# Decomposition of Stat/Sys Errors

ATLAS Statistics Forum
CERN, 12 September 2017

Glen Cowan
Physics Department
Royal Holloway, University of London
**www.pp.rhul.ac.uk/~cowan**
**g.cowan@rhul.ac.uk**

# Background info

Recall prototypical analysis:

primary data $\boldsymbol{x}$,

parameter of interest $\mu$,

nuisance parameter(s) $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$,

control measurements $\boldsymbol{y}$,

model $P(\boldsymbol{x}, \boldsymbol{y} \,|\, \mu, \boldsymbol{\theta}) = L(\mu, \boldsymbol{\theta})$ (the likelihood).

Often control measurements are designed to constrain a particular nuisance parameter, e.g., $y_i = \tilde{\theta}_i$, could be "best guess" of $\theta_i$, but treated as a measurement with a sampling distribution $p(\tilde{\theta}_i|\theta_i)$.

Maximize the likelihood $\rightarrow \hat{\mu}$

Variance of $\hat{\mu}$ reflects total uncertainty, i.e., the model *with* nuisance parameters is "correct", no systematic uncertainty.

# Commonly used method

1) Identify source of systematic with nuisance parameter $\theta$.
2) Fix $\theta = \theta_0$
3) Repeat fit, get $\hat{\mu}_{\theta 0}$
4) Get variance $V[\hat{\mu}_{\theta 0}]$
5) $\sigma_{\mathrm{sys},\theta} = (V[\hat{\mu}] - V[\hat{\mu}_{\theta 0}])^{1/2}$

But what about nuisance parameters that we expect to be (at least partially) constrained by the data, e.g, background level/shape?

At least some portion of the uncertainty in such nuisance parameters is more logically regarded as a statistical error.

# Alternative approach / example

Goal (?) of stat/sys breakdown is to communicate how the uncertainty is expected to scale with luminosity, so, define ratio of lumi to that of actual measurement

$$\lambda = \mathcal{L}/\mathcal{L}_0$$

and rewrite model so as to include lambda. E.g.,

$$x \sim \mathrm{Gauss}(\lambda(\mu + \beta)e^{\theta}, \sqrt{\lambda}\sigma_x)$$

$$y \sim \mathrm{Gauss}(\lambda\beta, \sqrt{\lambda}\sigma_y)$$

$$z \sim \mathrm{Gauss}(\theta, \sigma_z)$$

Here $\mu$ is parameter of interest (~signal rate); $\beta$ (~background rate) and $\theta$ (scale factor) are nuisance parameters.

$x$ is "main" measurement; $y$ and $z$ are control measurements.

# Example (2)

Likelihood = product of 3 Gaussians → −2ln $L$ gives

$$\chi^2(\mu, \beta, \theta) = \frac{(x - \lambda(\mu + \beta)e^{\theta})^2}{\lambda\sigma_x^2} + \frac{(y - \lambda\beta)^2}{\lambda\sigma_y^2} + \frac{(z - \theta)^2}{\sigma_z^2}$$

Minimizing $\chi^2$ gives estimators

$$\hat{\mu} = \frac{1}{\lambda}(xe^{-z} - y) \qquad \hat{\beta} = y/\lambda \qquad \hat{\theta} = z$$

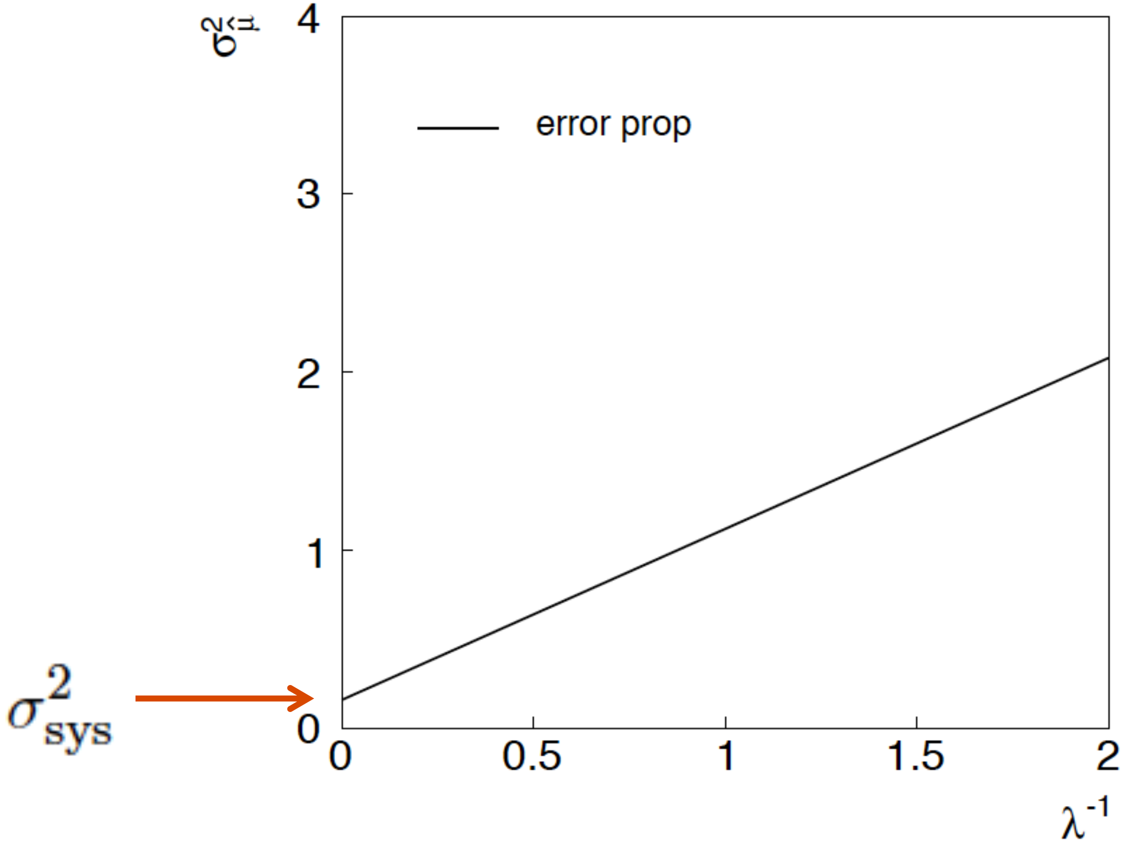
Linear error propagation gives the variance

$$V[\hat{\mu}] = \frac{1}{\lambda}\left(e^{-2\theta}\sigma_x^2 + \sigma_y^2\right) + (\mu + \beta)^2\sigma_z^2$$

stat sys

# Result of example (constant $\sigma_z$)

Plot variance versus $\lambda^{-1}$, intercept at zero (infinite lumi) corresponds to systematic error:



# Variation on example

But suppose the std. deviation of control measurement $z$ had been modeled as

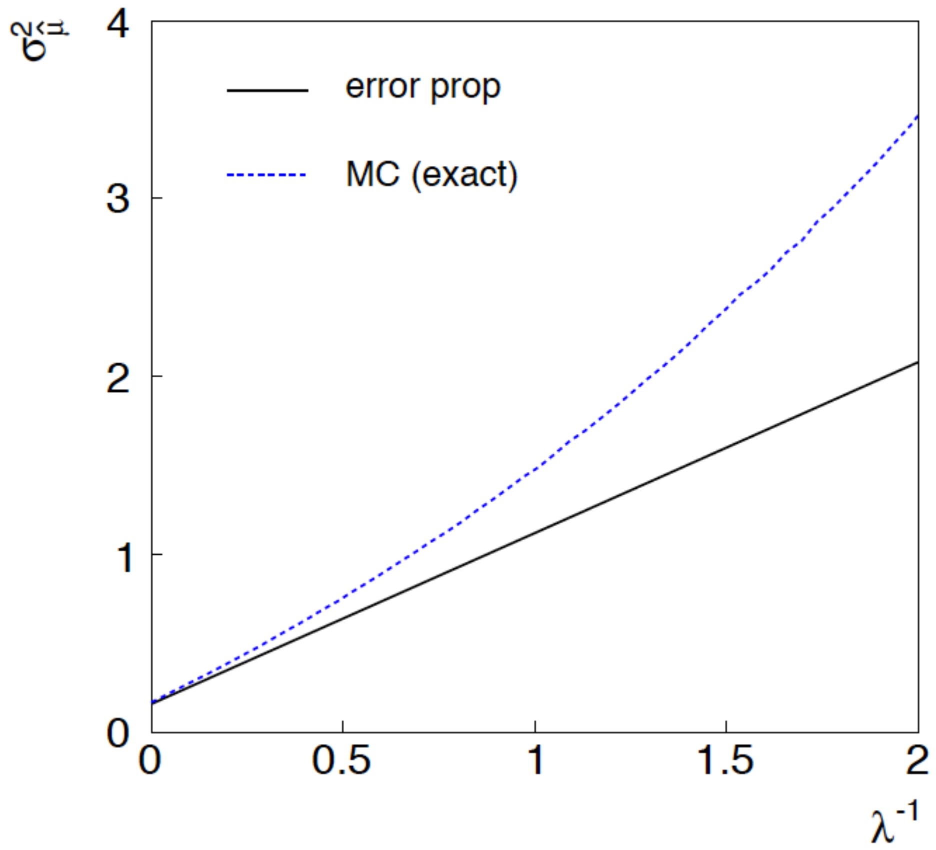
$$\sigma_z = \sqrt{\sigma_{z0}^2 + \frac{\sigma_{z1}^2}{\lambda}}$$

Here $\sigma_{z0}$ will contribute to the part that does not change with $\lambda$, (systematic error), $\sigma_{z1}$ to part that goes as $1/\lambda$ (stat error).

But in the first ("commonly used") method, fixing $\theta$ would in in effect treat both as part of the systematic uncertainty.

# MC (exact) determination of variance

Estimator for $\mu$ nonlinear in $z$, so error propagation not exact; use MC to get variance:



# Extrapolate or not?

Behaviour in region near nominal lumi may seem like reasonable basis for stat/sys decomposition, but may not give meaningful extrapolation to infinite lumi: