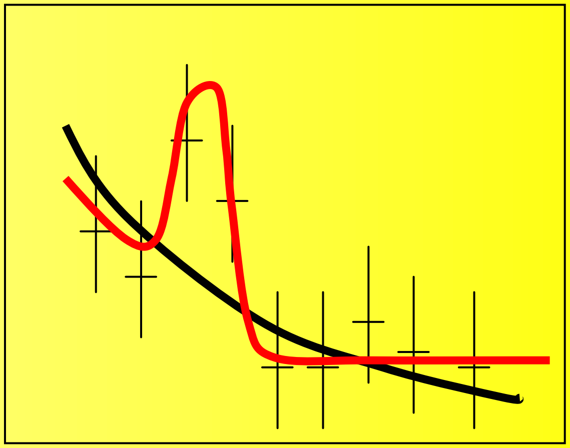
Significance: Does $\Delta\chi^2$ give σ^2?

Roger Barlow

Montreal Collaboration meeting

June 2006

How it started



Analysis looking for bumps

Pure background gives χ^2_{old} of 60 for 37 dof (Prob 1%).

Not good but not totally impossible

Fit to background+bump (4 new parameters) gives better χ^2_{new} of 28

Question:

Is this improvement significant?

Answer:

Significance is $\sqrt{(\chi^2_{new} - \chi^2_{old})}$

$= \sqrt{(60-28)} = 5.65$

Schematic only!!

No reference to any other resonance, real or fictitious

Puzzle. How does a 3 sigma discrepancy become a 5 sigma discovery?

Justification?

- ‘We always do it this way’
- ‘Belle does it this way’
- ‘CLEO does it this way’

Possible Justification

Likelihood Ratio Test

a.k.a. Maximum Likelihood Ratio Test

If M_1 and M_2 are models with max. likelihoods L_1 and L_2 for the data, then $2\ln(L_2 / L_1)$ is distributed as a χ^2 with $N_1 - N_2$ degrees of freedom

Provided that

1. M_2 contains M_1 ✔
2. Ns are large ✔
3. Errors are Gaussian ✔
4. Models are linear ✗

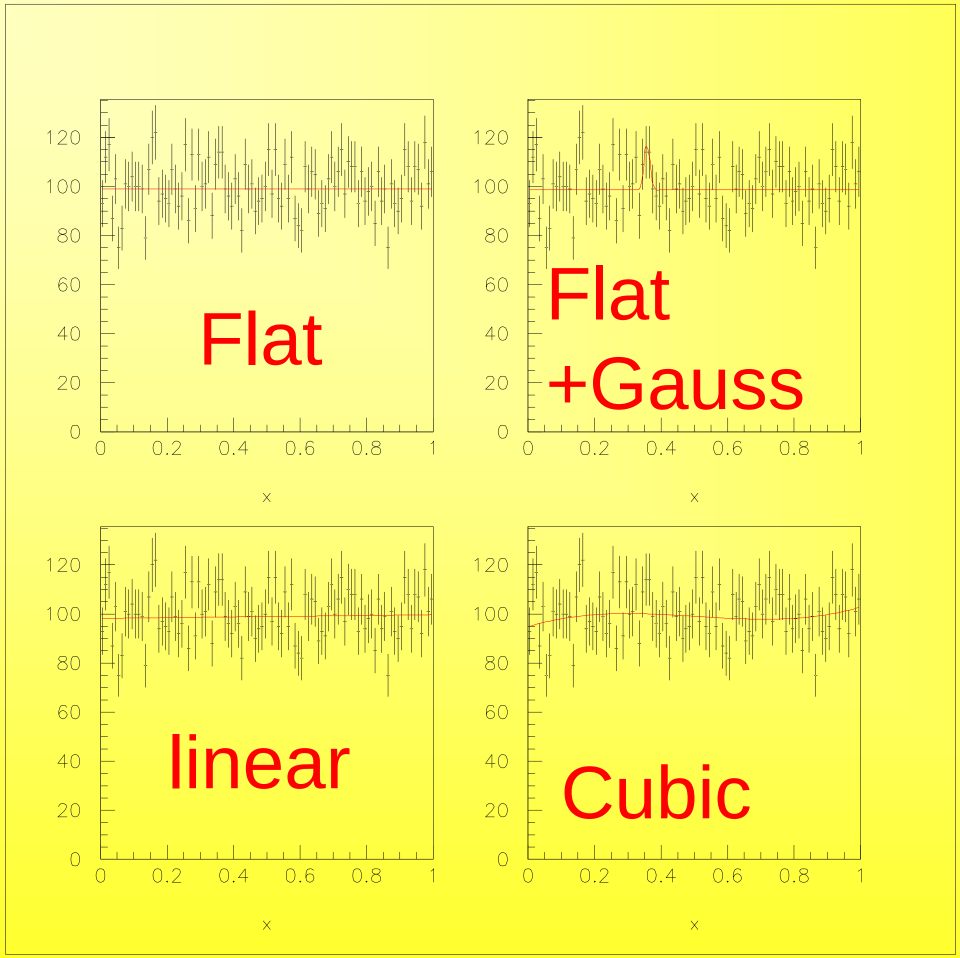
Does it matter?

- Investigate with toy MC
- Generate with Uniform distribution in 100 bins, <events/bin>=100. 100 is large and Poisson is reasonably Gaussian
- Fit with
  - Uniform distribution (99 dof)
  - Linear distribution (98 dof)
  - Cubic (96 dof)
  - Flat+Gaussian (96 dof)

Cubic is linear(!) $a_0+a_1\, x + a_2\, x^2 + a_3\, x^3$

Gaussian is not linear in μ and σ

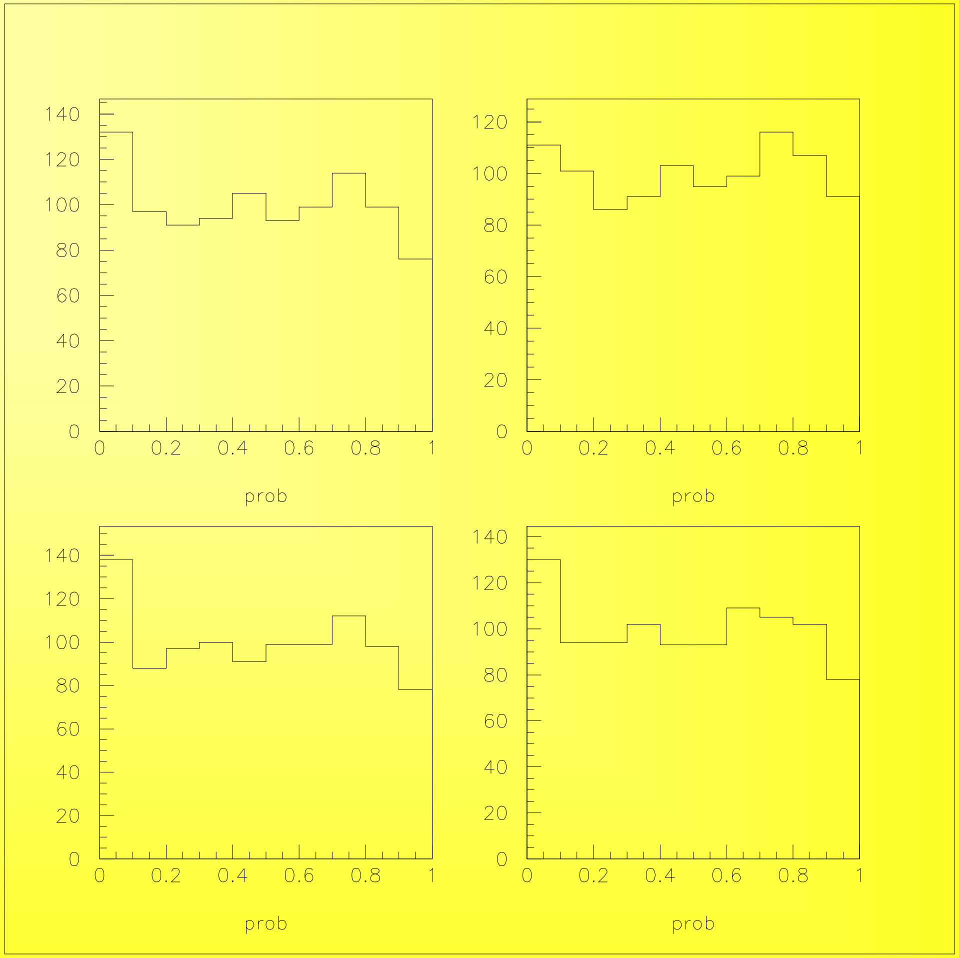
One ‘experiment’



Calculate χ^2 probabilities of models on their own

From the χ^2 and N_{dof} for the 4 models. Not bad. Show Gaussian approximation is working fairly well.

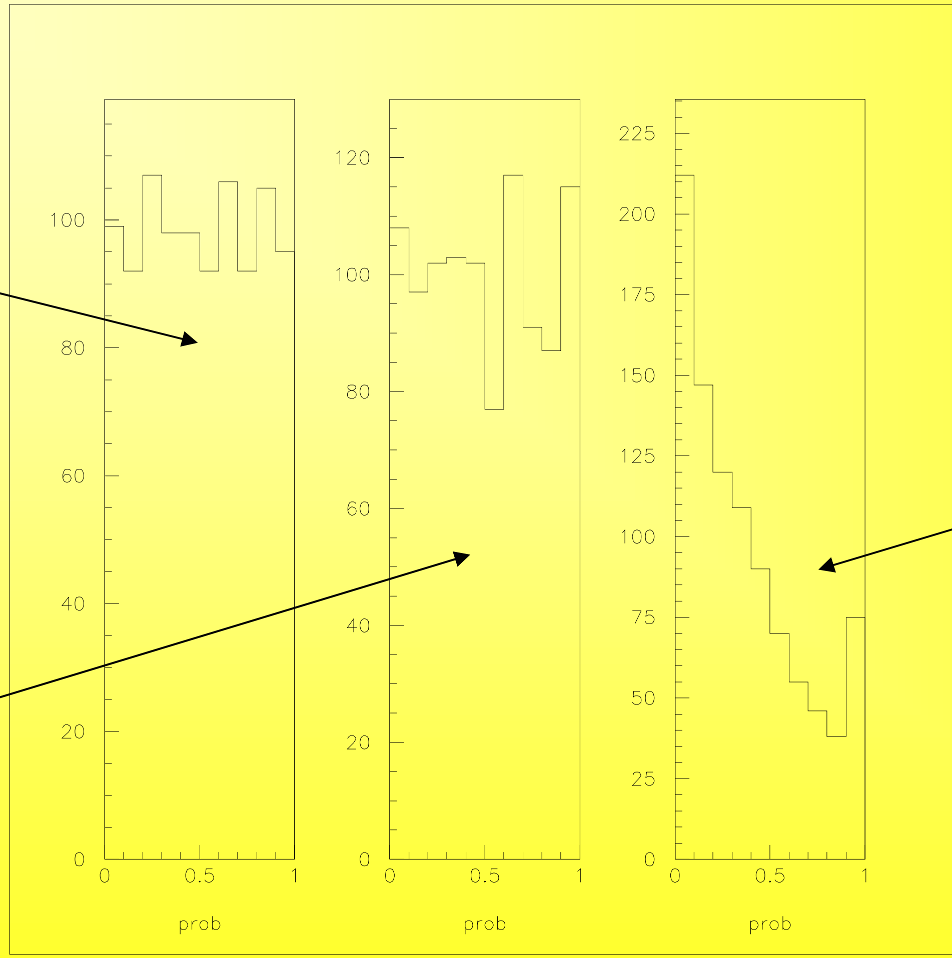
Ideal would be flat for all



Calculate χ^2 probabilities of differences in models

Compare linear and uniform models. 1 dof. Probability flat Method OK

Compare cubic and uniform models. 3 dof. Probability flat Method OK



Compare flat+gaussin and uniform models. 3 dof. Probability very unflat

Method invalid

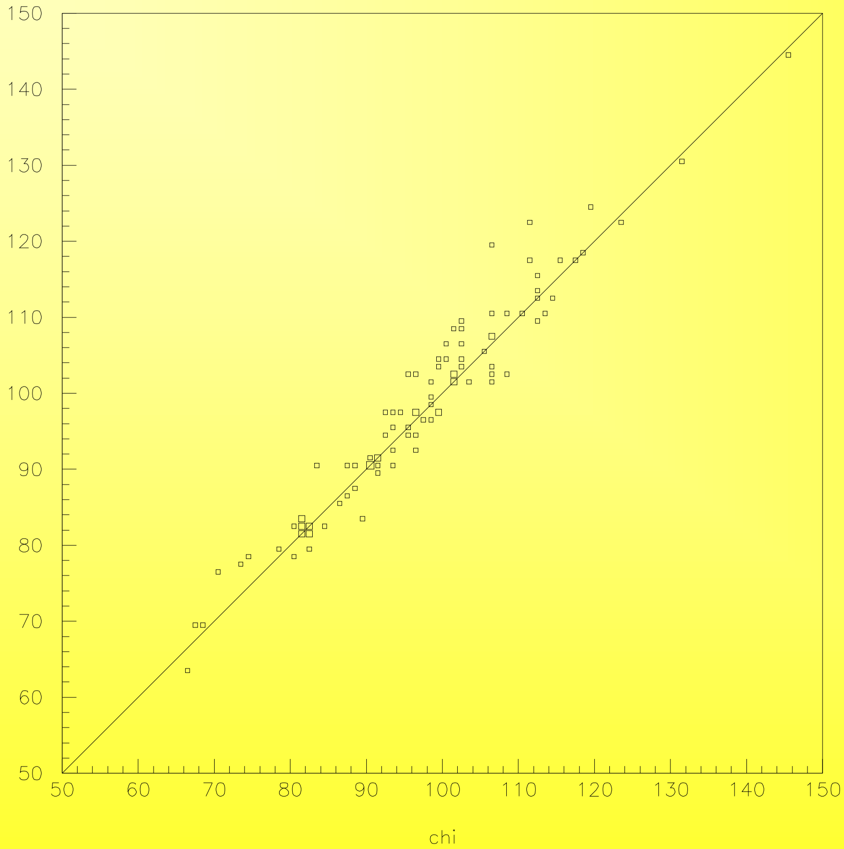
Peak at low P corresponds to large $\Delta\chi^2$ i.e. false claims of significant signal

Not all parameters are equally useful

Shows χ^2 for flat+gauss v. cubic

Same number of parameters

Flat+gauss tends to be lower



If 2 models have the same number of parameters and both contain the true model, one can give better results than the other.
This tells us nothing about the data

Helpful (?) way of thinking

- Method _is_ valid if you fix Gaussian position and width and just vary size (1 dof – and linear). OK for investigating a known peak
- Intuitively sensible for small σ: you fit the known bin exactly. Contribution ($\approx$1) to χ^2 is zapped.
- If you keep σ small and float μ this gives your fit the power to zap the bin with highest $\chi^{2.}$.That is larger, tricky to calculate, and depends on the number of bins. Result not guaranteed to be χ^2

Conclusion: Does $\Delta\chi^2$ give σ^2?

No