A Bayesian approach to evaluate confidence intervals in counting experiments with background

F. Loparco

Dipartimento Interateneo di Fisica “M. Merlin”, dell’Università degli Studi di Bari and
INFN Sezione di Bari, Via Amendola 173, I-70126 Bari, Italy

M. N. Mazziotta

INFN Sezione di Bari, Via Amendola 173, I-70126 Bari, Italy

Abstract

In this paper we propose a procedure to evaluate Bayesian confidence intervals in counting experiments where both signal and background fluctuations are described by the Poisson statistics. The results obtained when the method is applied to the calculation of upper limits will also be illustrated.

Keywords: Counting experiments, Confidence intervals, Upper limits

1. Introduction

The evaluation of confidence intervals and limits is a common task in particle physics. Usually the goal of an experiment is that of determining a parameter $\theta$ starting from a set of measurements of a random variable $x$ (the outcome of the experiment) and assuming an hypothesis for the probability distribution function (p.d.f.) $f(x|\theta)$. A confidence interval for $\theta$ at the confidence level (or coverage probability) $1-\beta$ includes the true value of $\theta$ with a probability $1-\beta$. This means that if the experiment were repeated many times, the estimates of $\theta$ will fall in the confidence interval in a fraction

Email addresses: loparco@ba.infn.it (F. Loparco), mazziotta@ba.infn.it (M. N. Mazziotta)

Preprint submitted to Elsevier January 21, 2013
1 − β of the experiments.

Confidence intervals can be evaluated following either the frequentist or the Bayesian approaches (a detailed discussion about these methods can be found in ref. [1]). One of the main issues arising when applying the frequentist approach is that the confidence intervals may include unphysical regions for the parameter. These cases can be fixed by introducing some “ad-hoc” corrections in the mathematical procedures used to evaluate the intervals (see for instance ref. [4, 5]). Such issues can be easily avoided when following the Bayesian approach, by properly choosing the prior p.d.f. for the parameter. However, the fact that the Bayesian approach requires a choice of the prior p.d.f. for the parameter introduces some degree of arbitrariness in the evaluation of confidence intervals.

In many cases the outcome of an experiment can be described in terms of a set parameters, not all being of any interest for the final result (nuisance parameters). In these cases the experimenter wishes to evaluate confidence regions on the parameters of interest, in a manner that is independent on the nuisance parameters. From a mathematical point of view, the Bayesian approach allows one to treat the nuisance parameters in a very simple and straightforward way. Indicating respectively with θ and ν the parameters of interest and the nuisance parameters, one has to write down their joint prior p.d.f. π(θ, ν) and to evaluate from it the marginal prior p.d.f. for θ. In most cases θ and ν are independent random variables, and their joint p.d.f. can be factorized as π(θ, ν) = π(θ)π(ν), thus making the calculation easier.

In the following sections we will illustrate an application of the Bayesian approach to the analysis of a counting experiment with background. We will assume that the outcome of the experiment can be modeled in terms of a parameter of interest (signal) and of a nuisance parameter (background), that will be supposed to be independent on each other. The posterior p.d.f.

---

1In the most general case θ is a vector of parameters, i.e. θ = (θ₁, θ₂, ..., θₙ) and x represents a set of observables, i.e. x = (x₁, x₂, ..., xₘ). Consequently, the p.d.f. f(x|θ) = f(x₁, ..., xₘ|θ₁, ..., θₙ) will be a function of m+n variables. Also, the confidence interval defined in the case of a single parameter becomes a n-dimensional confidence region for the vector of parameters θ.

2This can be done by assigning null probabilities to the values of θ in the unphysical regions.

3Hereafter we will us the greek symbol π to indicate prior p.d.f.s and the latin symbol p to indicate posterior p.d.f.s
for the signal will be derived starting from a set of simple and quite general assumptions on the prior p.d.f.s for both the signal and the background. The formulas to evaluate upper limits on the signal will also be derived and the results will be discussed. Our results are a generalization of the ones illustrated in ref. [1], where the same problem is discussed, but without taking background fluctuations into account.

2. Formulation of the problem

Let us consider a counting experiment in which one wants to measure the signal counts in presence of a background, that is measured in a subsidiary experiment. Such a situation can happen when measuring the activity of a weak radioactive source in presence of background with a counting device, like a Geiger-Muller detector. In this case two independent measurements are carried out, one for the signal and one for the background, and the results have to be properly combined. Another example is an experiment in which the signal and the background are evaluated looking at two different space regions. Also in this case, the signal is evaluated by combining the counts measured in the signal and background regions.\footnote{We prefer to use the word “combine” instead of “subtract” because in many cases the signal cannot be evaluated by simply subtracting the counts in the background region from the ones in the signal region. This may happen for instance when the counts in the background region are more than the ones in the signal region.}

In the following we shall denote with \( n \) and \( m \) the number of counts measured respectively in the signal and in the background regions. We will also indicate with \( s \) and \( b \) the true values of the signal and of the background counts respectively, and we will assume that the true value of the background counts in the signal region is given by \( c b \), where \( c \) is a constant value that is assumed to be exactly known. Let us consider, as a first example, an experiment to measure the activity of a radioactive source. In this case \( n \) and \( m \) are respectively the counts recorded during a time interval \( T_s \) in presence of the source, and during a time interval \( T_b \) when the source has been removed. The value of \( c \) can be estimated as \( c = T_s / T_b \), and the source activity can be evaluated as \( s / T_s \) once the value of \( s \) has been measured. A second example, taken from the astrophysics, is the measurement of the photon flux of a point source in the sky with a gamma-ray detector. In this case the signal region can be a cone centered on the source direction with a
given angular aperture, while the background region can be an annulus far away from the source. In this case $c$ can be evaluated as the ratio between the solid angles of the signal region and of the background region, eventually multiplied for the ratio between the live times of the two regions.

Under the above assumptions, the probability of measuring $m$ counts in the background region will be a Poisson distribution with mean value $b$, i.e.:

$$p(m|b) = e^{-b} \frac{b^m}{m!}$$  \hspace{1cm} (1)

while the probability of measuring $n$ counts in the signal region will be a Poisson distribution with mean value $s + cb$, i.e.:

$$p(n|s, b) = e^{-(s+cb)} \frac{(s + cb)^n}{n!}$$  \hspace{1cm} (2)

Since the two measurements are independent, the joint p.d.f. for $n$ and $m$ will be given by:

$$p(n, m|s, b) = e^{-(s+cb)} \frac{(s + cb)^n}{n!} e^{-b} \frac{b^m}{m!}$$  \hspace{1cm} (3)

Our problem is that of evaluating a Bayesian confidence interval (or an upper limit) for the parameter $s$, independently on $b$. An analogous problem is discussed in the textbook [6] and in ref. [1], where the background value is assumed to be exactly known. Our discussion will be therefore a generalization of refs. [1, 6]. A possible solution of the problem taking background fluctuations into account is given in ref. [7]. However, in ref. [7], a gaussian p.d.f. is used to model the background, the results being valid in the case of large counts. A similar problem is also illustrated in ref. [8], where Bayesian confidence interval are evaluated for a Poisson signal with known background and with fluctuations on the detection efficiency.

Another possible solution to our problem is also given in refs. [4, 5], where a frequentist approach is followed with the application of the profile likelihood method. However, while the procedure described in refs. [4, 5] requires some adjustments to handle the cases when $n < cm$ or when either $n = 0$ or $m = 0$, in our method the treatment of these cases is straightforward and does not require any adjustment. In fact, the formulas that will be derived in sections 3 and 4 are valid for all the values of $n$ and $m$. 

4
3. The Bayesian approach

The implementation of the Bayesian approach requires the “probabilistic inversion” of eq. 3 i.e. the evaluation of the conditional p.d.f. \( p(s, b|n, m) \) starting from \( p(n, m|s, b) \) and applying the Bayes’ theorem.

3.1. Choice of the priors

The application of the Bayes’ theorem starts from the assumption of a prior p.d.f. for both the random variables \( s \) and \( b \). In the following we will assume that \( s \) and \( b \) are independent.

For the true background value \( b \) we will assume a uniform prior:

\[
\pi(b) = \begin{cases} 
\pi_0, & b \geq 0 \\
0, & b < 0 
\end{cases} \tag{4}
\]

where \( \pi_0 > 0 \) is a constant. The parameter \( b \) is only constrained to be non-negative.

On the other hand, for the signal true value \( s \) we will assume a prior p.d.f. given by:

\[
\pi(s) = \begin{cases} 
ks^{-\alpha}, & s \geq 0 \\
0, & s < 0 
\end{cases} \tag{5}
\]

with \( k > 0 \). Also in this case, the only constraint on the parameter \( s \) is that it must be non-negative.

It is worth to point out at this stage that both \( \pi(b) \) and \( \pi(s) \) defined in eqs. 4 and 5 are improper priors \( \footnote{It’s easy to show that \( \int_0^\infty \pi(b) \, db = \infty \) and \( \int_0^\infty \pi(s) \, ds = \infty \) for any value of \( \alpha \).} \) and then they can lead to posterior p.d.f.s that are not normalizable. In particular, as it will be shown in sec. 3.3, this happens when setting \( \alpha \geq 1 \). Our calculations will therefore be valid only for \( \alpha < 1 \).

The uniform prior for the signal is obtained by setting \( \alpha = 0 \) and represents the natural choice when the experimenter does not have any model for the signal. On the other hand, the choice of a power-law prior with \( \alpha > 0 \) will reflect the experimenter’s belief that small signal values are more likely than larger ones. It is also possible to choose negative values of \( \alpha \); this choice, that is rather uncommon, would favour larger signal values with respect to smaller ones \( \footnote{Let \( s_1 < s_2 \) be two possible signal values and let us consider the intervals \([s_1, s_1 + \Delta s]\).} \)
3.2. Evaluation of the background posterior p.d.f.

Applying the Bayes’ theorem and using for $\pi(b)$ the expression in eq. 4 it is possible to obtain the following equation:

$$
p(b|m) = \frac{p(m|b)\pi(b)}{\int p(m|b)\pi(b)db} = \frac{p(m|b)}{\int_0^\infty p(m|b)db} \quad \text{(6)}
$$

Finally, replacing $p(m|b)$ with its expression given in eq. 1, it is straightforward to obtain the final result:

$$
p(b|m) = e^{-b} \frac{b^m}{m!} \quad \text{(7)}
$$

Note that even though the expression of $p(b|m)$ in eq. 7 is the same as that of $p(m|b)$ in eq. 1 their meanings are completely different. In fact, the random variable in eq. 1 is $m$, and the formula tells that $m$ follows a Poisson distribution; on the other hand, the random variable in eq. 7 is $b$, and the formula tells that $b$ follows a Gamma distribution.

Using eq. 7 one can also easily evaluate the average value of $b$ and its standard deviation. It is easy to show that $\langle b \rangle = m + 1$ and $\sigma_b = \sqrt{m + 1}$.

3.3. Evaluation of the signal posterior p.d.f.

The Bayes’ theorem can be applied to get the joint posterior p.d.f. for both $s$ and $b$:

$$
p(s, b|n, m) = \frac{p(n, m|s, b)\pi(s)\pi(b)}{\int ds \int db p(n, m|s, b)\pi(s)\pi(b)} \quad \text{(8)}
$$

Replacing $p(n, m|s, b)$ with its expression given in eq. 3 and $\pi(b)$ and $\pi(s)$ with their expressions given in eqs. 1 and 2 respectively, eq. 8 can be rewritten as follows:

$$
p(s, b|n, m) = \frac{e^{-s-(c+1)b}b^m s^{-\alpha} (s + cb)^n}{\int_0^\infty ds \int_0^\infty db e^{-s-(c+1)b}b^m s^{-\alpha} (s + cb)^n} \quad \text{(9)}
$$

and $[s_2, s_2 + \Delta s]$. The ratio between the probabilities of finding $s$ in the two intervals is given by $R = P(s_1 < s < s_1 + \Delta s)/P(s_2 < s < s_2 + \Delta s) = (s_2/s_1)^\alpha$. When $\alpha > 0$ ($\alpha < 0$) then $R > 1$ ($R < 1$) and small (large) signal values are more likely than larger (smaller) values. For a discussion about the choice of priors in the Bayesian approach see for instance the textbook [6].
Indicating with $N$ the denominator in the right-hand side of eq. 9, it can be rewritten as:

$$N = \int_0^\infty db \ b^m e^{-(c+1)b} \int_0^\infty ds \ e^{-s} s^{-\alpha} (s + cb)^n = \int_0^\infty db \ b^m e^{-(c+1)b} f(b) \quad (10)$$

where we have indicated with $f(b)$ the result of the integral in $ds$, that can be seen as a function of the variable $b$.

Applying the binomial theorem, the term $(s + cb)^n$ in the expression of $f(b)$ can be expanded as follows:

$$(s + cb)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} s^k (cb)^{n-k} \quad (11)$$

Using this result, the expression of $f(b)$ becomes:

$$f(b) = n! \sum_{k=0}^{n} \frac{(cb)^{n-k}}{k!(n-k)!} \int_0^\infty e^{-s} s^{k-\alpha} ds \quad (12)$$

and, taking into account the definition of the Gamma function the previous equation can be written as:

$$f(b) = \Gamma(n + 1) \sum_{k=0}^{n} \frac{(cb)^{n-k}}{k!(n-k)!} \frac{\Gamma(k - \alpha + 1)}{\Gamma(k + 1)\Gamma(n - k + 1)} \quad (13)$$

Introducing the expression of $f(b)$ given by eq. 13 in the expression of $N$ given by eq. 10 we get the following result:

$$N = \Gamma(n + 1) \sum_{k=0}^{n} \frac{\Gamma(k - \alpha + 1)}{\Gamma(k + 1)\Gamma(n - k + 1)} \int_0^\infty db \ b^{m+n-k} e^{-(c+1)b} \quad (14)$$

By making a proper change of variable, the integral in the right-hand side eq. 14 can be expressed in terms of a Gamma function. Hence eq. 14 can be rewritten as follows:

---

7 The Gamma function is defined as $\Gamma(z) = \int_0^\infty dt e^{-t} t^{z-1}$. It can be shown that $\Gamma(z) = (z - 1)!$ when $z$ is a positive integer.
\[ N = \frac{\Gamma(n + 1)}{(c + 1)^{m+1}} \sum_{k=0}^{n} \frac{\Gamma(k - \alpha + 1)\Gamma(m + n - k + 1)}{\Gamma(k + 1)\Gamma(n - k + 1)} \left( \frac{c}{c + 1} \right)^{n-k} \]  

(15)

The joint posterior p.d.f. for \( s \) and \( b \) is then given by:

\[ p(s, b|n, m) = \frac{1}{N} e^{-s-(c+1)b} b^m s^{-\alpha} (s + cb)^n \]  

(16)

with the expression of \( N \) given in eq. 15.

To evaluate the marginal p.d.f. for \( s \) we need to integrate the joint p.d.f. with respect to \( b \):

\[ p(s|n, m) = \int_0^\infty p(s, b|n, m)db = \frac{1}{N} e^{-s} s^{-\alpha} \int_0^\infty db e^{-(c+1)b} b^m (s + cb)^n \]  

(17)

Indicating with \( g(s) \) the integral in the right-hand side of eq. 17 it can be evaluated in a similar way to that used to calculate \( f(b) \). It is easy to show that:

\[ g(s) = \frac{\Gamma(n + 1)}{(c + 1)^{m+1}} \sum_{k=0}^{n} \frac{\Gamma(m + n - k + 1)}{\Gamma(k + 1)\Gamma(n - k + 1)} \left( \frac{c}{c + 1} \right)^{n-k} s^k \]  

(18)

Introducing in eq. 17 the expression of \( g(s) \) given by eq. 18 and the expression of \( N \) given by eq. 15 it is possible to show that the posterior p.d.f. for \( s \) is given by:

\[ p(s|n, m) = \frac{\sum_{k=0}^{n} \frac{\Gamma(m + n - k + 1)}{\Gamma(k + 1)\Gamma(n - k + 1)} \left( \frac{c}{c + 1} \right)^{n-k} s^k}{\sum_{k=0}^{n} \frac{\Gamma(k - \alpha + 1)\Gamma(m + n - k + 1)}{\Gamma(k + 1)\Gamma(n - k + 1)} \left( \frac{c}{c + 1} \right)^{n-k}} \]  

(19)

It is worth to point out here that eq. 19 is valid only for \( \alpha < 1 \). In fact, if \( \alpha \geq 1 \), the argument \( k - \alpha + 1 \) of the Gamma function in the sum in the right-hand side of the denominator will be null or negative, and consequently the Gamma function either will assume negative values or will diverge. As discussed in sec. 3.1 this behaviour is due to the fact that since the posterior
p.d.f. \( p(s|n, m) \) is obtained starting from improper priors, its normalization is not guaranteed.

Fig. 1 shows the posterior p.d.f.s for the signal \( s \) evaluated for some different values of \( n \) and \( m \) in the case \( c = 1 \), i.e. when the background region has the same size as the signal region. The calculations have been performed in the cases \( \alpha = 0.5 \) (small signal expected), \( \alpha = 0 \) (uniform prior) and \( \alpha = -0.5 \) (large signal expected). The differences between the three p.d.f.s become negligible when \( n \) is larger than \( cm \).

From eq. 19 one can also easily calculate the moments of the p.d.f. \( p(s|n, m) \). It’s easy to show that the \( h \)th moment of the p.d.f. is given by:

\[
\langle s^h | n, m \rangle = \frac{\sum_{k=0}^{n} \frac{\Gamma(k - \alpha + 1 + h)\Gamma(m + n - k + 1)}{\Gamma(k + 1)\Gamma(n - k + 1)}}{\sum_{k=0}^{n} \frac{\Gamma(k - \alpha + 1)\Gamma(m + n - k + 1)}{\Gamma(k + 1)\Gamma(n - k + 1)}} \left( \frac{c}{c + 1} \right)^{n-k}
\]

In particular, eq. 20 allows to calculate the expectation value of \( s \), i.e. \( \langle s \rangle \) and its variance, that can be evaluated as \( \text{var}(s) = \langle s^2 \rangle - \langle s \rangle^2 \).

Finally, we can consider the limiting case with absence of background, that can be obtained by setting \( c = 0 \) (background region with null size). In this limit, all the terms with \( c/(c+1) \)^{n-k} in both the summations of eq. 19 will vanish, but the ones with \( k = n \), where \( c/(c+1) \)^{n-k} = 1. Hence, in this case the posterior p.d.f. for the signal simplifies to:

\[
p(s|n, m) = \frac{s^{n-\alpha}e^{-s}}{\Gamma(n - \alpha + 1)}
\]

and, as expected, does not depend on \( m \). In particular, in the case \( \alpha = 0 \) the posterior p.d.f. for the signal becomes a Gamma distribution.

4. Evaluation of the upper limits on the signal

The posterior p.d.f. for the signal given by eq. 19 can be used to evaluate Bayesian confidence intervals for \( s \). In particular, in this section, we will apply the result of eq. 19 to the calculation of upper limits on \( s \).
To evaluate the upper limit $s_u$ at the confidence level $1 - \beta$ we have to solve the integral equation:

$$\int_0^{s_u} p(s|n, m) ds = 1 - \beta$$  \hspace{1cm} (22)

Taking advantage of the fact that

$$\int_0^{s_u} s^{k-\alpha} e^{-s} ds = \gamma(k - \alpha + 1, s_u)$$  \hspace{1cm} (23)

where we have indicated with $\gamma$ the incomplete Gamma function, eq. 22, can be rewritten as:

$$1 - \beta = \frac{\sum_{k=0}^{n} \frac{\gamma(k - \alpha + 1, s_u) \Gamma(m + n - k + 1)}{\Gamma(k + 1) \Gamma(n - k + 1)} \left( \frac{c}{c+1} \right)^{n-k}}{\sum_{k=0}^{n} \frac{\Gamma(k - \alpha + 1) \Gamma(m + n - k + 1)}{\Gamma(k + 1) \Gamma(n - k + 1)} \left( \frac{c}{c+1} \right)^{n-k}}$$  \hspace{1cm} (24)

Eq. 24 can be solved numerically and allows to obtain the Bayesian upper limits on $s$ at the confidence level $1 - \beta$ for any values of $n$ and $m$, given the values of $c$ and $\alpha$. We have performed our calculations using the CERN ROOT package \[9\]. In particular, we have implemented a code that allows to evaluate the numerical solutions $s_u$ of eq. 24 for any given value of $\beta$ with the Brent’s method, using the ROOT built-in tools.

In fig. 2 the upper limits on the signal at 90% confidence level are shown as a function of the observed counts in the signal ($n$) and background ($m$) regions in the case $c = 1$ for three different values of $\alpha$. As expected, the choice of the power law index $\alpha$ in the signal prior p.d.f. will affect the result on the upper limits. In particular, the upper limits on the signal will increase with decreasing $\alpha$. This result is in agreement with the fact that when $\alpha$ is positive and close to 1 small signal values are expected while, on the other hand, when $\alpha$ is negative, large signal values are expected. As pointed out in sec. 1, this is a general issue of the Bayesian approach, the results being influenced by the initial belief of the experimenter.

---

8The incomplete Gamma function is defined as $\gamma(z, x) = \int_0^x dt e^{-t} t^{z-1}$. According to this definition $\gamma(z, \infty) = \Gamma(z)$. 10
Fig. 3 shows the upper limits on the signal at 90% confidence level as a function of the background events $m$ for several values of the signal events $n$. The calculation has been performed in the case $c = 1$ with the uniform prior ($\alpha = 0$). For any given value of $n$, the upper limit on the signal decreases with increasing $m$, in agreement with the fact that (in the cases when $n > cm$) a rough estimate of $s$ is given by $n - cm$ and the upper limit is expected to be proportional to $n - cm$. It has also to be pointed out that in the case $n = 0$, i.e. when no events are observed in the signal region, the upper limit is always equal to 2.30, independently on $m$.

Finally, fig. 4 shows the upper limits on the signal at 90% confidence level as a function of the signal events $m$ for several values of the background events $m$. Again, the calculation has been performed in the case $c = 1$ with the uniform prior ($\alpha = 0$). The upper limits increase with increasing $n$, and the trend of the curves becomes almost linear when $n > cm$.

4.1. Study of the frequentist coverage

To study the frequentist coverage of the upper limits obtained with our procedure we implemented a dedicated simulation. For simplicity we studied only the case with $c = 1$, when the signal and the background regions have the same sizes.

For any given pair of values of $s$ and $b$, a sample of $10^5$ events was simulated, each corresponding to the outcome of an experiment. Each event consists of a couple of random integer numbers $(n, m)$, representing respectively the counts in the signal and in the background regions, that are generated according to the p.d.f. of eq. 3. For each couple of counts $(n, m)$ the corresponding upper limits on $s$ at 90% confidence level were evaluated by solving eq. 24 for different values of $\alpha$. The coverage was then evaluated as the fraction of events with an associated upper limits less than the true value of the signal $s$.

Fig. 5 shows, as an example, the results obtained with a simulated sample of events with $s = 3.5$ and $b = 2$. The distributions of the upper limits at 90% confidence level are shown for $\alpha = 0.5$, $\alpha = 0$ and $\alpha = -0.5$. As expected,

---

9It is evident that when no events (or a few events) are observed in the signal region and a large number of events are observed in the background region the prior p.d.f. given by eq. 5 is not adequate. In these cases a different prior p.d.f. should be used, that allows for negative values of $s$. 

11
the coverage increases with decreasing $\alpha$, since lower values of $\alpha$ correspond to more conservative upper limits.

We have studied the frequentist coverage of our upper limits as a function of both $s$ and $b$. The results are summarized in Fig. 6 where the coverage is plotted as a function of $s$ in the cases when $b = 0, b = 1, b = 1.5, b = 2, b = 5$ and $b = 10$. As expected, the frequentist coverage decreases with increasing $\alpha$. The choice of a uniform prior for the signal, i.e. $\alpha = 0$ guarantees a coverage that is larger than 90% for low signal values, and oscillates around 90% with increasing $s$. On the other hand, the choice of a less conservative prior with $\alpha = 0.5$ does not seem to affect significantly the coverage in case of an high background level.

5. Conclusions

We have developed a procedure that, following the Bayesian approach, allows to evaluate confidence intervals on the signal in experiments with background where both signal and background are modeled by the Poisson statistics. The implementation of the method is quite simple from the mathematical point of view, and does not require any adjustments to treat the cases when less events are observed in the signal region than those in the background region. The results obtained when our procedure is applied to the calculation of upper limits have been also discussed. The frequentist coverage of the upper limits evaluated with this procedure has been also studied.

References

[1] G. Cowan et al. (Particle Data Group), Journ. Phys. G 37 (2010), 075021 see also http://pdg.lbl.gov/2010/reviews/rpp2010-rev-statistics.pdf

[2] J. Neyman, Phil. Trans. Royal Soc. London, Series A 236, 333 (1937)

[3] G. J. Feldman, R. D. Cousins, “Unified approach to the classical statistical analysis of small signals”, Phys. Rev. D 57 (1998), 3873

[4] W. A. Rolke, A. M. Lopez, “Confidence intervals and upper bounds for small signals in presence of background noise”, Nucl. Inst. Meth. A458 (2000), 745
[5] W. A. Rolke, A. M. Lopez, J. Conrad, “Limits on Confidence Intervals in the Presence of Nuisance Parameters”, ArXiv:physics/0403059 (2009)

[6] G. Cowan, “Statistical Data Analysis”, Oxford University Press (1998), ISBN 0-19-850155-2

[7] O. Helene, “Determination of the upper limit of a peak area”, Nucl. Inst. Meth. 212 (1983), 319

[8] J. Heinrich et al., “Interval estimation in the presence of nuisance parameters. Bayesian approach”, ArXiv:physics/0409129 (2004)

[9] R. Brun and F. Rademakers, “ROOT - An object oriented data analysis framework”, Nucl. Inst. Meth. A389 (1997), 81 (see also http://root.cern.ch)
Figure 1: The plots show the posterior p.d.f.s for the signal \( p(s|n, m) \) for different values of \( n, m \) and \( \alpha \) in the case \( c = 1 \). Going from top to bottom, the plots correspond to \( n = 0, 1, 2, 5 \); going from left to right the plots correspond to \( m = 0, 1, 2, 5 \). The p.d.f.s have been evaluated assuming \( \alpha = 0.5 \) (continuous lines), \( \alpha = 0 \) (dashed lines) and \( \alpha = -0.5 \) (dotted lines). In each plot three vertical lines are also drawn with the same styles specified above, each indicating the average value of \( s \) from the corresponding p.d.f.s.
Figure 2: The plots show the values of the signal upper limits at 90% confidence level as a function of the counts observed in the signal and background regions, $n$ and $m$, in the case $c = 1$. Going from top to bottom, the plots correspond to $\alpha = 0.5$, $\alpha = 0$ and $\alpha = -0.5$. 
Figure 3: The plot shows the values of the signal upper limits at 90% confidence level as a function of the counts observed in the background region, $m$, in the case $c = 1$ and $\alpha = 0$, for some different values of the counts in the signal region. The calculation has been performed for $n = 0$ ($\bullet$), 1 ($\blacksquare$), 2 ($\blacktriangle$), 3 ($\blacktriangledown$), 5 ($\circ$), 10 ($\square$) and 20 ($\blacktriangledown$).
Figure 4: The plot shows the values of the signal upper limits at 90% confidence level as a function of the counts observed in the signal region, $n$, in the case $c = 1$ and $\alpha = 0$, for some different values of the counts in the background region. The calculation has been performed for $m = 0$ (•), 1 (■), 2 (▲), 3 (▼), 5 (◦), 10 (□) and 20 (△).
Figure 5: The plots show the distributions of the signal upper limits at 90% confidence level obtained from a sample of $10^5$ simulated experiments with $s = 3.5$, $b = 2$ and $c = 1$. Going from top to bottom, the plots correspond to the upper limits evaluated by setting $\alpha = 0.5$, $\alpha = 0$ and $\alpha = -0.5$. The dashed lines indicate the true value of the signal. The coverage is graphically represented by the area of the histogram at the right of the dashed line normalized to the total area.
Figure 6: The plots show the frequentist coverage of the upper limits at 90% confidence level as a function of the true signal value $s$ for different values of $b$ with $c = 1$. From top left to bottom right the plots correspond to the cases $b = 0$, $b = 1$, $b = 1.5$, $b = 2$, $b = 5$ and $b = 10$. The upper limits have been evaluated by setting $\alpha = 0.5$ (continuous lines), $\alpha = 0$ (dashed lines) and $\alpha = -0.5$ (dotted lines). The dash-dotted lines indicate the 90% coverage.