Nuclear Instruments and Methods in Physics Research A 681 (2012) 55–60

Contents lists available at SciVerse ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

A real-time statistical alarm method for mobile gamma spectrometry—Combining counts of pulses with spectral distribution of pulses

Peder Kock [a,*], Jan Lanke [b], Christer Samuelsson [a]

[a] Medical Radiation Physics, Department of Clinical Sciences Malmö, Lund University, Skåne University Hospital Malmö, SE-205 02 Malmö, Sweden
[b] Department of Statistics, Lund University, Sweden

ARTICLE INFO

Article history:
Received 9 September 2011
Received in revised form
2 April 2012
Accepted 13 April 2012
Available online 21 April 2012

Keywords:
Gamma spectrometry
Hypothesis testing
Low counts
Multinomial distribution

ABSTRACT

A well-founded decision needs to take into account as much information from a sample as possible. In gamma spectrometry, the number of photons and their energy are the two quantities readily accessible to the physicist and both should be used in order to increase the power of a statistical test. While the problem of counts of pulses has been much studied the problem of spectral distribution of pulses has been generally overlooked. This work presents a statistical test combining tests on count rate and tests on spectral distribution. The proposed method is shown to have an acceptable false positive rate and, when compared with two other test statistics found in the literature, greater power.

1. Introduction

Statistical inference about analyte activity present in a sample is an important research topic in health physics and part of the more fundamental question: is there a signal present? To answer this question, using statistical inference, one either accepts or rejects the null hypothesis

H_0 : No signal present in sample
versus
H_1 : Signal present in sample

at an *a priori* determined significance level, α. The test statistic used can vary, but ideally the probability of rejecting H_0 when it is in fact true, i.e. a false positive or type I error, should be α [5].

Strom and MacLellan [19] evaluated eight test statistics with respect to their actual false positive rates, α'. For the lowest count rates (typically a Poisson mean $\mu_b < 2$), they found that no method satisfied the predefined significance level, α. It has long been known that this result is due to the discrete nature of counting statistics and the effects are especially severe in the low-level region (see e.g. [2,7,14]). Interestingly, the most well-known method in the health physics field, given by Currie [9], also showed the worst result with regard to α', even for intermediate count rates, while the method of Stapleton showed good results, i.e. $\alpha' \approx \alpha$ for $\mu_b > 5$ [19].

Taking into account the spectral information in a gamma-ray spectrum (or histogram) should increase the power of the test, but going from one bin to multiple bins also increases the complexity of the problem. One method, which calculates the probability for each possible pulse configuration, given some background distribution, was presented by Méray [15]. Compared with the single-bin method of Currie this approach significantly lowered the *detection limit* (see Ref. [9] for a definition) [16].

This work presents a new test based on a combination of a count rate test, viz. a modification of the Sumerling and Darby (S&D) test [20], and a likelihood ratio test of the spectral distribution of the counts. The two *p*-values so obtained are subjected to Fisher's method for combining *p*-values. Both the false and true positive rates for the proposed method are evaluated and compared with those of several other methods.

The method described in this work is designed for, and evaluated in, a mobile gamma spectrometry context. This typically means conducting repeated short-term measurements, possibly for an extended period of time, while searching for a radioactive source. To avoid too many false positive alarms the chosen α is small (0.1–1%) and the count rates in the simulations are low to intermediate ($5 \le \mu_b \le 30$). High count rate environments, where pulses are abundant, can provide many challenges but generally not with respect to the problem addressed in this

* Corresponding author. Tel.: +46 40 33 86 64.
E-mail address: peder.kock@med.lu.se (P. Kock).

0168-9002/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.nima.2012.04.038

work. The present work might still be useful in other scientific fields, despite the chosen context.

2. Theory and methods

Starting with the basic model of radioactive counting, the Poisson distribution, we show that the spectral distribution of pulses, given the total count, is described by binomial or multinomial probabilities, depending on the number of channels used. We then present two hypotheses that split the radioactive counting problem into two parts: first, the problem of pulse sums, and secondly, the problem of spectral distribution.

2.1. Single-channel Poisson model

Suppose we have a radioactive counting experiment with two samples. These samples will henceforth be referred to as *background* and *sample*. Suppose also that the experiment involves only one channel in which pulses are registered. The probability of observing k pulses from *sample* is a Poisson probability

$$P(X=k)=\frac{e^{-\mu}\mu^k}{k!} \tag{1}$$

where μ is the true mean. Substituting μ by λ in Eq. (1) then gives the probability of observing k pulses from *background*. By combining the counts from sample, x, and background, y, so that $z=x+y$, the conditional probability of observing a sample-background pair can be written

$$P(X=x,Y=y|Z=z)=\binom{z}{x}q^x(1-q)^y \tag{2}$$

where $q=\mu/(\mu+\lambda)$. For a derivation, which is straight-forward using two Poisson distributions, see e.g. [6,19].

An interesting observation is that in high energy physics (HEP) and gamma-ray astronomy (GRA) the single-channel problem of Poisson ratios is called signal-bin/sideband and the on/off problem respectively. It is an old problem that has got much attention, see e.g. Cousins et al. [6] for a comprehensive review. The problem is also well known in the health physics/gamma-ray spectroscopy field, see e.g. [1,9,10,18,19].

2.2. Dual-channel properties

If the pulses from sample and background are split into two separate channels, c_1 and c_2, then for each channel the probability of observing k pulses from background or sample is given by Eq. (1), substituting μ by the appropriate true mean. The joint probability of observing x_1 and x_2 pulses from sample in the two channels is then

$$P(X_1=x_1,X_2=x_2)=P(x_1)P(x_2)=\frac{e^{-\mu_1}\mu_1^{x_1}}{x_1!}\frac{e^{-\mu_2}\mu_2^{x_2}}{x_2!} \tag{3}$$

where the first step can be carried out since the random variables X_1,X_2 are assumed to be independent. The background pulses are also Poisson distributed with true means λ_1,λ_2 in c_1 and c_2 respectively. The probability of observing $Y_1=y_1$ and $Y_2=y_2$ counts in background is also given by Eq. (3), substituting μ_i by λ_i. The conditional probability of observing a pair of counts from sample, given the sum of the counts, can be shown to be

$$P(X_1=x_1|X_1+X_2=x)=\binom{x}{x_1}\left(\frac{\mu_1}{\mu_1+\mu_2}\right)^{x_1}\left(\frac{\mu_2}{\mu_1+\mu_2}\right)^{x-x_1} \tag{4}$$

and analogously for the background

$$P(Y_1=y_1|Y_1+Y_2=y)=\binom{y}{y_1}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{y_1}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{y-y_1} \tag{5}$$

2.3. Multichannel properties

Moving on to k channels and using the notation

$$x=x_1+x_2+\cdots+x_k \tag{6a}$$

$$\mu=\mu_1+\mu_2+\cdots+\mu_k \tag{6b}$$

$$q_1=\frac{\mu_1}{\mu},\ q_2=\frac{\mu_2}{\mu},\ldots,q_k=\frac{\mu_k}{\mu} \tag{6c}$$

$$\mathbf{x}=(x_1,x_2,\ldots,x_k) \tag{6d}$$

$$\mathbf{q}=(q_1,q_2,\ldots,q_k) \tag{6e}$$

the probability of observing $\mathbf{x}$ in channels $1,2,\ldots,k$ is

$$\begin{aligned}P(\mathbf{x})&=\frac{e^{-\mu}\mu^x}{x_1!x_2!\cdots x_k!}\left(\frac{\mu_1}{\mu}\right)^{x_1}\left(\frac{\mu_2}{\mu}\right)^{x_2}\cdots\left(\frac{\mu_k}{\mu}\right)^{x_k}\\&=\frac{e^{-\mu}\mu^x}{x!}\binom{x}{x_1,x_2,\ldots,x_k}q_1^{x_1}q_2^{x_2}\cdots q_k^{x_k}\end{aligned} \tag{7}$$

where

$$\binom{x}{x_1,x_2,\ldots,x_k}=\frac{x!}{x_1!x_2!\cdots x_k!}$$

is a multinomial coefficient. The conditional probability of observing $\mathbf{x}$, given a total of x pulses, is then

$$P(\mathbf{x}|x)=\frac{P(\mathbf{x})}{P(x)}=\binom{x}{x_1,x_2,\ldots,x_k}q_1^{x_1}q_2^{x_2}\cdots q_k^{x_k} \tag{8}$$

which is a probability in a multinomial distribution, $\text{Mult}_k(x;\mathbf{q})$. Note that if we are interested only in the ith frequency in Eq. (8) it is binomially distributed

$$(x_i|x)\in\text{Bin}(x,q_i). \tag{9}$$

The probability of observing a spectral distribution $\mathbf{y}$ within background can be derived using Eqs. (6) and (7), substituting x_i and μ_i by y_i and λ_i.

2.4. Hypotheses

To test if sample and background are two samples from the same underlying distribution several tests and hypotheses can be constructed. In this work we choose to study two different hypotheses, the first one, $H_0^{(S)}$, concerning the pulse sum and the second, $H_0^{(R)}$, concerning the spectral distribution within the samples

$$H_0^{(S)}:\sum_{i=1}^{k}\mu_i=\sum_{i=1}^{k}\lambda_i \tag{10a}$$

$$H_0^{(R)}:\mu_i/\mu=\lambda_i/\lambda\quad\text{for } i=1\ldots k \tag{10b}$$

$$H_0^{(SR)}:\mu_i=\lambda_i\quad\text{for } i=1\ldots k \tag{10c}$$

where, as easily seen, $H_0^{(SR)}$ is the combination of $H_0^{(S)}$ and $H_0^{(R)}$.

2.5. Test statistics for $H_0^{(S)}$

2.5.1. Sumerling and Darby's method

The probability mass function (pmf) of S&D is given in Eq. (2). Summing over all probabilities from x up to $z=x+y$ gives the