}\frac{t_i}{\tau}$$

$$\frac{\mathrm{d}\,L}{\mathrm{d}\tau} = \frac{1}{\tau^2}\left(-N\tau + \sum_{i=1}^{N} t_1\right) \Rightarrow \tau = \frac{\sum_{i=1}^{N} t_i}{N}$$

$$\frac{\mathrm{d}^2 L}{\mathrm{d}\tau^2} = -\frac{2}{\tau}\frac{\mathrm{d}L}{\mathrm{d}\tau} - \frac{1}{\tau^2}N \Rightarrow \sigma_\tau = \frac{\tau}{\sqrt{N}}$$

parabolic assumption

A new problem with a finite measurement range

- New pdf:
  - T has a smaller excursion [tmin, tmax] with Δt = tmax-tmin
  - $f_{cut}(t)$ = A f(t), with A=normalisation /

$$\int_{tmin}^{tmax} A.f(t)\,\mathrm{d}t = 1 \quad \Rightarrow \quad A = \exp^{\frac{t_{min}}{\tau}} \left(1 - \exp^{-\frac{\Delta t}{\tau}}\right)^{-1}$$

$$f_{cut}(t) = \frac{1}{\tau\left(1 - \exp^{-\frac{\Delta t}{\tau}}\right)} \exp^{-\frac{(t - t_{min})}{\tau}}$$

$$L = -N\log(\tau) + \sum_{i=1}^{N} \frac{(t_i - t_{min})}{\tau} - \sum_{i=1}^{N} \log\left(1 - \exp^{-\frac{\Delta t}{\tau}}\right)$$

Solution with a finite measurement range

$$\frac{\mathrm{d}\,L}{\mathrm{d}\tau} = \frac{1}{\tau^2}\Big(-N\tau + \sum_{i=1}^{N}(t_i - t_{min}) + \sum_{i=1}^{N} \frac{\Delta t}{\exp^{-\frac{\Delta t}{\tau}} - 1}\Big)$$

$\underline{\tau}$ not so easy...

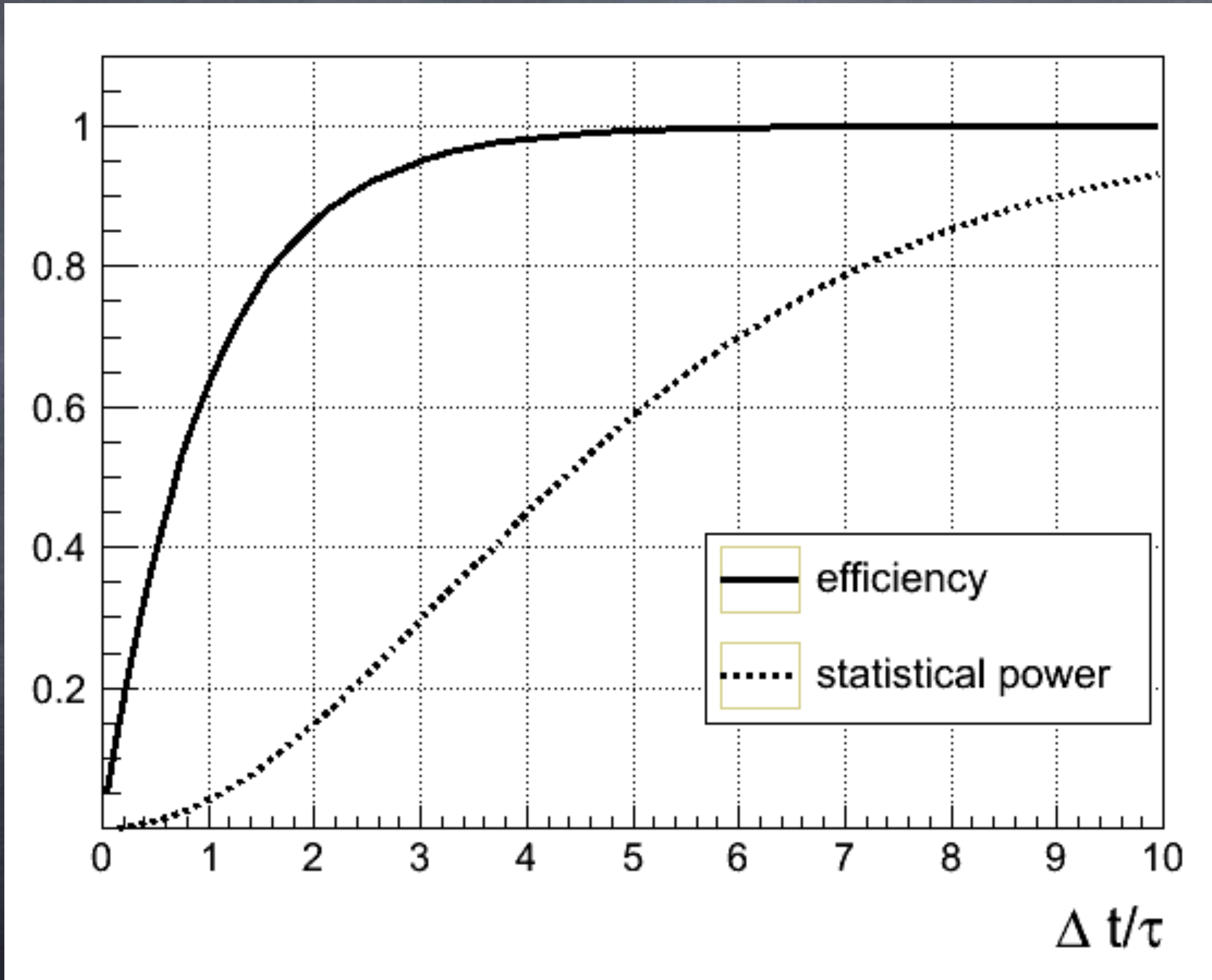
$$\frac{\mathrm{d}^2 L}{\mathrm{d}\tau^2} = -\frac{2}{\tau}\frac{\mathrm{d}L}{\mathrm{d}\tau} - \frac{1}{\tau^2}\left(N - \sum_{i=1}^{N}\left(\frac{\Delta t/2\tau}{\sinh(\Delta t/2\tau)}\right)^2\right)$$

$$\sigma_{\underline{\tau}} = \frac{\tau}{\sqrt{N}} \frac{1}{\sqrt{1 - \left(\frac{1}{\sinh(\Delta t/2\tau)}\right)^2}}$$

Statistical power = $\mathcal{P} = 1 - \left(\frac{1}{\sinh(\Delta t/2\tau)}\right)^2$

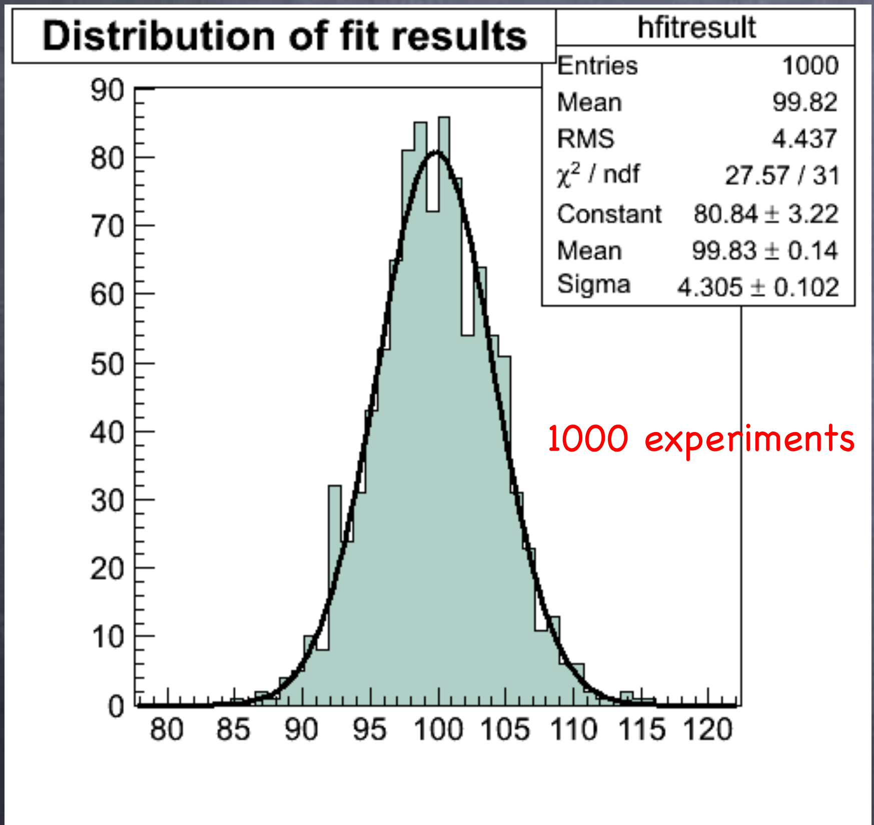
N → N × $\mathcal{P}$

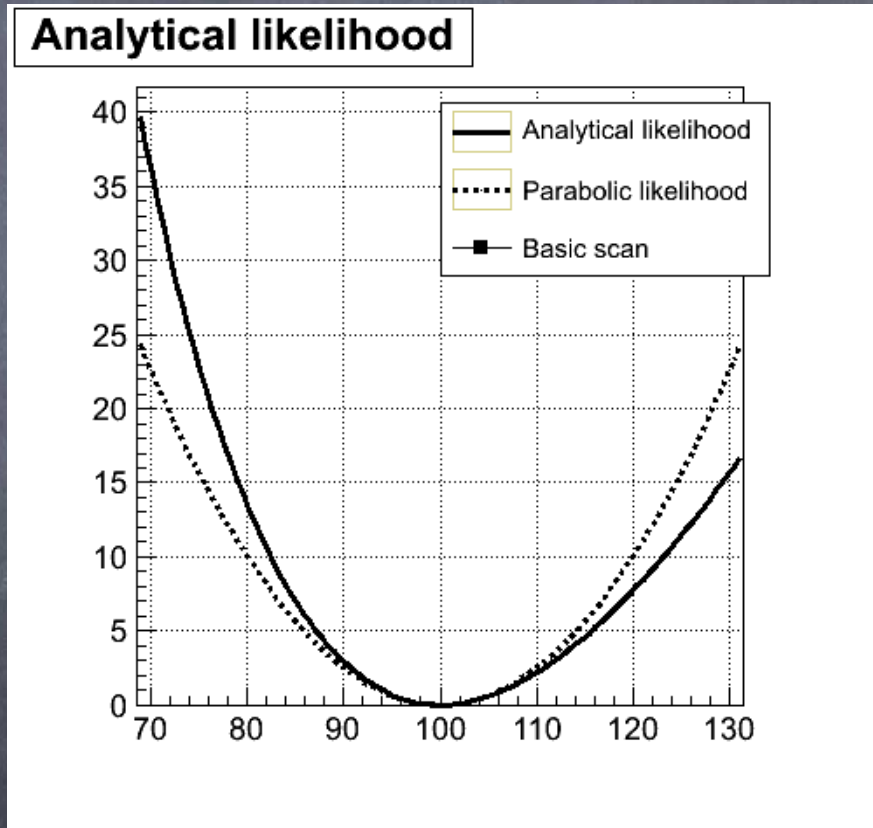
Statistical power



Crosscheck

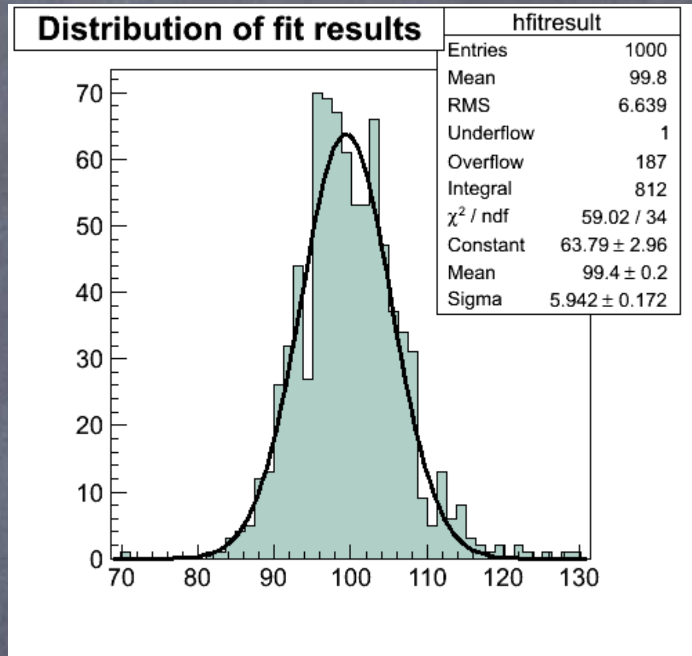
- Toy MC
  - Use your computer to see how it works !
  - Reproduce many time the same experiment = same statistic size
    ➡ distribution of fit results provide bias and statistical uncertainty





True τ = 100, N = 500, $t \in [0,1000]$: expected $\sigma_\tau = 4.4$

More trials



True τ = 100, N = 500, $t \in [500,800]$:
expected $\sigma_\tau = 6.3$

True τ = 100, N = 500: expected $\sigma_\tau = 4.4$

Conclusion / Oulooks

- For the maximum likelihood method
  - The quality of your estimator depends on the correctness of the pdf used
  - Trivial to say...might be more difficult to crosscheck:
    - Control the normalisation on the real range of your measurements
    - Use a toy MonteCarlo

- Do it yourself !
  - Macro minuitFitMultiExp.C

- Next problem: how to deal with nuisance parameters?
  - Efficiency dependent on t
  - Finite resolution on t measurement
  - See next lecture: Profile Likelyhood