Systematic Uncertainties

Matthew Kenzie
CERN

Precision Measurements School
Meinerzhagen

September 23, 2015

1. The joke

A classic joke

- A mathematician, an engineer and a physicist are all presented with a cow
- They are asked to use the cow to determine the average mass of all cows

A classic joke

- A mathematician, an engineer and a physicist are all presented with a cow
- They are asked to use the cow to determine the average mass of all cows

1. The mathematician: cow function $m = \int_V D(x, y, z) dxdydz$

A classic joke

- A mathematician, an engineer and a physicist are all presented with a cow
- They are asked to use the cow to determine the average mass of all cows

1. The mathematician: cow function $m = \int_V D(x,y,z) dxdydz$
2. The engineer: submerge the cow in water

A classic joke

- A mathematician, an engineer and a physicist are all presented with a cow
- They are asked to use the cow to determine the average mass of all cows

1. The mathematician: cow function $m = \int_V D(x, y, z) dxdydz$
2. The engineer: submerge the cow in water
3. The physicist: assume the cow is a uniform sphere of water $m = \frac{4}{3}\pi r^3 \cdot 1\text{kg}/\text{m}^3$

What's the point?

- The point is that systematic uncertainties are not particularly well defined as a problem
- As physicists we usually do as simple a job as we can with the statistics we have available
- We ALWAYS make assumptions - you cannot possibly know everything there is to know about an experiment.
- As long as you can quantify how big an effect these assumptions COULD have (and correct for them)
- There are many ways of evaluating them - shall we evaluate them for the cow?

2. Introduction

Jargon

This is some jargon which I assume you already know

If you don't know what one of these means then SPEAK NOW! (I won't bite[i])

Cow	A large farm yard mammal used mainly for the production of milk, cheese, yoghurt, beef and leather
PDF	Probability Density Function
Likelihood	Likelihood function
DLL (ΔLL)	Difference in log likelihood from minimum
χ^2 ($\Delta\chi^2$)	Goodness of fit (difference in χ^2 from minimum)
MC	Monte Carlo simulation
Toy	A pseudo-experiment generated from a distribution (PDF) i.e. a single MC event
Toys	A set of the above
Pull	Difference between fitted value and true (generated) value divided by the uncertainty, $(\hat{\mu} - \mu)/\sigma$

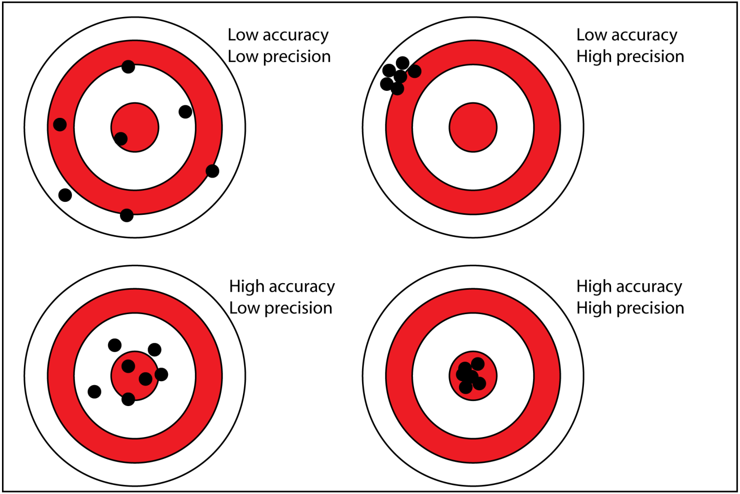
I also assume you have seen the following:

- Central limit theorem (i.e why things are Gaussian distributed)
- Propagation of uncertainties

[i] at least not hard ;)

Precision vs Accuracy

- Is it better to be accurate or precise?
- Statistical / systematic uncertainty

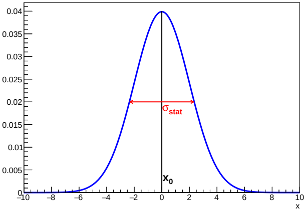


Precision vs Accuracy

In any experiment there are usually two sources of error

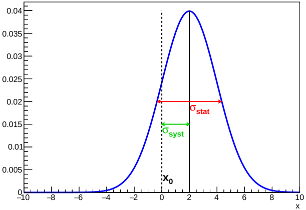
Statistical - accuracy

- The inherent random error for an experiment
- Related to the accuracy of the equipment
  - A standard wrist watch measures to within one second
- Also related to the size of the data sample
- Should be randomly distributed around the true value



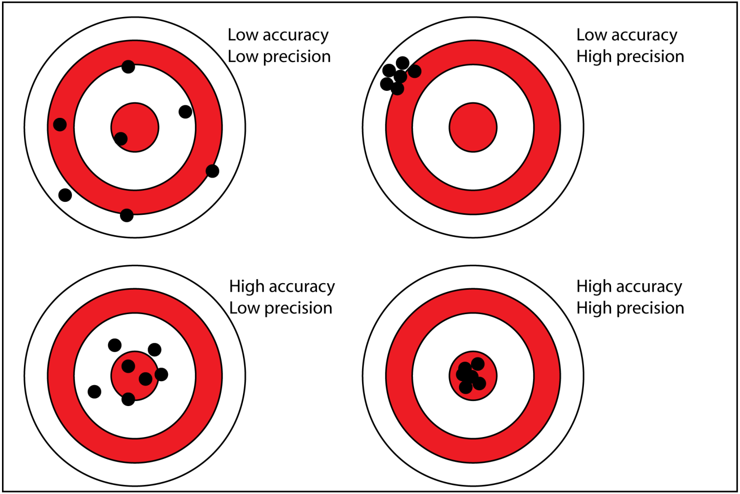
Systematic - precision

- For whatever reason measurement is simply wrong
  - Watch is running ten minutes slow
- Will be wrong everytime
- Offset from the true value
- Typically has a statistical component
- Usually we correct for them and the statistical uncertainty on the correction is the systematic



Precision vs Accuracy

- The picture in real data analysis is usually a mixture of these two
- Often use a subset of data for calibration, checks and systematic study



Systematic effect and systematic error

Know the distinction

- A systematic effect is a general category of things which can cause a shift or bias in your result
  - energy resolution
  - background
  - efficiency
  - dead time
  - etc.
- A systematic error is the uncertainty in the estimation of a systematic effect
  - Correct for the systematic effect
  - The uncertainty on the correction is the systematic uncertainty
  - Sometimes you cannot correct so you have to try and cover with an appropriate uncertainty

Know the difference between a check and a systematic

- Checks can be things like
  - Split data into seperate years / magnet config / run numbers and check consistency
- A check is not a systematic study
- If you do a check and cannot explain the outcome then the last resort should be a systematic uncertainty

Which systematics are important?

How do I know which systematics to include?

YOU DON'T!

Known unknowns

- Think!
- Talk to colleagues
- Talk to the most miserable and aggersive collaborator you know
- Check to see what other similar analyses have done
- Keep thinking

Unknown unknowns

- These are what the checks are for
- Checks should convince you (and others) you haven't missed something that can have a big effect
- Subsequently think of good checks
- Should be in a position where anything you might have missed must have a small effect such that it doesn't need to be considered

What reduces uncertainties?

Statistical

Systematic

What reduces uncertainties?

Statistical

- More data / more events ?
- More signal events, $\sigma_S \approx S/(\sqrt{S+B})$
- Fewer free parameters

Systematic

What reduces uncertainties?

Statistical

- More data / more events ?
- More signal events, $\sigma_S \approx S/(\sqrt{S+B})$
- Fewer free parameters

Systematic

- Often more data also (control channels, alignment etc.)
- Better simulation
- Better ideas (data driven methods)
- Understanding your equipment

Estimating systematic errors

What the experts say

Louis Lyons

"In general, there are no simple rules or prescriptions for eliminating systematic errors. To a large extent it is simply a matter of common sense plus experience."

Statistics for nuclear and particle physicists

Roger Barlow

"Systematic errors cover a spectrum from the mildly inconvenient to the downright catastrophic"

Statistics: A guide to the use of statistical methods in the physical sciences

3. Evaluating systematic uncertainties

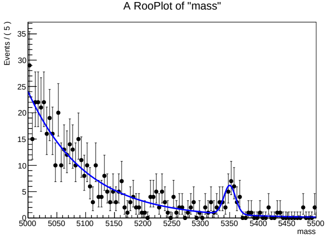
Evaluating sytematic uncertainties

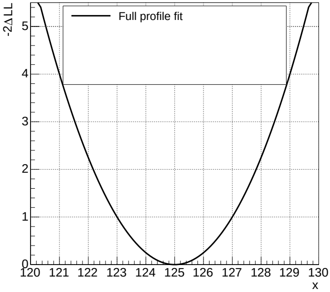
- Think carefully about the whole experiment
- Particularly any numbers applied to the data
  - Efficiencies, calibration etc.
- Study what effect these have
- If the effect is small then apply a systematic uncertainty to incorporate this
- If the effect is large then it should be corrected for and an uncertainty introduced for the correction
- Often there is a trade off between statistical and systematic uncertainties
  - Optimal point statistically is $S/\sqrt{S+B}$ but you may want to be tighter or looser than this
  - How much freedom do you give to your model

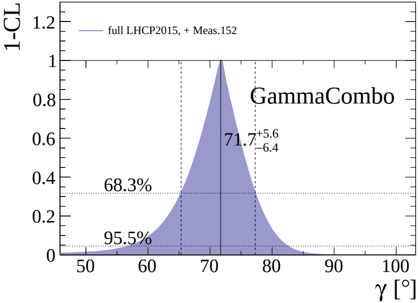
How do we do quote systematic uncertainties?

Example

- We have some data we've selected
- We have a model we've constructed to describe the data
- Perform an NLL or χ^2 fit
- Get statistical uncertainty
- Systematic?







How do we quote systematic uncertainties?

Method 1 - add in quadrature

- Estimate the size of various systematic uncertainties
- Add them in quadrature
- Sometimes this is not possible (or at least difficult) e.g. limit setting

Method 2 - incorporate in the model

- Estimate the size of various systematic uncertainties
- Add terms to your likelihood which parametrise your "ignorance"
- Profile likelihood
- Also tells you how well your data "agree" with the systematic (nuisance parameter pull)
- Sometimes this is not possible (or at least difficult) e.g. complicated fits

$$\mathcal{L}(\theta, \nu) = P_x(x|\theta, \nu) P_y(y|\nu)$$

Evaluating systematic uncertainties

- Look at difference between data with guessed effect and without
- Generate MC with the effect in and assess the average difference when extracting using the nominal model
  - very common and well accepted
- Use a control channel
- Use simulation
- Use some data driven method

There is no right way

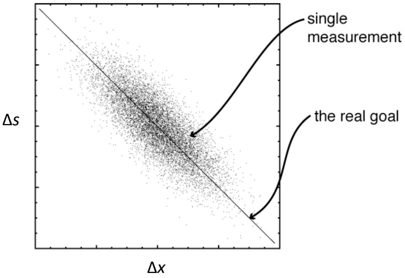
- There are just come commonly used / accepted methods
- If you find a better one then inform the community

Evaluating systematic uncertainties

- In general once you have a source of systematic uncertainty we want to know how does a change from this effect alter my result
- In other words we want to know $\frac{\Delta\hat{x}}{\Delta s}$ for our measured quantiy, $\hat{x}$ and a systematic effect, s
- Then we can evalute the systematic error:

$$\sigma_{\hat{x}} = \left(\frac{\Delta\hat{x}}{\Delta s}\right)\sigma_s$$

- The differential $\frac{\Delta\hat{x}}{\Delta s}$ can be calculated by generating pseudo-data
  - Make a change Δ_s, generate toys, calculate the effect on $\hat{x}$
- Often knowledge of σ_s is difficult
  - E.g. size of uncertainty in theoretical prediction from factorisation assumption?

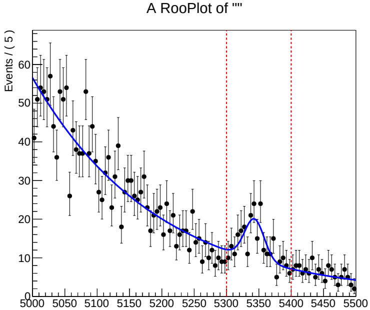


Some examples

- Systematics are difficult
  - If you say that you like them people will think you're weird
  - They usually take up the most time for a precision analysis (the selection, fitting etc. are easy in comparison)
- There is no common / predefined approach
  - Many experiments have guidelines or common methods for particular types of systematic
- Most textbooks only mention them passingly
- Of course this makes it rather difficult to teach
- I will try and cover a few basic examples which are common to HEP

Efficiency correction

Measure the branching ratio of a decay



- Select events in mass window which gives, $N_S \pm \sigma_{N_S}$
- Have determined the efficiency of all selection criteria from MC simulation, $\epsilon = 10 \pm 1\%$
  - Including detector acceptance, trigger and selection requirements
- Measured the luminosity as $4\,\text{fb}^{-1} \pm 4\%$
- Know the cross section for producing such particles in my accelerator, $xs \pm \sigma_{xs}$

Result

$\mathcal{B} = X \pm \sigma_{\text{stat}} \pm \sigma_{\text{syst}}$

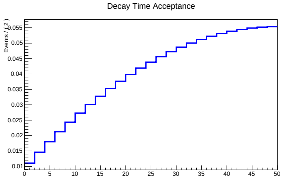
- Statistical uncertainty from number of events (propagated from σ_{N_S})
- Systematic uncertainty propagated from ϵ, luminosity, cross section

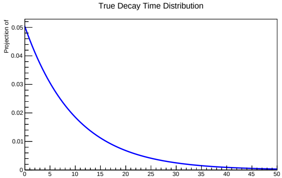
Efficiency correction

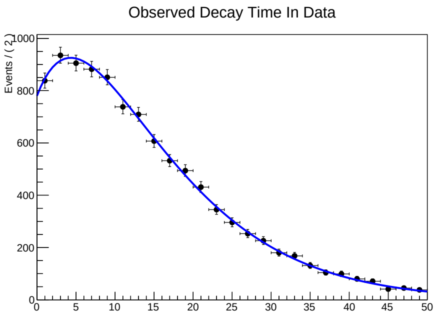
- Sometimes can be more complicated
- Not correcting a number but instead a distribution
- For example extracting decay time acceptance from MC
- Often incorporate into PDF (don't change the data)

$$\mathcal{P} = p(t) \times \epsilon(t)$$

- For the systematic assess what happens with decay time acceptance distribution







Efficiency systematic

Checks

These check for *mistakes* - do not take their difference in quadrature as a systematic

- Fit for a known lifetime in a control channel
- Tighten / loosen cuts which affect the lifetime acceptance, does this change the result
- Try and check consistency with a data driven method (e.g. "swimming" in LHCb)

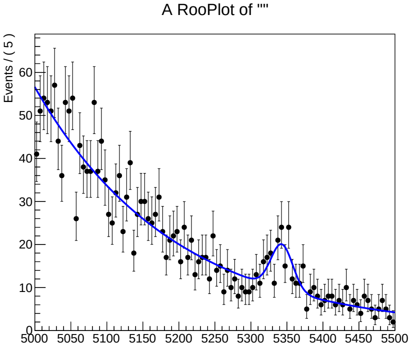
Systematics

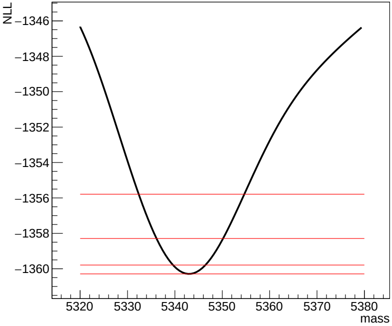
- How can my assumed shape of the acceptance be different from the truth?
- Consider a correction for data / MC differences and resulting systematic uncertainty
- What is my decay time resolution?
- Compute the outcome of their effect with toy MC
  - Generate toys from an "extreme" case, fit back with the nominal model and examine the pull
- Even better let the decay time acceptance have some freedom in the fit - very difficult

Energy scale/resolution

New particle mass measurement

- Find new resonance in $\mu\mu$ invariant mass spectra
- Want to measure its mass
- Fit data with an exponential (background) + Gaussian (signal)
- Stat only: $m = 5343.0 \pm 3.4\ \mathrm{GeV}/c^2$





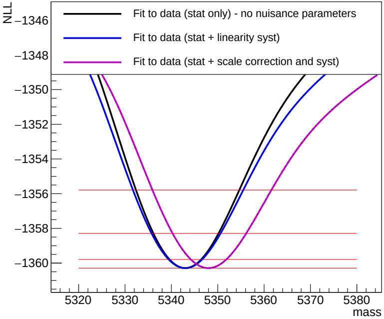
Energy scale/resolution

Systematic uncertainties

- Energy scale calibration ($J/\psi \to \mu\mu$)
- Energy scale linearity (from $m(J/\psi)$ to $m(\mathrm{newparticle})$)
- These could be propagated through to the effect on the energy of the muon to the mass of the resonance and added in quadrature
- Or add terms to the likelihood:

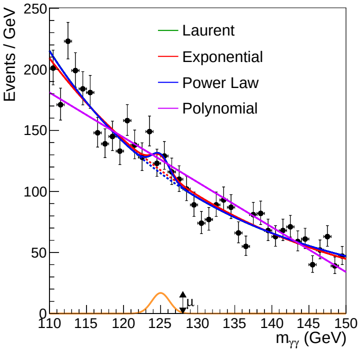
$$\mu = \mu_{fit} + G(\Delta\mu_{calib}, \sigma_{calib}) + G(0, \sigma_{linear})$$

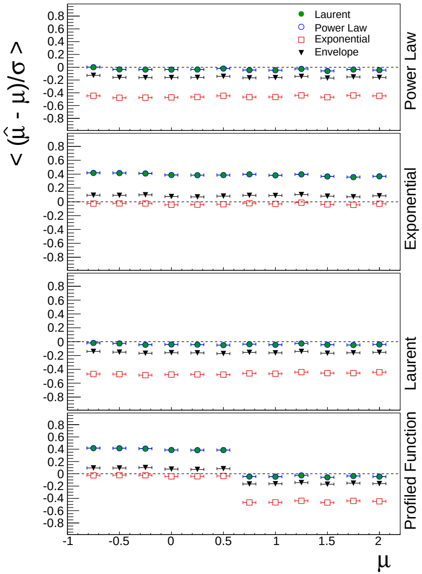
- Profile likelihood
- Uncertainty increases from more freedom in the fit
- E.g. case of energy scale and resolution uncertainty from two photons to Higgs mass



Model uncertainties

- These can often be large
- For example I have a large background which I model with an exponential
  - How do I really know it's an exponential?
- This can have a large effect on the signal yield





Model uncertainties

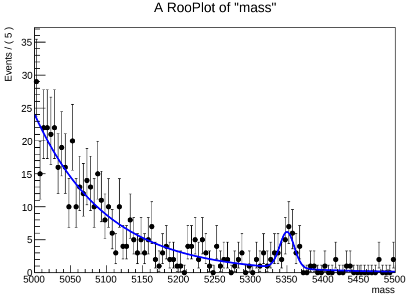
- This is something you cannot correct for
- Common thing to do is add a large systematic from the pull plots shown above
  - I.e. generate from one model and fit back with a bunch of others
  - Look at the spread in these and use this as a systematic
- However there are some alternative methods

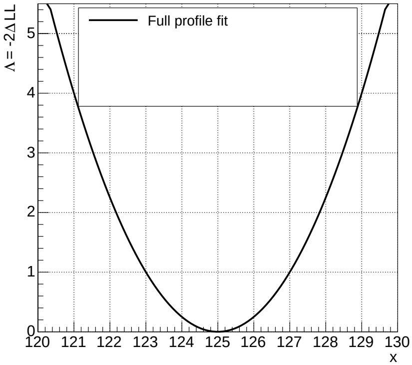
Concept of a nuisance parameter

Consider a simple situation:

- Fit a simple Gaussian signal and exponential background model to data with
  - one parameter of interest (observable) - e.g signal yield, x
  - one nuisance parameter - e.g. background exponential slope, θ
  - all other parameters fixed

1. **Scan $\Lambda = -2LL$ of parameter x whilst profiling θ**

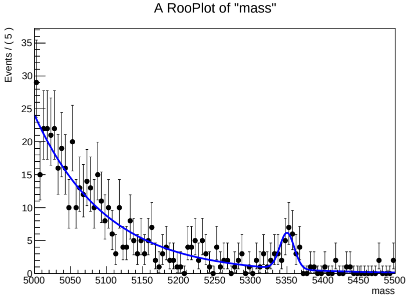


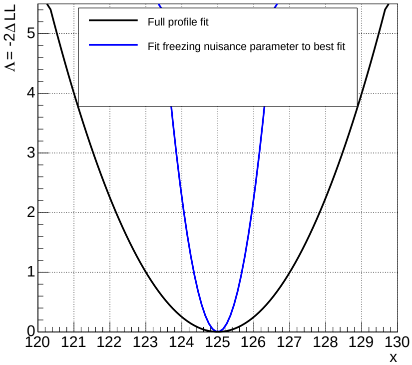


Concept of a nuisance parameter

- Now imagine the background parameter is perfectly known
  - fix nuisance parameter which now has no variation
  - equivalent to the statistical only error

2. **Fix θ to it's best fit value**
   - blue line

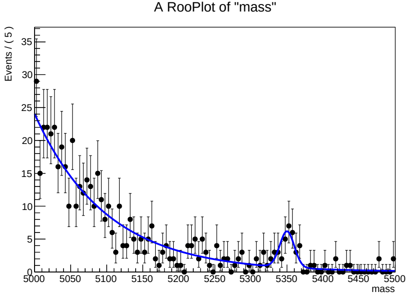


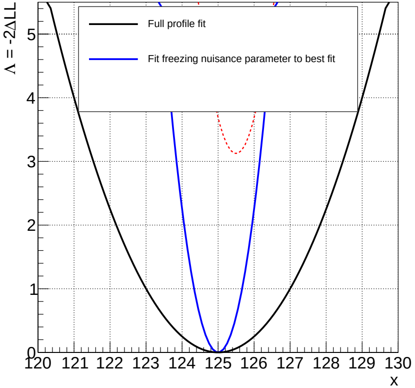


Concept of a nuisance parameter

- What about if we fix the background parameter to some other value?
  - this gives some other curve
  - not necessarily near the minimum

3. **Fix θ to a random value**
   - red dashed line

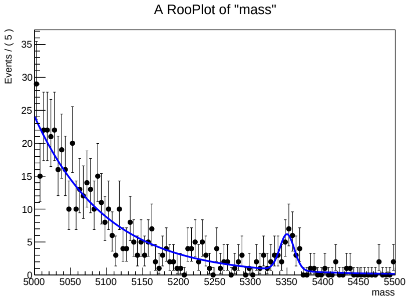


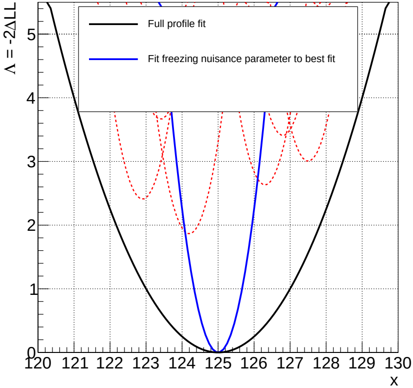


Concept of a nuisance parameter

- Can do this for a few different values of the background parameter

2. **Fix θ to a few random values**
   - red dashed lines

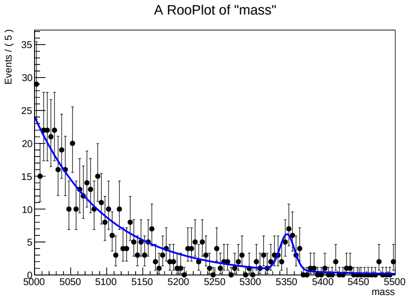


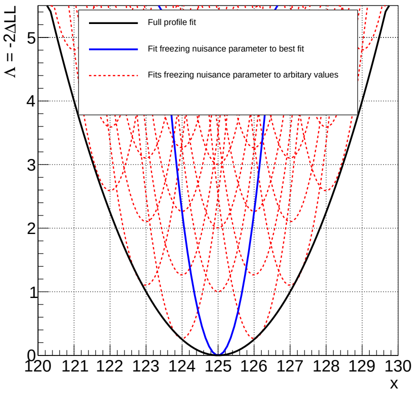


Concept of a nuisance parameter

- And even more values...

2. **Fix θ to a few random values**
   - red dashed lines

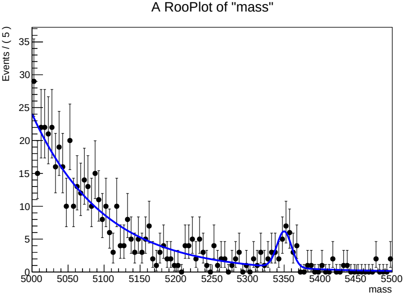


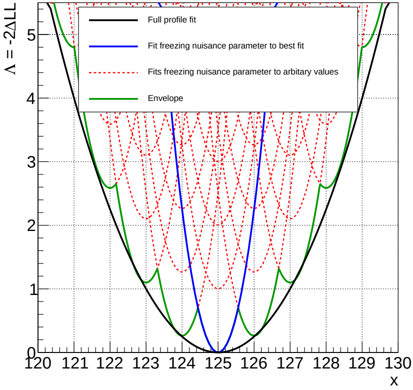


Concept of a nuisance parameter

- If you draw the minimum contour around all of the red dashed lines you begin to recover the original curve
  - In this case it doesn't matter because θ is a continuous nuisance parameter
  - But if we have a parameter that can *ONLY* take discrete values then we can make a profile likelihood in this way
    - For example we have ten different models (we can label them as having discrete value of a nuisance parameter $n = 1 - 10$)

2. **Draw minimum "envelope"**
   - green line

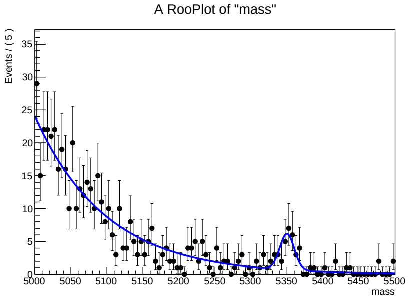


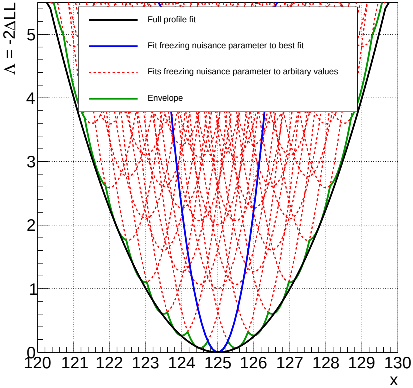


Concept of a nuisance parameter

- Clearly the more discrete values we sample the closer we get to the original

2. **Draw minimum "envelope"**
   - green line

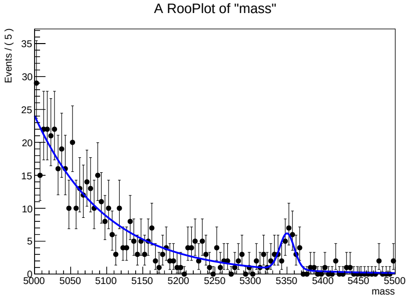


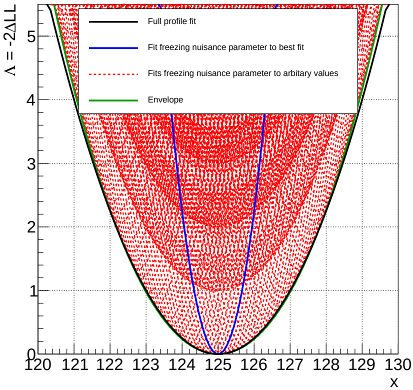


Concept of a nuisance parameter

- Clearly the more discrete values we sample the closer we get to the original
- *IMPORTANTLY - you can mix discrete nuisance parameters with continous ones*

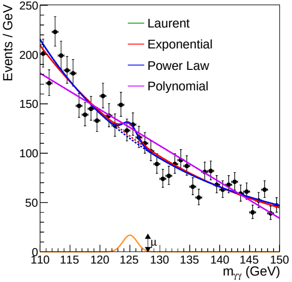
2. **Draw minimum "envelope"**
   - green line





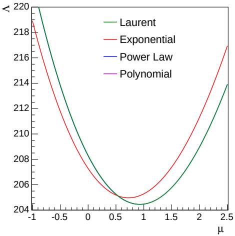
Back to the more realistic example

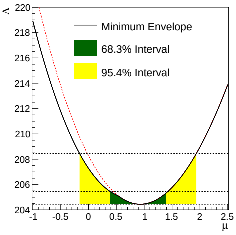
- A small signal component
- Some realistic (and one unrealistic) background models
- Do a profile scan for each model and take the envelope
  - Choices which are very similar have no effect (**Laurent** and **Power Law**)
  - Choices which are bad have no effect (**Polynomial**)
  - Choices which compete increase the uncertainty (**Exponential**)
- Uncertainty is increased if models are different
- *NOTE:* No explicit model choice has to be made
  - We don't actually care what model "is the best"



Result:

- A best fit value ✔
- A confidence interval ✔
- A systematic from the model choice ✔



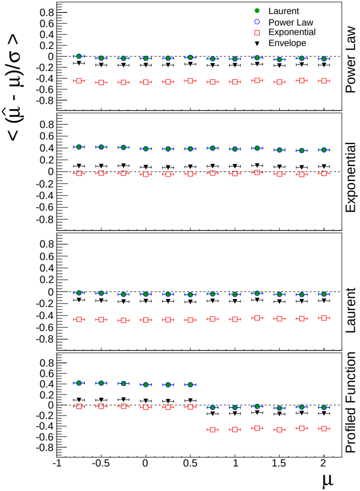


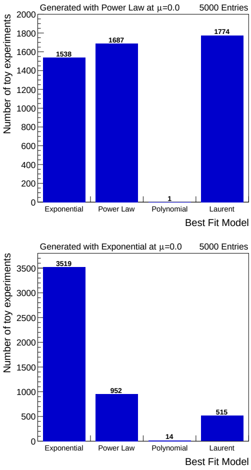
Bias and Coverage properites

Generate toy MC from various background hypotheses and then refit to asses the bias (using the pull) and the coverage

Bias:

- When you generate and fit back with *the same* (or similar) background function the bias is neglible (**green points** in top panel, **red points** in second panel)
- When you generate and fit back with *different* functions the bias is large (**red points** in top panel, **green points** in second panel)
- Using the profile envelope (**black points**) you find a small bias **for all cases**



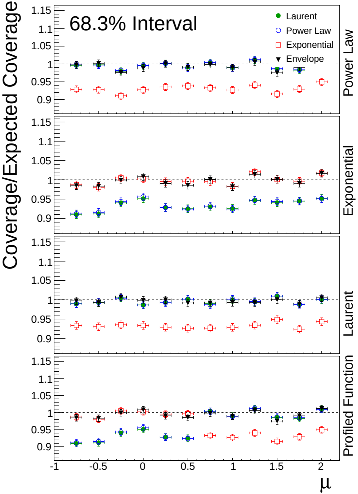


Bias and Coverage properites

Generate toy MC from various background hypotheses and then refit to asses the bias (using the pull) and the coverage

Coverage:

- When you generate and fit back with *the same* (or similar) background function the coverage is good (**green points** in top panel, **red points** in second panel)
- When you generate and fit back with *different* functions there can be under-coverage (**red points** in top panel, **green points** in second panel)
- Using the profile envelope (**black points**) you recover good coverage **for all cases**



4. Correlation, covariance and all that stuff

Correlation and covariance

In most cases we present our results as Gaussian observables

- We present our observables in the form:
  $v = XX \pm YY\ (\mathrm{stat}) \pm ZZ\ (\mathrm{syst})$
- For $n > 1$ dimensions should also provide correlation matrices:
  - Can calculate sytematic correlations using toys
  - Generate with one effect and caluclate how this correlates with another
  - Most people don't bother doing this (sad face)

Central values and uncertainties (stat and syst)

$$x_1 = v_1 \pm \sigma_1^{\mathrm{stat}} \pm \sigma_1^{\mathrm{syst}} \tag{1}$$

$$x_2 = v_2 \pm \sigma_2^{\mathrm{stat}} \pm \sigma_2^{\mathrm{syst}} \tag{2}$$

$$\ldots \tag{3}$$

$$x_n = v_n \pm \sigma_n^{\mathrm{stat}} \pm \sigma_n^{\mathrm{syst}} \tag{4}$$

Statistical correlation

$$\begin{pmatrix} 1 & \rho_{1,2} & \ldots & \rho_{1,n} \\ \rho_{2,1} & 1 & \ldots & \rho_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ \rho_{n,1} & \rho_{n,2} & \ldots & 1 \end{pmatrix}$$

Systematic correlation

$$\begin{pmatrix} 1 & \rho_{1,2} & \ldots & \rho_{1,n} \\ \rho_{2,1} & 1 & \ldots & \rho_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ \rho_{n,1} & \rho_{n,2} & \ldots & 1 \end{pmatrix}$$

Building a covariance matrix

Correlation, ρ, dictates the correlation between uncertainties.
Covariance, $\mathcal{C}$, is the full error matrix
Convert from a correlation matrix to a covariance matrix $\hat{\mathcal{C}} = \vec{\sigma}\hat{\rho}\vec{\sigma}$

$$\hat{\mathcal{C}} = \begin{pmatrix} \sigma_i^2 & \sigma_i\sigma_j\rho_{ij} \\ \sigma_i\sigma_j\rho_{ij} & \sigma_j^2 \end{pmatrix} = \underbrace{\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_j \end{pmatrix}}_{\hat{\sigma}} \underbrace{\begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix}}_{\hat{\rho}} \underbrace{\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_j \end{pmatrix}}_{\hat{\sigma}} \tag{5}$$

Convert from covariance matrix to correlation matrix $\hat{\rho} = \hat{\sigma}^{-1}\hat{\mathcal{C}}\hat{\sigma}^{-1}$

$$\hat{\rho} = \begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1/\sigma_i & 0 \\ 0 & 1/\sigma_j \end{pmatrix}}_{\hat{\sigma}^{-1}} \underbrace{\begin{pmatrix} \sigma_i^2 & \sigma_i\sigma_j\rho_{ij} \\ \sigma_i\sigma_j\rho_{ij} & \sigma_j^2 \end{pmatrix}}_{\hat{\mathcal{C}}} \underbrace{\begin{pmatrix} 1/\sigma_i & 0 \\ 0 & 1/\sigma_j \end{pmatrix}}_{\hat{\sigma}^{-1}} \tag{6}$$

Note: $\hat{\sigma}^{-1} = \mathrm{inv}\left(\sqrt{\mathrm{diag}(\hat{\mathcal{C}})}\right)$

Remember covariance matrices add linearly (not in quadrature)

Propagation of uncertainty

Assuming no correlations:

$$x = f(a, b, c)$$
$$\sigma_x = \sqrt{\left(\frac{\partial f}{\partial a}\right)^2 \sigma_a^2 + \left(\frac{\partial f}{\partial b}\right)^2 \sigma_b^2 + \left(\frac{\partial f}{\partial c}\right)^2 \sigma_c^2} \tag{7}$$

For two correlated variables:

$$x = f(a, b)$$
$$\sigma_x = \sqrt{\left(\frac{\partial f}{\partial a}\right)^2 \sigma_a^2 + \left(\frac{\partial f}{\partial b}\right)^2 \sigma_b^2 + 2\frac{\partial f}{\partial a}\frac{\partial f}{\partial b}\sigma_a \sigma_b \rho_{ab}}$$

In general....

Transforming to another variable

Completely general

- Have a vector of observables, $\vec{A}$, and a vector of functions which depend on the elements of $\vec{A}$, $\vec{X}(\vec{A})$.
- Uncertainties on the vector $\vec{A}$ are described by the covariance matrix $\mathcal{C}_A$.
- We would like to know the covariance matrix after transformation by $\vec{X}$.
- The Jacobian matrix will transform the errors (to first order) following a Taylor expansion.

$$\vec{X}(\vec{A}+\delta\vec{A}) = \vec{X}(\vec{A}) + \left.\frac{\partial\vec{X}}{\partial\vec{A}}\right|_{\vec{A}} \delta\vec{A} + \mathcal{O}(\delta\vec{A}^2)$$

$\frac{\partial\vec{X}}{\partial\vec{A}}$ is simply the Jacobian matrix, $\mathcal{J}$.

- To transform from the covariance in $\vec{A}$ to the covariance in $\vec{X}$ then transform via the Jacobian:

$$\mathcal{C}_{\vec{X}} = \mathcal{J}\mathcal{C}_{\vec{A}}\mathcal{J}^{\mathrm{T}}$$

Transforming to another variable

- General formula:

$$\mathcal{C}_{\vec{X}} = \mathcal{J}\mathcal{C}_{\vec{A}}\mathcal{J}^{\mathrm{T}}, \text{ where } \mathcal{J} = \left.\frac{\partial \vec{X}}{\partial \vec{A}}\right|_{\vec{A}}$$

- For two variables:

$$x = x(a, b)$$
$$y = y(a, b)$$
$$\mathcal{C}_{xy} = \begin{pmatrix} \frac{\partial x}{\partial a} & \frac{\partial x}{\partial b} \\ \frac{\partial y}{\partial a} & \frac{\partial y}{\partial b} \end{pmatrix} \begin{pmatrix} \sigma_a^2 & \sigma_a\sigma_b\rho_{ab} \\ \sigma_a\sigma_b\rho_{ab} & \sigma_b^2 \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial a} & \frac{\partial y}{\partial a} \\ \frac{\partial x}{\partial b} & \frac{\partial y}{\partial b} \end{pmatrix}$$

Transforming to another variable

Another example

- We measure two observables a, b with σ_a, σ_b and ρ_{ab}
- Transform to $x = a(1+b)$ and $y = a(1-b)$
- Know that:

$$\mathcal{J} = \begin{pmatrix} \frac{\partial x}{\partial a} & \frac{\partial x}{\partial b} \\ \frac{\partial y}{\partial a} & \frac{\partial y}{\partial b} \end{pmatrix} = \begin{pmatrix} 1+b & a \\ 1-b & -a \end{pmatrix}, \quad \mathcal{C}_{ab} = \begin{pmatrix} \sigma_a^2 & \sigma_a\sigma_b\rho_{ab} \\ \sigma_a\sigma_b\rho_{ab} & \sigma_b^2 \end{pmatrix}$$

- And so:

$$\mathcal{C}_{xy} = \begin{pmatrix} 1+b & a \\ 1-b & -a \end{pmatrix} \begin{pmatrix} \sigma_a^2 & \sigma_a\sigma_b\rho_{ab} \\ \sigma_a\sigma_b\rho_{ab} & \sigma_b^2 \end{pmatrix} \begin{pmatrix} 1+b & 1-b \\ a & -a \end{pmatrix}$$

5. Summary

Summary

- Systematics are difficult
- There is no common approach so it takes some common sense and experience
- There is often confusion between:
  1. A systematic effect
  2. A systematic uncertainty
  3. A check for consistency
- The most robust ways of evaluting systematic are by generating toys
- Incorporate nuisance parameters into your likelihood where possible
  - This allows the data the freedom to pick
- Consider carefully how you evaluate uncertainties from model choices you might make