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

Kock, Peder; Lanke, Jan; Samuelsson, Christer

Published in:
Nuclear Instruments & Methods in Physics Research. Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment

DOI:
10.1016/j.nima.2012.04.038

2012

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] and Christer Samuelsson[a]

[a]*Medical Radiation Physics, Department of Clinical Sciences Lund, Lund University*
[b]*Department of Statistics, Lund University*

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 Meray et al. [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 & 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 to 1%) and the count rates in the simulations are low to intermediate ($5 \leq \mu_b \leq 30$). High count rate environments, where pulses are abundant, can provide many challenges but generally not with respect to the problem addressed in this 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

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

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

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

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

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

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, $\mathrm{Mult}_k(x\,; \mathbf{q})$. Note that if we are interested only in the i'th frequency in Eq. (8) it is binomially distributed

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

The probability of observing a spectral distribution $\mathbf{y}$ within background can be derived using Eqs. (6-7), 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:

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

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

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

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 probability of observing x or more pulses in sample, given the sum, z, of the pulses in sample and background:

$$P_{\mathrm{SD}} = \sum_{i=x}^{z} \binom{z}{i} q^i (1-q)^{z-i} \tag{11}$$

which is the test statistic [20]. If the counting times of sample and background are equal, then $q = 1/2$. In the case of different counting times for sample (t_x) and background (t_y), q becomes

$$q = \frac{t_x}{t_x + t_y} \tag{12}$$

The null hypothesis, $H_0^{(S)}$, that the sample is blank is rejected if a blank sample would have produced a gross count at least as large as the one observed at most $100\alpha\%$ of the time [20]; that is, if

$$P_{\mathrm{SD}} \leq \alpha \tag{13}$$

As noted by Refs. [2, 7], the binomial test is known to be conservative, i.e. it will always reject H_0 less than $100\alpha\%$ of the time under H_0.

2.5.2 Currie's method

The test proposed by Currie [9] is based on a comparison of the net count $x - y$ and the quantity

$$C_{\text{currie}}(y) = k_\alpha \sqrt{2y} \tag{14}$$

where k_α is the quantile of a standard normal distribution such that $P(X > k_\alpha) = \alpha$; $H_0^{(S)}$ is rejected at the α-level if: $x - y > C_{\text{currie}}$.

ANSI/HPS N13.30-1996 [12] presented Eq. (14) in a more general form, where the counting times of background (t_y) and sample (t_x) can be different

$$C_{\text{N13.30}}(y, t_x, t_y) = k_\alpha \sqrt{\frac{y}{t_y}\left(\frac{1}{t_x} + \frac{1}{t_y}\right)} \tag{15}$$

[19] therefore dubbed this variant of Currie's rule to "N13.30". In this form the null hypothesis, $H_0^{(S)}$, is rejected if the net count rate $R_n = x/t_x - y/t_y > C_{\text{N13.30}}$.

2.5.3 Stapleton's method

Stapleton's method, described in [19], computes a standard normal deviate, $z_{\text{stapleton}}$, from the observed counts x, y and counting times t_x, t_y

$$z_{stapleton}(x, y, t_x, t_y) = 2\frac{\sqrt{\frac{x+d}{t_x}} - \sqrt{\frac{y+d}{t_y}}}{\sqrt{\frac{1}{t_x} + \frac{1}{t_y}}} \tag{16}$$

where d is a parameter, $0 < d < 1$. Throughout this work, d was 0.4 to be comparable to the results of Strom and MacLellan [19]. Using this test, $H_0^{(S)}$ is rejected if $z_{\text{stapleton}} > k_\alpha$.

2.5.4 Sumerling and Darby, mid-p version

As noted above, the S&D method is known to be conservative. A proposition to remove the conservativeness of discrete test statistics was proposed by Lancaster [14]; the method has been evaluated and recommended by e.g. Refs. [2, 3, 7] . In our case the technique amounts to taking only half the probability in the first term of Eq. (11):

$$\begin{aligned} P_{\text{SDmidp}} &= \frac{1}{2}P(X = x \mid z) + P(X > x \mid z) \\ &= \frac{1}{2}\binom{z}{x} q^x (1-q)^{z-x} + \sum_{i=x+1}^{z} \binom{z}{i} q^i (1-q)^{z-i} \end{aligned} \tag{17}$$

where the second term is interpreted as zero if $x = z$. As for the S&D-method, $H_0^{(S)}$ is rejected at the α-level if $P_{\text{SDmidp}} \leq \alpha$. Since $P_{\text{SD}} > P_{\text{SDmidp}}$ the latter version will be less conservative.

2.6 Test statistics for $H_0^{(R)}$

2.6.1 Goodness-of-fit test

The fit of a multinomial model to a data set, $\mathbf{x}$, can be tested using a goodness-of-fit (GoF) test [4]. The probability of observing a data set at least as extreme as the one observed, $P(\mathbf{x})$, can be written as

$$P_{\text{fit}} = \sum_{\mathbf{a}:P(\mathbf{a})\leq P(\mathbf{x})} P(\mathbf{a}) \tag{18}$$

To test if an observed outcome $\mathbf{x}$ came from a distribution according to the null, [8] gives the following four-step recipe:

1. For every possible outcome $\mathbf{a}$, calculate the probability $P(\mathbf{a})$ according to Eq. (8)
2. Rank the probabilities from smallest to largest
3. Starting with the smallest rank, add the consecutive probabilities up to and including that associated with $\mathbf{x}$; this cumulative probability gives the chance of obtaining an outcome that is no more probable than $\mathbf{x}$
4. Reject H_0 if this cumulative probability is at most α

Naively traversing all combinations at least as extreme as $\mathbf{x}$ (step 1-2) quickly becomes computationally infeasible, as the number of channels k, and pulse sum x, grow. To overcome this, an approximation to the distribution of the log-likelihood ratio statistic G^2 can be used instead. One such approximation, based on Fast Fourier Transform (FFT), is given by Keich and Nagarajan. The p-value of the fit is approximated from the entropy score

$$s_0 = \sum_{i=1}^{k} x_i \ln\left(\frac{x_i}{\pi_i x}\right) \tag{19}$$

where π_i are the expected background cell probabilities and $x = \sum x_i$ [13, 17]. However, in order to use the multinomial GoF test, the background probabilities π_i have to be known.

2.6.2 Test using likelihood ratio

A test of the pulse distribution can be performed using the likelihood function. The multinomial probability function was given in Eq. (8). Using the same notation we now have two multinomial observations, the sample $\mathbf{x}$ and the background, $\mathbf{y}$, which are independent. The likelihood function is

$$P(\mathbf{x}|x)P(\mathbf{y}|y)$$

where

$$P(\mathbf{x}|x) = c_x \prod_{i=1}^{k} \left(\frac{\mu_i}{\mu}\right)^{x_i} \tag{20}$$

$$P(\mathbf{y}|y) = c_y \prod_{i=1}^{k} \left(\frac{\lambda_i}{\lambda}\right)^{y_i}; \tag{21}$$

here c_x and c_y are irrelevant constants.

If we introduce

$$\xi_i = \frac{\mu_i}{\mu}, \quad \eta_i = \frac{\lambda_i}{\lambda}$$

the likelihood function can be written

$$\mathcal{L}(\boldsymbol{\xi}, \boldsymbol{\eta}) = c_x c_y \left(\prod_{i=1}^{k} \xi_i^{x_i}\right) \left(\prod_{i=1}^{k} \eta_i^{y_i}\right) \tag{22}$$

and under the null hypothesis $\boldsymbol{\xi} = \boldsymbol{\eta}$ we get

$$\sup_{H_0} \mathcal{L} = c_x c_y \prod_{i=1}^{k} \left(\frac{z_i}{z}\right)^{z_i} \tag{23}$$

where $z_i = x_i + y_i$, $z = x + y$, while

$$\sup_{H_1} \mathcal{L} = c_x c_y \left(\prod_{i=1}^{k} \left(\frac{x_i}{x}\right)^{x_i}\right) \left(\prod_{i=1}^{k} \left(\frac{y_i}{y}\right)^{y_i}\right) \tag{24}$$

Using the notation

$$\Lambda = \frac{\sup_{H_0} \mathcal{L}}{\sup_{H_1} \mathcal{L}} \tag{25}$$

the log-likelihood ratio statistic is obtained as

$$-2 \ln \Lambda = 2 \sum_{i=1}^{k} \left\{ x_i \ln\left(\frac{x_i}{x}\right) + y_i \ln\left(\frac{y_i}{y}\right) - z_i \ln\left(\frac{z_i}{z}\right) \right\} \tag{26}$$

This statistic has an approximate $\chi^2(k-1)$ distribution for large counts [21]. $H_0^{(R)}$ is rejected at the α level if

$$1 - F_{k-1}(-2\ln\Lambda) < \alpha \tag{27}$$

where F_{k-1} is the cumulative chi-square function for $k-1$ degrees of freedom. Note that the method only traverses the sum given in Eq. (26) once, it therefore has linear complexity i.e. $\mathcal{O}(k)$.

In the special case $k = 2$ there is a possibility to distinguish between one- and two-sided tests; the version described above is two-sided. However, there are situations where a one-sided test could be useful. Suppose we have two bins, the first one covering a region of interest and the second a region where only background pulses are expected. The alternative hypothesis would then be

$$H_{1,>} : \frac{\mu_1}{\mu} > \frac{\lambda_1}{\lambda} \tag{28}$$

Clearly $\sup_{H_0} \mathcal{L}$ is still given by Eq. (23) while, cf. Eq. (24),

$$\sup_{H_{1,>}} \mathcal{L} = \begin{cases} \sup_{H_1} \mathcal{L} & \text{if } \dfrac{x_1}{x} > \dfrac{y_1}{y} \\ \sup_{H_0} \mathcal{L} & \text{otherwise} \end{cases} \tag{29}$$

and hence

$$\Lambda = \begin{cases} \dfrac{\sup_{H_0} \mathcal{L}}{\sup_{H_1} \mathcal{L}} & \text{if } \dfrac{x_1}{x} > \dfrac{y_1}{y} \\ 1 & \text{otherwise} \end{cases} \tag{30}$$

Now $-2\ln\Lambda$ has, under H_0, an approximate distribution that is no longer $\chi^2(1)$; rather, it is a mixture in equal proportions of a $\chi^2(1)$ and a $\chi^2(0)$ where $\chi^2(0)$ is a one-point distribution at zero.

2.7 The proposed method

Combining a test on the pulse sum, $H_0^{(S)}$, with a test on the pulse distribution, $H_0^{(R)}$, can be expected to present advantages: utilising more of the information contained in the sample will likely increase the power of the test. Given p-values from two independent tests, one can combine them according to Fisher [11]

$$p_{\mathrm{SR}} = (1 - \ln(p_{\mathrm{S}} p_{\mathrm{R}}))\, p_{\mathrm{S}} p_{\mathrm{R}} \tag{31}$$

where p_{SR} is the p-value of the combined test. Another possibility is to derive the likelihood ratio directly for $H_0^{(SR)}$, cf. section 2.4; however, that option is not explored further in this work. Instead we propose using the mid-p version of the S&D method given in Eq. (17) in conjunction with a test on the likelihood ratio, given in Eq. (30). In the present work we consider only the case $k = 2$ and use the one-sided version of the likelihood ratio test.

3 Numerical Calculations and Monte Carlo Simulations

3.1 False positives

The actual false positive rates α' were derived by direct calculations or estimated through paired-blanks Monte Carlo simulations. A recipe for the direct probability calculations of α' is given in Appendix D.1 by Cousins et al. [6]. The direct calculation method was used for all methods testing the pulse sum, i.e. $H_0^{(S)}$. For the methods testing $H_0^{(R)}$ or $H_0^{(SR)}$ the direct method was not used because of the complexity of the multinomial problem.

The false positive rate was evaluated for Poisson true means $\mu = \lambda$, ranging from 5 to 30 with α at 0.1% and 1%. These are strict significance levels; for example Strom and MacLellan evaluated α's from 0.1% to 5%. However, for mobile gamma spectrometry a false alarm rate of more than 1% would be inadequate as many short-term measurements are carried out, often for an extended period of time. For each λ in the simulations, $2k \times 10^6$ samples (one sample and one background, for each channel) were drawn from an appropriate Poisson distribution.

3.2 True positives

While the actual false positive rate simulations evaluate the test statistics with respect to their false alarm rate, another set of numerical calculations and simulations were required to evaluate the method's sensitivities. By simulating a ^{137}Cs point source at $r = 20$ m from a virtual detector, the different methods were evaluated with respect to their true positive rates, i.e. their sensitivities. In all simulations and calculations 1 s observations were used.

The source emitted photons isotropically which, after attenuation in air, gave rise to a mean count rate, N s^{-1}, in the detector

$$N(A) = \frac{\epsilon\, p_\gamma\, e^{-ar}}{4\pi r^2}\, A = 1.683 \times 10^{-7} A \tag{32}$$

where A is the ^{137}Cs activity (Bq), $a = 9.4 \times 10^{-3}$ m^{-1} the assumed linear attenuation coefficient in air, $p_\gamma = 0.851$ the probability of 662 keV photon emission given a ^{137}Cs decay and $\epsilon = 1.2 \times 10^{-3}$ m^2 the assumed detector efficiency. The latter corresponds to a detector with a relative efficiency of roughly 50% in relation to a 3"x3" NaI(Tl).

The true positive Monte Carlo simulations were constructed using the following scheme; for each activity A:

1. Perform the following n times:

  i. Draw 2k background samples $a_i, b_i \in \text{Po}(\lambda_i)$

  ii. Draw ^{137}Cs pulses $a_\gamma \in \text{Po}(N(A))$, cf. Eq. (32)

  iii. Add a_γ to $\mathbf{a}$ according to $\mathbf{p} = (p_1, \ldots, p_k)$

  iv. Check whether H_0 is rejected at the α-level using $\mathbf{x} = \mathbf{a}$ and $\mathbf{y} = \mathbf{b}$

2. Estimate the power of the test for activity A

$$1 - \beta = \frac{n_{rejceted}}{n} \tag{33}$$

where β is the probability of a false negative.

The range of activities tested was $A = [20, 200]$ MBq, effectively adding $N = [3.37, 33.7]$ pulses per second to the sample. The background count rate, λ_i, was 10. As in the previous section, the results for $H_0^{(S)}$ were instead derived using direct calculation.

4 Results and Discussion

4.1 False positives - results under H_0

Results from the false positive calculations and simulations are given in Tables 1-3. Standard errors are given for the simulated results. The results are split into three tables: Table 1 holds the results for test statistics for $H_0^{(S)}$, Table 2 the results for test statistics for $H_0^{(R)}$ and Table 3 for the proposed method, testing $H_0^{(SR)}$. The tests on spectral distribution were done using two bins which had equal probabilities, $p_i = 1/2$, given an event. This has been performed using both the one- and the two-sided version.

With two exceptions the tests perform better at the 1%-level than at the 0.1%-level. The relation between the reference tests, as regards their performance, are the same at the two levels. The results of Currie's test are, as

	$\alpha = 0.1$ %				$\alpha = 1$ %			
	Currie	S&D	S&D mid-p	Stapleton	Currie	S&D	S&D mid-p	Stapleton
$\mu = \lambda$								
5	2.768	0.029	0.067	0.143	5.79	0.33	0.65	1.19
10	1.483	0.044	0.078	0.136	3.57	0.55	0.91	1.02
15	0.948	0.051	0.085	0.120	3.08	0.60	0.90	1.00
20	0.742	0.059	0.091	0.108	2.70	0.65	0.94	1.01
25	0.618	0.062	0.092	0.106	2.43	0.70	0.97	1.00
30	0.532	0.065	0.094	0.103	2.22	0.70	0.95	1.02

Table 1: Actual false positive rates, α', given in percent (%), for the tests on $H_0^{(S)}$.

already demonstrated by e.g. Strom and MacLellan, unacceptable. Sumerling and Darby's test is very conservative, but most of this conservativeness is gone when using the mid-p version. Stapleton's test shows the overall best performance, but the proposed test and S&D mid-p test also yield false positive rates close to α. These latter two methods show relatively better results at the 0.1%-level, than at the 1%-level.

	$\alpha = 0.1$ %		$\alpha = 1$ %	
	Likelihood ratio	Goodness-of-fit[a]	Likelihood ratio	Goodness-of-fit[a]
$\mu_i = \lambda_i$				
2.5	0.190 ± 0.004	0.001 ± 0.000	1.82 ± 0.01	0.12 ± 0.00
5	0.196 ± 0.004	0.060 ± 0.002	1.47 ± 0.01	0.95 ± 0.01
7.5	0.145 ± 0.004	0.126 ± 0.004	1.22 ± 0.01	1.31 ± 0.01
10	0.125 ± 0.004	0.107 ± 0.003	1.14 ± 0.01	1.25 ± 0.01
12.5	0.124 ± 0.004	0.100 ± 0.003	1.11 ± 0.01	1.12 ± 0.01
15	0.109 ± 0.003	0.103 ± 0.003	1.08 ± 0.01	1.07 ± 0.01

[a] assumes known π_i.

Table 2: Actual false positive rates, α', given in percent (%), for the tests on $H_0^{(R)}$ using two equiprobable bins. For the cases shown in the table the one-sided and two-sided likelihood ratio tests have identical α'.

One can imagine scenarios were it might be appropriate or advantageous to use the goodness-of-fit test, given by Keich and Nagarajan (K&N). One condition must, however, be met in order to use that method: the expected spectral distribution of the background has to be known. One would then expect the GoF test to perform better than the likelihood ratio tests at low values of λ, because the chi-square approximation, which is central to the likelihood ratio tests, is not very good at the lowest values of λ. However, as seen in Table 2, the results of the K&N GoF test are in the same range as the much simpler likelihood ratio tests at the 1%-level.

	Proposed method	
	$\alpha = 0.1$ %	$\alpha = 1$ %
$\mu_i = \lambda_i$		
2.5	0.070 ± 0.003	1.305 ± 0.011
5	0.153 ± 0.003	1.227 ± 0.011
7.5	0.121 ± 0.003	1.088 ± 0.010
10	0.107 ± 0.003	1.031 ± 0.010
12.5	0.103 ± 0.003	1.025 ± 0.010
15	0.097 ± 0.003	0.999 ± 0.010

Table 3: Actual false positive rates, α', given in percent (%), for the proposed method using two equiprobable bins.

4.2 True positives - results under H_1

Figs. 1-2 present the results from the true positive simulations and calculations at the $\alpha = 0.1$%-level. Fig. 1 shows the power $(1-\beta)$ of the different methods as a function of the activity, as described in section 3.2. The methods of Currie and S&D are omitted from Fig. 1 due to their unacceptable false positive rates (cf. Table 2). As shown in Fig. 1 the proposed method has a higher power than the reference methods, for all levels of signal (activities) added.

Fig. 1 shows a best-case scenario for the special case k=2; that is, all ^{137}Cs pulses are placed in the first bin. This is favourable for the proposed method. However, and as described in the likelihood ratio test section (2.6.2), it is not an unrealistic scenario. The effect of different signal probabilities, i.e. different $\boldsymbol{p}$, is shown in Fig. 2. As can be seen in Fig 2., the power of the proposed method is higher than that of the S&D mid-p version as long as 70% or more ($p_1 \geq 0.7$) of the ^{137}Cs pulses are placed in the first bin. For values of p_1 below this critical level the proposed method performs worse than the methods testing for $H_0^{(S)}$ alone. This can be understood by considering that the proposed method is using the one-sided likelihood ratio test. This test, as shown in Eq. (30), is sensitive to a higher proportion of the total pulses in the first bin than the second. Cases where this is not true (50% if $\boldsymbol{p} = (0.5, 0.5)$) are therefore ignored by the proposed method.

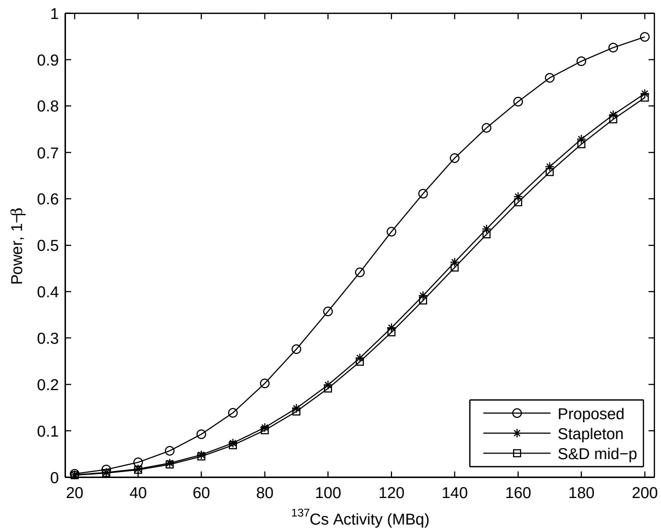

Figure 1: Actual true positive rates, $1-\beta$, as a function of the ^{137}Cs activity.

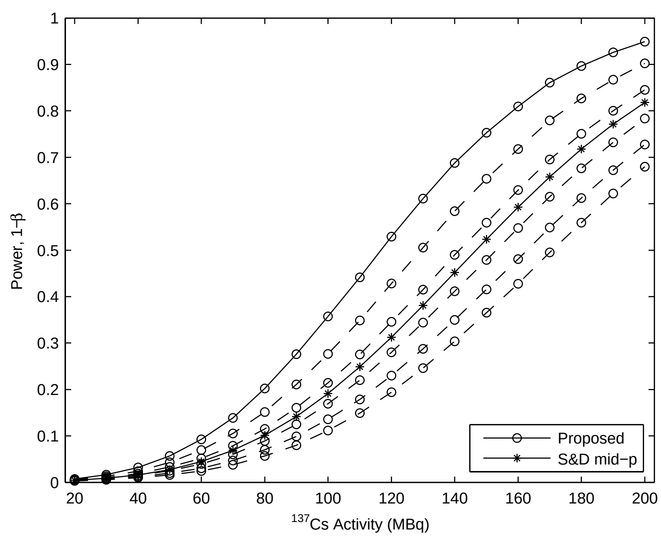

Figure 2: Actual true positive rates, $1-\beta$, as a function of the ^{137}Cs activity, for a range of p_i parmeter values. The mid-p version of Sumerling & Darby's test is also given, for reference. Top curve (solid line, circles): $\mathbf{p} = \{1.0, 0\}$, then in steps of 0.1 (dashed lines, circles) to bottom curve (dashed line, circles) $\mathbf{p} = \{0.5, 0.5\}$.

5 Conclusions

This work has presented a novel method consisting in combining counts of pulses with spectral distribution of pulses. The proposed method showed an acceptable false positive rate (α'), in relation to the given α, and good power $(1-\beta)$, when compared with the reference methods. It is fast to use, with linear complexity in relation to the number of channels used.

Acknowledgements

The authors would also like to thank an anonymous reviewer for insightful suggestions and comments, especially concerning the single-channel on/off problem. This work was supported by the Swedish Radiation Safety Authority, SSM.

References

References

[1] Alvarez, J. L. (2007), Poisson-based detection limit and signal confidence intervals for few total counts. *Health Phys.*, 93(2); pp. 120–126

[2] Agresti, A. and Gottard, A. (2005), Comment: Randomized confidence intervals and the mid-p approach. *Stat. Sci.*, 20(4); pp. 367–371

[3] Agresti, A. and Gottard, A. (2007), Nonconservative exact small-sample inference for discrete data. *Comput. Stat. Data. An.*, 51(12); pp. 6447–6458

[4] Baker, S. and Cousins, R.D. (1984), Clarification of the use of chi-square and likelihood functions in fits to histograms. *Nucl. Instrum. Meth. A*, 221(2); pp. 437–442

[5] Barlow, R.J. (1989), *Statistics: A guide to the use of statistical methods in the physical sciences.* New York: John Wiley & Sons Ltd., 1st ed.

[6] Cousins, R.D., Linnemann, J.T. and Tucker, J. (2008), Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a poisson process. *Nucl. Instrum. Meth. A.*, 595(2); pp. 480–501

[7] Cousins, R.D., Hymes, K.E. and Tucker, J. (2010), Frequentist evaluation of intervals estimated for a binomial parameter and for the ratio of poisson means. *Nucl. Instrum. Meth. A.*, 612(2); pp. 388–398

[8] Cressie, N. and Read, T.R.C. (1989), Pearson's χ^2 and the loglikelihood ratio statistic G^2: A comparative review. *Int. Stat. Review*, 57(1); pp. 19–43

[9] Currie, L.A. (1968), Limits from qualitative detection and quantitative determination. *Anal. Chem.*, 40(3); pp. 586–593

[10] De Geer, L.-E. (2004), Currie detection limits in gamma-ray spectroscopy. *Appl. Radiat. Isotopes*, 61(2-3); pp. 151–160

[11] Fisher, R. A. (1932), *Statistical methods for research workers.* London: Oliver and Boyd, 4th ed.

[12] Health Physics Society (1996), *Performance criteria for radiobioassay. An American national standard. ANSI/HPS N13.30-1996.* American National Standards Institute, New York, USA

[13] Keich, U. and Nagarajan, N. (2006), A fast and numerically robust method for exact multinomial goodness-of-fit test. *J. Comput. Graph. Stat.*, 15(4); pp. 779–802

[14] Lancaster, H.O. (1961), Significance tests in discrete distributions. *J. Am. Stat. Assoc.*, 56(294); pp. 223–234

[15] Méray, L. (1994), Deconvolution of nuclear spectra of low counts. *Nucl. Instrum. Meth. A.*, 353(1-3); pp. 272–275

[16] Méray, L. and Demény, O. (2001), Detection limit and decision thresholds in spectrometry. *Appl. Spectrosc.*, 55(8); pp. 1102–1108

[17] Nagarajan, N., Jones, N. and Keich, U. (2005), Computing the P-value of the information content from an alignment of multiple sequences. *Bioinformatics*, 21(suppl 1); pp. 311–318

[18] Potter, W.E (2005), Comparison of approximate formulas for decision levels and detection limits for paired counting with the exact results. *J. Radioanal. Nucl. Ch.*, 263(2); pp. 321–325

[19] Strom, D.J. and MacLellan, J.A. (2001), Evaluation of eight decision rules for low-level radioactivity counting. *Health Phys.*, 81(1); pp. 27–34

[20] Sumerling, T.J. and Darby, S.C. (1981), *Statistical aspects of the interpretation of counting experiments designed to detect low levels of radioactivity.* National Radiological Protection Board (NRPB); NRPB-R113, Chilton, UK

[21] Wilks, S.S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses, *Ann. Math. Stat.*, 9; pp. 60–62,