Testing time variability of gamma-ray flux

DALIBOR NOSEK¹, STANISLAV STEFANÍK¹, JANA NOSKOVÁ².

¹ Institute of Particle and Nuclear Physics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic
² Department of Mathematics, Faculty of Civil Engineering, Czech Technical University, Prague, Czech Republic

stefanik_stanislav@yahoo.com

Abstract: A way of examining a hypothetical non-zero γ-ray signal for the time changes is presented. The time variability of the recently observed γ-ray source PKS 2155–304 is discussed. Several measurements were found to be excessive or deficient with large significances on time scales of months and days.

Keywords: on–off method, γ-rays, time variability.

1 Introduction

The asymptotic Li–Ma technique [1] is traditionally used in γ-astronomy to confirm positive results in searching for discrete sources. In high energy physics, different signal-to-background methods or more advanced techniques based, for example, on Bayesian reasoning are applied to distinguish signal from background [2]. To our knowledge, no simple procedure aiming at testing the change in a beforehand proven source activity has been established on statistical grounds, however.

In the following, we describe briefly the on–off problem and introduce two different statistical measures for refusing no–source or no–sink hypotheses, binomial and Li–Ma significances. As a main goal, we present a modified on–off method to test the change of a given non–zero source activity. This method is applied to demonstrate the time variability of the γ-ray flux observed in the direction of the source PKS 2155–304 on different time scales [3, 4].

2 The on-off method

The on–off analysis is widely used in many branches of physics, especially in γ-ray astronomy or in high energy physics, see e.g. Refs. [1, 2]. In γ-ray astronomy, the on–off problem is communicated in the following way. Consider an observation of events in a given time aiming to detect their source in the on–source region on the sky the choice of which is motivated by previous observations or by other arguments. Their on–source excess is judged by comparing the number of events with arrival directions pointing to the on–source region, N_on, to the number of events with arrival directions in a control off–source region, N_off, where no signal events are expected. Both these observations are treated as Poissonian random variables. The expected ratio of the count numbers in these regions is assumed to be known provided no source is present in the on–source region. This ratio, the on–off parameter, is simply α = w_on / w_off where w_on and w_off are given weights of the on– and off–source regions, respectively [3]. In this sense, the on–off problem consists in constructing a hypothesis test for the ratio of two unknown Poissonian parameters, μ_on and μ_off, assumed to give observed on– and off–source counts, respectively. The null hypothesis of no source in the on–source region, expressed by the equality μ_on = αμ_off, is then tested against a one–sided alternative of an excess of on–source counts, i.e. μ_on > αμ_off. This test is easily modified to judge a deficit of events in a region where their sink is expected.

Let us assume in the following that the source is ascertained with a given activity such that μ_on = αβμ_off where a source parameter β known in advance describes the excess of on–source counts above off–source ones. Then a question arises whether follow–up counts do agree with previous observations or not. In this case, we propose a test of the null hypothesis expressed by μ_on = βαμ_off against a two–sided alternative of an excess or deficit of on–source counts, i.e. μ_on > βαμ_off or μ_on < βαμ_off.

A suggested method is described in Appendix A. A simple binomial treatment of the on–off problem relying on the aforementioned non–zero source conjecture is described in Appendix A.1. A modified asymptotic Li–Ma significance that may be used to express deviation of observed counts from a given non–zero source is introduced in Appendix A.2. We show that within a classical statistical treatment reasonable modifications of the standard test statistics can be deduced for a predefined non–zero source activity represented by a parameter β. Resultant asymptotic significances are given by a transformation when α → αβ is replaced in the standard significance formulas [1]. Further details of our derivation and its interesting features will be discussed elsewhere.

3 Time variability

3.1 Monte Carlo simulations

The statistical on–off tests were applied to the data simulated within the Monte Carlo (MC) method. For this purpose, we generated 1000 Poissonian events separately in on– and off–regions of equal weights w_on = w_off giving α = 1. The reliability of our method was intentionally demonstrated on statistics based on small numbers. The on– and off–source means and intensities were μ_on = λ_on = 6.3 and μ_off = λ_off = 2.1, respectively.

1. Usually, a time process of observation where events occur continuously and independently of one another in time is considered. In such a case, the weights are simply given by observational times, w_on = t_on and w_off = t_off. One can also deal with the spatial process. It is, for example, introduced as a spatial dependence of exposure–weighted areas enclosed in circles with radii given by elements of an ordered set of separation angles of events as measured from a source direction. Then, the weights are given by these exposure–weighted areas.
We have applied the modified on–off tests to demonstrate what extend the words, the null no–source hypothesis (\(S_{on} \sim N(0,1)\)). Small blue and magenta points represent Li–Ma and binomial significances for the null no–source hypothesis (\(\beta = 1\)) to be rejected. The diagonal of the first quadrant is depicted by the dotted line.

First, we tested the MC data for the no–source hypothesis (\(\beta = 1\)). Our results are shown in Fig.1 in a quantile–quantil plot (QQ plot). Small blue and magenta points refer to asymptotic Li–Ma and binomial significances, respectively. Since these points do not lie on the indicated diagonal of the first quadrant, the sample statistics \(S_{LM}\) and \(S_{Bi}\) defined in Appendix A are shown not to come from Gaussian distribution with zero mean and unit variance. In other words, the null no–source hypothesis (\(\beta = 1\)) is demonstrated not to be true as expected. Nonetheless, because the blue and magenta points in Fig.1 approximately lie on a line, the distributions of the asymptotic Li–Ma and binomial statistics, \(S_{LM}\) and \(S_{Bi}\), are to be linearly related to the standardized Gaussian distribution.

In the second step, we tested the non–zero source hypothesis assuming a parameter \(\beta = 3\), i.e. \(\mu_{on} = 3\mu_{off}\). Used as the input of MC simulations. Resultant QQ plots for the asymptotic Li–Ma (big black points) and binomial (big red points) statistics are also depicted in Fig.1. In this case, the relationship between studied sample statistics and the standardized Gaussian distribution is well demonstrated. These statistics lie on the diagonal of the first quadrant implying that \(S_{LM} \sim N(0,1)\) and \(S_{Bi} \sim N(0,1)\) as well.

We conclude that the standard asymptotic measure of the level of significance of a source based on the null no–source conjecture can be trustworthy modified to search for an excess or deficit of events with respect to a preassigned source strength using a non–zero source hypothesis.

3.2 PKS 2155–304

We have applied the modified on–off tests to demonstrate the change in the non–zero \(\gamma\)–activity of several observed sources. In particular, here we present results showing to what extend the \(\gamma\)–ray variability observed in the direction of PKS 2155–304 on time scales of months and days can be verified on statistical grounds. To this end, we have adopted experimental data collected by the H.E.S.S. telescopes in the 2002–2003 and 2005–2007 campaigns [3,4].

Results of our analysis are presented in Figs.2 and 3. In these figures, the distributions of asymptotic sample significances are compared to the standardized Gaussian distribution in QQ plots. Both Li–Ma (black points) and binomial (red circles) significances are shown. Increasing sizes of marks refer to observational times of events for which asymptotic significances are depicted.

In Fig.2 the results of our analysis of the experimental data collected in the period 2003–2005 by the H.E.S.S. instrument are shown. In that time, 8 independent observations were recorded, see Table 3 in Ref. [3]. For each of these events we calculated both asymptotic significances for the non–zero source. The hypothetical \(\gamma\)–activity was characterized by a parameter \(\beta = 1.75\). This parameter roughly equals to an average observed on–source signal when compared to a background read out from the off–source region, i.e. \(\beta \approx \frac{\mu_{on}}{\mu_{off}}\). This way, heavier tails observed in the sample significances are more skewed than the reference distribution. Moreover, the significance distributions are more dispersed than the standardized Gaussian distribution. Except for one measurement at the end of 2002, the number of observed on– or off–source counts ranged from several hundreds to over ten thousand.

The agreement between both studied statistics depicted in Fig.2 is a salient feature. It is well visible that their distributions are more dispersed than the reference standardized Gaussian distribution. Except for the last three measurements with significances which are not inconsistent with the chosen average \(\gamma\)–activity of the source, the sample values of the remaining significances, and its trend in particular, suggest that the collected data set is not drawn from the reference distribution. This feature is to be interpreted as an unambiguous sign of monthly changes of the observed \(\gamma\)–ray flux from the investigated source.

Much more data on the PKS 2155–304 \(\gamma\)–activity has been collected in the 2005–2007 H.E.S.S. campaign [4]. In our analysis, only 29 events with MJD=53618–53705 and 54264–54376 were observed before and after the huge flare have been included. Using relevant data from Table A.1 in Ref. [4], we obtained an average non–zero source activity expressed by a parameter \(\beta = 1.38\). In all cases, the number of observed on– or off–source counts were above five hundred.

In Fig.3 our results obtained from the later H.E.S.S. campaign are summarized. Both studied asymptotic sample statistics, Li–Ma and binomial significances, agree with one another. Their distributions are more dispersed than the standardized Gaussian distribution. Moreover, their QQ plots are arced indicating that the sample significances distributions are more skewed than the reference distribution. This way, heavier tails observed in the sample distributions of the Li–Ma and binomial statistics suggest visible fluctuations in the inspected data set. Hence, not a few observations yielding the sample distributions of significances inconsistent with the reference distribution are to be considered as a signature of the time variability of the \(\gamma\)–ray flux observed from the studied source.

4 Conclusions

We introduced the on–off tests with the null hypothesis that assumes a predefined source present in the on–source region. Basic features of the Li–Ma as well as binomial
We adopt notation used in We focus on a level of significance associated with a statistical test trying to reject the null hypothesis stating that there is a source with a given activity in a suspected region. We adopt notation used in γ-ray astronomy. We deal with a hypothesis test in which the statistical significance of an excess or deficit of events with respect to a predefined source activity in a given region is established.

Let us assume that measured counts in the on–source region, \(N_{\text{on}}\), come from a Poissonian distribution with a mean \(\mu_{\text{on}} = w_{\text{on}}\lambda_{\text{on}}\), while off–source counts observed in the control off–source region, \(N_{\text{off}}\), come from a background Poissonian distribution with a mean \(\mu_{\text{off}} = w_{\text{off}}\lambda_{\text{off}}\). Here, \(\lambda_{\text{on}}\) and \(\lambda_{\text{off}}\) are unknown intensities of observed on–source and off–source counts, i.e. \(N_{\text{on}} \sim \text{Po}(w_{\text{on}}\lambda_{\text{on}})\) and \(N_{\text{off}} \sim \text{Po}(w_{\text{off}}\lambda_{\text{off}})\). Parameters \(w_{\text{on}}\) and \(w_{\text{off}}\) assign known on–source and off–source weights, respectively.

We will verify the null hypothesis stating \(\lambda_{\text{on}} = \lambda_{\text{off}}\), i.e. \(\mu_{\text{on}} = w_{\text{on}}\lambda_{\text{on}} = \alpha \mu_{\text{off}}\lambda_{\text{off}} = \alpha \mu_{\text{off}}\), where \(\alpha = \frac{w_{\text{on}}}{w_{\text{off}}}\) is the on–off parameter and a parameter \(\beta > 0\) that is chosen in advance is responsible for an excess or deficit of events in the on–source region when compared to the off–source one if \(\beta \neq 1\).

### A. The on–off method with non–zero source

We focus on a level of significance associated with a statistical test trying to reject the null hypothesis stating that there is a source with a given activity in a suspected region. We adopt notation used in γ-ray astronomy. We deal with a hypothesis test in which the statistical significance of an excess or deficit of events with respect to a predefined source activity in a given region is established.

Let us assume that measured counts in the on–source region, \(N_{\text{on}}\), come from a Poissonian distribution with a mean \(\mu_{\text{on}} = w_{\text{on}}\lambda_{\text{on}}\), while off–source counts observed in the control off–source region, \(N_{\text{off}}\), come from a background Poissonian distribution with a mean \(\mu_{\text{off}} = w_{\text{off}}\lambda_{\text{off}}\). Here, \(\lambda_{\text{on}}\) and \(\lambda_{\text{off}}\) are unknown intensities of observed on–source and off–source counts, i.e. \(N_{\text{on}} \sim \text{Po}(w_{\text{on}}\lambda_{\text{on}})\) and \(N_{\text{off}} \sim \text{Po}(w_{\text{off}}\lambda_{\text{off}})\). Parameters \(w_{\text{on}}\) and \(w_{\text{off}}\) assign known on–source and off–source weights, respectively.

We will verify the null hypothesis stating \(\lambda_{\text{on}} = \lambda_{\text{off}}\), i.e. \(\mu_{\text{on}} = w_{\text{on}}\lambda_{\text{on}} = \alpha \mu_{\text{off}}\lambda_{\text{off}} = \alpha \mu_{\text{off}}\), where \(\alpha = \frac{w_{\text{on}}}{w_{\text{off}}}\) is the on–off parameter and a parameter \(\beta > 0\) that is chosen in advance is responsible for an excess or deficit of events in the on–source region when compared to the off–source one if \(\beta \neq 1\).

### A.1 Binomial significance

Let us assume a statistical test based on a conditional distribution. Then, the probability that \(N_{\text{on}}\) events is observed in the on–source region provided that \(N = N_{\text{on}} + N_{\text{off}}\) counts measured in the whole inspected region follow the Poissonian distribution with the mean parameter \(\mu = \mu_{\text{on}} + \mu_{\text{off}}\), is given by

\[
P(N_{\text{on}} \mid N) = \frac{P_{N_{\text{on}}} (\mu_{\text{on}}) P_{N_{\text{off}}} (\mu_{\text{off}})}{P_N (\mu)} = \frac{N_{\text{on}}^{N_{\text{on}}} (1 - q)^{N_{\text{off}}}}{N^{N}}. \tag{1}
\]

Here, \(P_k (\mu) = \frac{\mu^k e^{-\mu}}{k!}\) assigns the Poissonian probability to observe \(k\) events if the mean is \(\mu\). The number of on–source counts obeys the binomial distribution with the binomial parameter \(q\). If \(\lambda_{\text{on}} = \beta \lambda_{\text{off}}\) holds, then in addition

\[
q = \frac{\mu_{\text{on}}}{\mu_{\text{on}} + \mu_{\text{off}}} = \frac{\alpha \beta}{1 + \alpha \beta}. \tag{2}
\]

This way, the on–off problem is reduced to judge whether measured on–source counts \(N_{\text{on}}\) can be viewed as a realization of the binomial distribution with the known parameter \(q_0\) provided \(N\) events are measured in the whole region under considerations. The level of significance, the probability with which the null hypothesis \((\lambda_{\text{on}} = \beta \lambda_{\text{off}})\) is refused in favor of an excess of events in the on–source region \((\lambda_{\text{on}} > \beta \lambda_{\text{off}})\) if it is true (excess \(p\)-value), is

\[
p_e = \sum_{k=N_{\text{off}}}^{N} \binom{N}{k} q_0^k (1 - q_0)^{N-k}. \tag{3}
\]

More precisely, the number of observed counts \(N_{\text{on}}\) attains a value the probability of which is under the stated hypothesis less than the predefined level of significance \(\delta\), \(p_e < \delta\). Accordingly, the level of significance with which the null hypothesis \((\lambda_{\text{on}} = \beta \lambda_{\text{off}})\) is rejected in favor of a deficit of
The maximum likelihood estimates of the on– and off–source means satisfy 
\[ \mu_{\text{on}} = \alpha \mu_{\text{off}}, \quad \mu_{\text{off}} = \beta \mu_{\text{off}}, \] 
and variances 
\[ \sigma^2_{\text{on}} = \alpha^2 \sigma^2_{\text{off}}, \quad \sigma^2_{\text{off}} = \beta^2 \sigma^2_{\text{off}}. \]

with one degree of freedom, 
\[ S_{\text{LM}}^2 \sim \chi^2(1). \]

\[ S_{\text{LM}} = \alpha N_{\text{on}} + \beta N_{\text{off}}. \]