Classification and Distributional properties of Gamma Ray Bursts

Shinjon Chakraborty & Prithwish Ghosh
University of Calcutta & Visva Bharati University
Department of Statistics
email: shinjonchakraborty07@gmail.com,
ghosh.prithwish1999@gmail.com

Plan

Plan of the presentation :

1. Introduction
2. Gamma Ray Burst
3. How statistics helps us to classify GRB
4. Conclusion
5. References

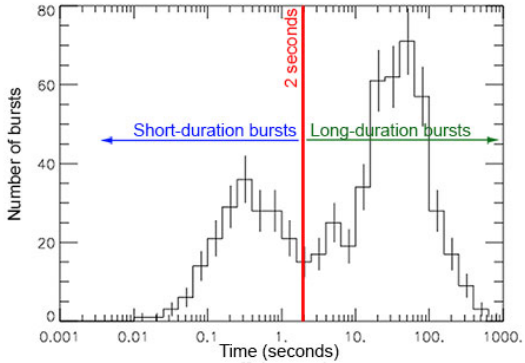
Introduction

In gamma-ray astronomy, gamma-ray bursts (GRBs) are immensely energetic explosions that have been observed in distant galaxies. They are the most energetic and luminous electromagnetic events since the Big Bang. Bursts can last from ten milliseconds to several hours. After an initial flash of gamma rays, a longer-lived ”afterglow” is usually emitted at longer wavelengths (X-ray, ultraviolet, optical, infrared, microwave and radio). Gamma Ray Burst can be classified into two parts which are given below:

- Short GRB
- Long GRB

Introduction(Cont..)

These two classes are likely created by different processes, but the end result in both cases is a brand new black hole. This is the Graph of the time versus number of bursts for the gamma-ray bursts observed by the BATSE instrument on the Compton Gamma-ray Telescope.



Short Gamma Ray Burst

Events with a duration of less than about two seconds are classified as short gamma-ray bursts. These account for about 30% of gamma-ray bursts, but until 2005, no afterglow had been successfully detected from any short event and little was known about their origins. Since then, several dozen short gamma-ray burst afterglows have been detected and localized, several of which are associated with regions of little or no star formation, such as large elliptical galaxies. This rules out a link to massive stars, confirming that short events are physically distinct from long events.

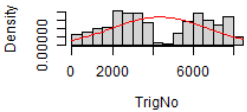
Long Gamma Ray Burst

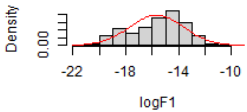
Most observed events (70%) have a duration of greater than two seconds and are classified as long gamma-ray bursts. Because these events constitute the majority of the population and because they tend to have the brightest afterglows, they have been observed in much greater detail than their short counterparts. Almost every well-studied long gamma-ray burst has been linked to a galaxy with rapid star formation, and in many cases to a core-collapse supernova as well, unambiguously associating long GRB's with the deaths of massive stars.

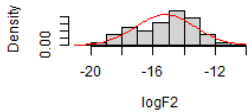
Insight of the Data Set

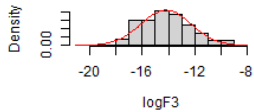
The primary source of astronomical data on GRB is collected from the Burst and Transient Source Experiment(BATSE) catalogue which is associated to the COMPTON Gamma-Ray Observatory(CGRO) mission. The data set is of dimension 1963 x 13 i.e it has information on 1963 GRB's quantified by the 13 parameters which are listed below. **TrigNo**, $\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{F}_3$, $\mathbf{F}_4$, **T50**, **T90**, **F64ms**, **F256ms**, **F1024ms**, $\mathbf{H}_{32}$, $\mathbf{H}_{321}$, $\mathbf{F}_t$
All the details about the parameters will be discussed during the presentation.

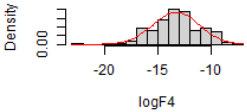
Exploratory data Analysis

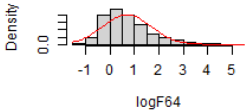


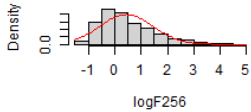


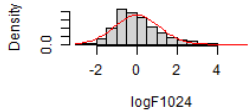


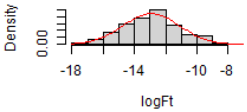


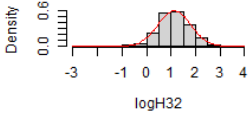


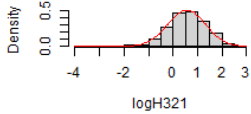


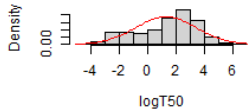












Exploratory Data analysis(cont..)

We plotted the histograms for every parameter in consideration for the study. The histograms along with a reference Normal density plot are shown in Figure . We made a figure set containing all figures—a total of 13 figures. Readers can see all the figures in the figure set accompanying figure .

We also checked the equivalence of distributional properties of the parameters through the non-parametric test Kolmogorov-Smirnov and also to verify the normality of the parameters, which is an important assumption for further parametric inferences, we applied the Shapiro-Wilks Test for Normality.

Exploratory Data analysis(cont..)

	logFt	logH32	logH321	logT50	logT90	logF64
logFt	1					
logH32	-0.04	1				
logH321	-0.05	0.94	1			
logT50	0.55	-0.39	-0.41	1		
logT90	0.59	-0.40	-0.43	0.94	1	
logF64	0.43	0.23	0.25	-0.17	-0.11	1

Table – Correlation Matrix

Exploratory Data analysis(cont..)

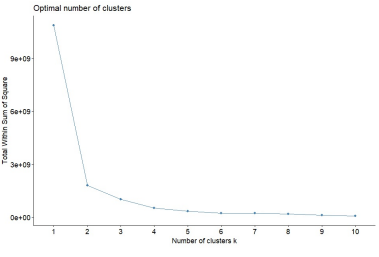
Above we have incorporated the Pearson's Correlation Measure to compute the correlation between the Concerned parameters of the data-set and the correlation matrix is presented in form of above Table. The Pearson's Correlation measure is defined as-

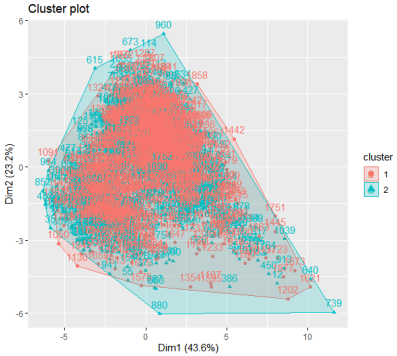
$$\gamma = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2}}$$

From the correlation matrix we can infer that, since there is a high positive correlation between T50 and T90, the time interval in which 90% burst count is observed is dependent on the time interval in which 50% burst count is observed.

Statistical Analysis(Kmeans)

Clustering Algorithm k-means clustering is a technique which aims to find more homogeneous subgroups within the data. The idea is to divide the GRB data into k distinct groups so that observations within a group are similar. Now with the help of the distortion curve or mostly known as elbow curve we can conclude that how many clusters are there in the data.





Statistical Analysis(Kmeans Cont..)

The K-means clustering visualises the presence of the two distinctive clusters with cluster sizes 907 and 1055. Figure shows the formation of two clusters within the dataset

	Cluster 1	Cluster 2
$\log F_t$	-12.86583	-12.92351
$\log F_1$	-15.39730	-15.49397
$\log F_2$	-15.20724	-15.30785
$\log F_3$	-14.11814	-14.15698
$\log F_4$	-14.02948	-14.03668
$\log F_{256}$	0.4874346	0.4149041
$\log F_{1024}$	0.09169627	0.01075617
$\log H_{32}$	1.137909	1.141418
$\log H_{321}$	0.5100424	0.6020447
logT50	1.3893963	1.224692
logT90	2.529343	2.204098

Table – Cluster Means

Classification

Next to get an idea of the classification of the data we perform the *Linear Discriminant Analysis*. The results of the discriminant analysis are tabulated.

	predicted Cluster 1	predicted Cluster 2
Actual Cluster 1	407	270
Actual Cluster 2	500	785
Total	907	1055

Table – Discriminant Analysis

Why we are studying this?

Earth's atmosphere is very effective at absorbing high energy electromagnetic radiation such as x-rays and gamma rays, so these types of radiation would not reach any dangerous levels at the surface during the burst event itself. GRB within a few kiloparsecs, with its energy directed towards Earth, will mostly damage life by raising the UV levels during the burst itself and for a few years thereafter. Models show that the destructive effects of this increase can cause up to 16 times the normal levels of DNA damage. It has proved difficult to assess a reliable evaluation of the consequences of this on the terrestrial ecosystem, because of the uncertainty in biological field and laboratory data.

Conclusion

- Some of the parameters follow normality
- The time intervals in which 90% of the burst counts are detected is highly positive correlated with the time intervals in which 50% of the burst counts are detected.
- From the Kolmogorv - Smirnov non-parametric test it is evident that the distribution of the time intervals for the percentage of the burst count changes with time
- From the partitioning(kmeans) method we can evidently confirm from the data the presence of long burst and short burst as the primary types of Gamma Ray Burst.
- Also from the Linear Discriminant Analysis we can conclude that there about 43% misclassification in the clusters.

References

- Feifei Wang, Yuan-Chuan Zou, Fuxiang Liu, Bin Liao, Yu Liu, Yating Chai, and Lei Xia, "A Comprehensive Statistical Study of Gamma-Ray Bursts"
- Soma Mukherjee, Eric D. Feigelson, Gutti Jogesh Babu, Fionn Murtagh, Chris Fraley and Adrian Raftery, "Three types of gamma ray bursts"
- Tanuka chattopadhyay, Ranjeev misra, Asis kumar Chattopadhyay and Malay naskar, "Statistical evidence for three classes of gamma-ray bursts"
- Asis Kumar Chattopadhyay, Tanuka Chattopadhyay, "Statistical Methods for Astronomical Data Analysis"
- Luis Manuel Sarro, Laurent Eyer William O'Mullane, Joris De Ridder, "Astrostatistics and Data Mining"
- Joseph M. Hilbe, "Astrostatistical Challenges for the New Astronomy"

References

- Baidyanath Basu, Tanuka Chattopadhyay, Sudhindra Nath Biswas,"An introduction to Astrophysics"
- W J Azzam, H Zitouni and N Guessoum, "Gamma-Ray Bursts: Characteristics and Prospects"
- Liang Zhang, Juan-Juan Luo, Yong-Feng Huang, Yu-Jun Gong, Sheng Wu, "Distribution of Gamma-Ray Bursts on the T90 plane and their classification revisited"
- Pavel Minaev , Alexei Pozanenko,"Classification Problem and Parameter Estimating of Gamma-Ray Bursts"