

Brief Article

The Author

August 26, 2019

1 Types of Errors

There are two main types of errors, random and systematic. Random errors, also referred to as statistical errors, are unavoidable and arise from the randomized statistical fluctuations in all measurements. These can be reduced by repeating the measurements over and over a large number of times. In many cases, this is not realistically possible which is why Monte Carlo simulations are handy. For the case of systematic errors, repeating the experiment many times or employing simulations does nothing to reduce this type of error. Systematic errors, also often referred to as bias, are caused by imperfections related to the experiment's conditions; such as setup, tools, and procedures used by the analyst. The individual choices that each analyst makes will alter the conditions and lead to an inconsistent estimate of the measurand. In our work, the measurand is the Isotropic Equivalent Energy (E_{iso}) of Gamma-ray Bursts (GRBs) and the type of error we focus on is systematic. In this dissertation we identify multiple sources of systematic errors that may lead to different E_{iso} energies for the same GRB. We discuss a wide breadth of possible sources of error in Chapter X, although we do not necessarily calculate them all. Systematic uncertainties are routinely omitted from most GRB studies and the majority of errors published (on both model parameters and energetics) are purely statistical in nature. We focus on identifying and quantifying several sources of systematic uncertainties caused by choices an analyst would make while studying the time-integrated ("fluence") spectrum of GRBs. We discuss the relevance of the individual components and how they depend on the GRBs in the sample. We then combine the components to provide a profile for the overall uncertainty in E_{iso} . Any uncertainty associated with this energetic will be transferred to the Amati and Ghirlanda correlations which adopt it for probing cosmological distances and pseudo-redshift estimations of other GRBs.

2 Identifying the Sources of Systematic Uncertainty

In this dissertation, we primarily investigate the affects of the analyst's choices made while modeling the fluence spectrum. Some of these are blah blah blah. We discuss these sources

of error in detail in Chapter X.

3 Propagation of Error

3.1 Addition or Subtraction

Suppose you have several sources of uncertainty (σ_a , σ_b , σ_c , and σ_d) and you want to find the total uncertainty (σ_{Total}).

3.2 Multiplication or Division

3.3 Raising to a Power

4 Calculating the Combined Uncertainty

Combined standard uncertainty ($u_c(y)$) is found by the following equation:

$$u_c(y(x_1, x_2, \dots)) = \sqrt{\sum_{i=1,n} c_i^2 u(x_i)^2} = \sqrt{\sum_{i=1,n} u(y, x_i)^2}^* \quad (1)$$

where $y(x_1, x_2, \dots)$ is a function of several parameters x_1, x_2, \dots , c_i is a sensitivity coefficient evaluated as $c_i = \partial y / \partial x_i$, the partial derivative of y with respect to x_i and $u(y, x_i)$ denotes the uncertainty in y arising from the uncertainty in x_i . We will now rewrite this equation in a more familiar light;

$$\sigma_{f(E)} = \sqrt{\sum_{i=1,n} \left(\frac{\partial f}{\partial \theta_i} \right)^2 \sigma_{\theta_i}^2} \quad (2)$$

Equations 1 and 2 are identical, we simply replaced several variables so that the equation is more familiar to us. This is the first part of the propagation of uncertainty equation for our spectral modeling functions; the part that does not rely on a covariance matrix. In equation 2, $f(E)$ could represent the *Band* function and θ_i would then represent the α , β , E_{pk} , and N parameters. Each variable's contribution $u(y, x_i)$ is just the square of the associated uncertainty expressed as a standard deviation multiplied by the square of the relevant sensitivity coefficient. This sensitivity coefficient (partial derivatives) describe how the value of the function (y) varies with changes in the parameters (x_i). When the x_i (or θ_i) variables are not independent (i.e., they are covariant), then we add on an additional term to equations 1 and 2;

$$u_c(y(x_1, x_2, \dots)) = \sqrt{\sum_{i=1,n} c_i^2 u(x_i)^2} + \sqrt{\sum_{\substack{i,k=1,n \\ i \neq k}} c_i c_k \cdot u(x_i, x_k)} \quad (3)$$

and

$$\sigma_{f(E)} = \sqrt{\sum_{i=1,n} \left(\frac{\partial f}{\partial \theta_i} \right)^2 \sigma_{\theta_i}^2} + \sqrt{\sum_{\substack{i,k=1,n \\ i \neq k}} \left(\frac{\partial f}{\partial \theta_i} \right) \left(\frac{\partial f}{\partial \theta_k} \right) \cdot \sigma_{\theta_i} \sigma_{\theta_k}} \quad . \quad (4)$$

If we expand the second summation in this equation and combine like terms (i.e., there will be an ik and a ki), the terms having covariance are all preceded by a factor of 2.

When an uncertainty contribution is expressed for a procedure, it is expressed as having some impact on the final result. In this case, the sensitivity coefficient (partial derivative) is equal to 1;

$$\sigma_{f(E)} = \sqrt{\sum_{i=1,n} \sigma_{\theta_i}^2} + \sqrt{\sum_{\substack{i,k=1,n \\ i \neq k}} \sigma_{\theta_i} \sigma_{\theta_k}} \quad . \quad (5)$$

This is the form we want since we are expressing a combined uncertainty of an analyst's method; thus θ_i are not longer model parameter uncertainties, but instead an uncertainty in a procedure (P_i). A list of these may be: choice of spectral model, choice of modeling program, choice of data to use, etc. Since the idea behind combining procedural uncertainties is that they should not have covariance between them, we can drop the second term in equation 5 and we are left with a sum of squares of the procedural uncertainties:

$$\sigma_{Total} = \sqrt{\sum_{i=1,n} \sigma_{P_i}^2} \quad . \quad (6)$$

This is how we get to the equation for combining uncertainties of multiple sources of uncertainties on a measurand caused by analyst decisions from the propagation of uncertainty equation.

4.1 Combining Uncertainties Procedures

In certain cases, combining uncertainties is reduced to a simple form. These cases are for addition and subtraction,

Our goal is to find the combined uncertainty on the measurement of E_{iso} . In order to do that, we have to first identify the individual sources of uncertainties and then quantify them.

Uncertainty evaluated from the dispersion of repeated measurements, whether it be your own or several published values.

5 Relevance to Collazzi et al. (2011), The Total Errors in Measuring E_{pk} for Gamma-ray Bursts

We used this paper as a guide for our work on E_{iso} .

5.1 Multiple published values for the same GRB

In section 3 of Collazzi et al. (2011), they take four different values of E_{pk} for the same GRB event (GRB 910503) that were published by different authors. These E_{pk} energies are 465 keV, 741 keV, 621 keV, and 586 keV. We want to find the uncertainty (or “scatter”) between these four energies. To do this we use equation 1.9 from Bevington’s “Data Reduction and Error Analysis for Physical Sciences” book;

$$s = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2}, \quad (7)$$

where \bar{x} is the mean of the four energies, x_i is each i th energy in the list, and N is the total number of energies ($N = 4$ in this case). In Python code, equation 7 would be

```
import numpy as np
def standard_deviation(values):
    N = len(values)
    mean_value = np.mean(values)
    sum_term = sum([(i-mean_value)**2 for i in values])
    return np.sqrt(sum_term/(N-1))
```

The Python numpy standard deviation function will match this one if we use `ddof=1` (`np.std(values, ddof=1)`), where `ddof` stands for delta degrees of freedom. The equation and code above only allows for `ddof=1` since $N - 1$, when they should realistically be set up with $N - \text{ddof}$ instead (how numpy’s is) in order to accept multiple delta degrees of freedom. When using `np.std(values)`, the default is `ddof=0`, which would use N as the denominator term. The text states that the factor of $N - 1$ (or `ddof=1`) is due to the fact that \bar{x} (the mean) is determined from the data and not independently. One important thing to point out is that equation 7, using `ddof=1`, is equivalent to the root-mean-square over root 2; 7, using `ddof=1`, is equivalent to the root-mean-square over root 2;

$$\sigma_{Choice} = \frac{\sigma_{rms}}{\sqrt{2}} \quad (8)$$

where

$$\sigma_{rms} = \sqrt{\frac{\sum_{i=1}^N (\Delta_i^2)}{N}} \quad \text{and where} \quad \Delta_i = \log_{10}(E_{pk,1}) - \log_{10}(E_{pk,2}),$$

which is the equation that Collazzi et al. (2011) uses in his examples. The Δ_i are the logged differences between the values. When there are four values for the same GRB, the

list of Δ_i will be the differences between each value in the list against each of the others. Thus, we'd have $\Delta_i = [A-B, A-C, A-D, B-C, B-D, C-D]$ if the list of values was $[A, B, C, D]$.

these are the differences between each individual value and the with respect to each of the others in the list.

requires the energies to first be logged, giving us 2.668 keV, 2.87 keV, 2.793 keV, and 2.768 keV, respectively. Using

In section 3 of their paper, they take four different published values of E_{pk} (465 keV, 741 keV, 621 keV, and 586 keV) for the same single GRB (GRB 910503) and report a rms scatter of $\sigma_{Choice} = 0.083$. The list of E_{pk} energies must first be logged, so these values become (2.668, 2.87, 2.793, 2.768). This value of $\sigma_{Choice} = 0.083$ was found through the following equations;

$$\sigma_{f(E)} = \sqrt{\sum_{i=1,n} \sigma_{\theta_i}^2} + \sqrt{\sum_{\substack{i,k=1,n \\ i \neq k}} \sigma_{\theta_i} \sigma_{\theta_k}} \quad . \quad (9)$$

6 Appendix

6.1 Definitions

quadrature - A fancy word for the square root of the sum of squares; $\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_n^2}$.
“Errors add in quadrature.”

Average Deviation - absolute values of the deviations. A measure of the dispersion of the expected observations about the mean.

Standard Deviation -
Variance -

6.2 Equations

6.2.1 Sum of Squares

$$\sigma_{rms} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^n \sigma_i^2}{n}} \quad (10)$$

`sum([i**2 for i in values])`

6.2.2 Root Mean Square

Also known by the names:

- quadratic mean

•

$$\sigma_{rms} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^n \sigma_i^2}{n}} \quad (11)$$

```
import numpy as np
np.sqrt( sum([i**2 for i in values]) / len(values) )
```