Incorporating Systematics and Statistical Uncertainties into Exclusion Limits

S.I.Bityukov
IHEP, Protvino, 141281, Russia
N.V. Krasnikov
INR RAS, Prospect 60-letiya Octyabrya 7a, Moscow, 117312, Russia

It is important to know the range in which a planned experiment can exclude the presence of a signal at a given confidence level $1-\epsilon$. We propose to use the probability of making a correct decision in future hypothesis testing about observation of a signal in planned experiments as the confidence level in the determination of exclusion limits.

1. INTRODUCTION

It is important to know the range in which a planned experiment can exclude the presence of a signal at a given confidence level $1-\epsilon$. It means that we plan to have an uncertainty not more than ϵ in our conclusion about observation or non-observation of a signal. The estimation of this uncertainty in future hypothesis testing allows the determination of exclusion limits.

Let us consider a statistical hypothesis

$$H_0\text{: } \textit{new physics is present in Nature}$$

against an alternative hypothesis

$$H_1\text{: } \textit{new physics is absent in Nature.}$$

The value of uncertainty is defined by the probability to reject H_0 when it is true (Type I error)

$$\alpha = P(reject\ H_0|H_0\ is\ true)$$

and the probability to accept H_0 when H_1 is true (Type II error)

$$\beta = P(accept\ H_0|H_0\ is\ false).$$

There are different approaches to the construction of exclusion regions in planned experiments [1–5]. For example, Hernandez et al. [1] propose the following criteria for the definition of exclusion limits:

$$\beta < \Delta \text{ and } \frac{\alpha}{1-\beta} < \epsilon \ ,$$

i.e. the experiment will observe with probability at least $1-\Delta$ at most a number of events such that the limit obtained at the $1-\epsilon$ confidence level excludes the corresponding signal.

In a recent note [4], the authors also propose to construct the exclusion region using two values: the magnitudes of significance level (α) and power of test ($1-\beta$) in hypothesis testing.

We propose to construct the exclusion region using only one value: the estimator

$$\hat{\kappa} = \frac{\hat{\alpha} + \hat{\beta}}{2} \qquad (1)$$

of the uncertainty [3] $\kappa = \alpha + \beta$, when testing H_0 versus H_1 with an equal-tailed test. It is the probability of making an incorrect choice in favor of one of the hypotheses in future hypothesis testing. Here $\hat{\alpha}$ and $\hat{\beta}$ are the estimators of possible Type I error (α) and Type II error (β) obtained by direct calculations. Then $\hat{\kappa}$ is independent of which hypothesis is chosen as H_0, and which is H_1. The estimator $\hat{\kappa}$ differs from the estimator $\tilde{\kappa} = \hat{\kappa}/(1-\hat{\kappa})$ which was used in [6]. For Poisson distributions, we propose to use an equal probability test [7] as a good approximation to the equal-tailed test for estimation of $\hat{\kappa}$. Note that our approach is close to that proposed in ref [5].

Ref [8] suggests a Monte Carlo method for taking into account several types of systematics in construction of confidence limits. We consider here systematics which do not have statistical properties and hence cannot be taken into account by traditional methods for estimating their influence on exclusion limits.

2. THE PROBABILITY OF MAKING A CORRECT DECISION

Suppose that the probability of observing n events in an experiment is described by the function $f(n;\mu)$ with parameter μ, and that we know the expected numbers of signal and background events (μ_s and μ_b respectively).

Let us specify what we mean by the probability of making a correct decision about the presence or absence of a new phenomenon in a planned experiment. Let us define the criterion for the hypothesis choice and calculate the probability of making a correct decision. This is possible, because we construct the critical region in such a way that the probability of an incorrect choice in favor of one of the hypotheses is independent of whether H_0 or H_1 is true. We consider two conditional distributions of probabilities

$$\begin{cases} f_0(n) = f(n; \mu_s + \mu_b), \\ f_1(n) = f(n; \mu_b) \end{cases} \quad (2)$$

We suppose that any prior suppositions about H_0 and H_1 can be included in $f_0(n)$ and $f_1(n)$. After choosing a critical region in some way, we can estimate the Type I ($\hat{\alpha}$) and Type II errors ($\hat{\beta}$). In the case of applying the equal-tailed test, their combination Eq. 1 is the probability of making incorrect choice in favour of one of the hypotheses.

In actuality we must estimate the random value $\kappa = \alpha + \beta = \hat{\kappa} + e$, where $\hat{\kappa}$ is a constant and e is a stochastic term. α is the fraction of incorrect decisions if H_0 is true. Then β is absent because H_1 is not realised in Nature. Correspondingly, β is the fraction of incorrect decisions if H_1 takes place; then α is absent. If H_0 is true, the Type I error equals $\hat{\alpha}$ and the error of our estimator (Eq. 1) is $\hat{e} = \hat{\kappa} - \hat{\alpha} = \dfrac{\hat{\alpha} + \hat{\beta}}{2} - \hat{\alpha} = -\dfrac{\hat{\alpha} - \hat{\beta}}{2}$. Similarly, if H_1 is true, the Type II error equals $\hat{\beta}$ and the error of the estimator is $\hat{e} = \hat{\kappa} - \hat{\beta} = \dfrac{\hat{\alpha} + \hat{\beta}}{2} - \hat{\beta} = \dfrac{\hat{\alpha} - \hat{\beta}}{2}$. Thus the stochastic term takes the values $\pm\dfrac{\hat{\alpha} - \hat{\beta}}{2}$. If we require $\hat{\alpha} = \hat{\beta}$, both errors of the estimation are equal to 0 $(\hat{\kappa} - \hat{\alpha} = \hat{\kappa} - \hat{\beta} = 0)$. As a result the estimator (Eq. 1) gives the probability of making an incorrect decision in future hypothesis testing.

Accordingly, $1 - \hat{\kappa}$ is the probability to make a correct choice with the given critical value. Note that the equal probability test gives results close to the equal-tailed test in the case of Poisson distributions and we use an equal probability test henceforth.

Let the probability of observing n events in an experiment be described by a Poisson distribution with parameter μ, i.e.

$$f(n; \mu) = \frac{\mu^n}{n!} e^{-\mu}. \quad (3)$$

Then the Type I and II errors can be written as:

$$\begin{cases} \hat{\alpha} = \sum_{i=0}^{n_c} f(i; \mu_s + \mu_b) = \sum_{i=0}^{n_c} f_0(i), \\ \hat{\beta} = 1 - \sum_{i=0}^{n_c} f(i; \mu_b) = 1 - \sum_{i=0}^{n_c} f_1(i), \end{cases} \quad (4)$$

where n_c is a critical value.

$\hat{\kappa}$ has a minimum if we choose n_c such that $f_0(n_c) = f_1(n_c)$. (For the discrete Poisson distribution, $n_c =$ largest integer i such that $f_0(i) \leq f_1(i)$). This follows directly from

$$\hat{\kappa} = \frac{\hat{\alpha} + \hat{\beta}}{2} = \frac{1}{2}(1 - \sum_{i=0}^{n_c}(f_1(i) - f_0(i))). \quad (5)$$

The value of $\hat{\kappa}$ decreases as i increases from 0 up to n_c. As soon as $f_0(i) > f_1(i)$, the value of $\hat{\kappa}$ increases. Thus $\hat{\kappa}$ will have its minimal value when applying the equal probability test [3], and

$$n_c = [\frac{\mu_s}{ln(\mu_s + \mu_b) - ln(\mu_b)}], \quad (6)$$

where square brackets mean the integer part of a number.

3. SIGNAL SIGNIFICANCE AND EXCLUSION LIMITS

$\hat{\kappa}$ plays the role of ϵ in the definition of the confidence level and, correspondingly, of the significance S of an excess of signal events above background [10] in planned experiments. In the case of Poisson distributions the definition of significance as

$$\hat{\kappa} = \frac{1}{\sqrt{2\pi}} \int_{S_{12}}^{\infty} e^{-\frac{x^2}{2}} dx. \quad (7)$$

leads to the formula [3, 6, 7]

$$S_{12} = \sqrt{\mu_s + \mu_b} - \sqrt{\mu_b}. \quad (8)$$

A factor two is needed (i.e. $S = 2S_{12}$) to correspond with common practice. As shown in [11] this approximation has good statistical properties as the significance for Poisson distributions.

Let us define the exclusion limit of a planned experiment: *the planned experiment can exclude the presence of a signal at a given confidence level ϵ if the probability of a wrong decision about the observation of the signal will be equal to or less than ϵ.*

Thus to determine the exclusion limit μ_s we must solve equations (9) for μ_s, with given ϵ and μ_b, where n_c is determined by eqn. (6).

$$\begin{cases} \hat{\alpha} = \sum_{i=0}^{n_c} f(i; \mu_s + \mu_b) = \sum_{i=0}^{n_c} f_0(i), \\ \hat{\beta} = 1 - \sum_{i=0}^{n_c} f(i; \mu_b) = 1 - \sum_{i=0}^{n_c} f_1(i), \\ \hat{\kappa} = \dfrac{\hat{\alpha} + \hat{\beta}}{2} \leq \epsilon. \end{cases} \quad (9)$$

4. SYSTEMATICS OF THEORETICAL ORIGIN

We consider here forthcoming experiments to search for new physics. In this case we must take into account the systematic uncertainty which have theoretical origin without statistical properties. For example, two loop corrections for most reactions at present are not known. It means that we can only estimate the scale of influence of background uncertainty on the observability of signal, i.e. we can determine the admissible level of uncertainty in theoretical calculations for a given experiment proposal.

Suppose the background cross section is known to be in the interval $(\sigma_b, \sigma_b(1+\delta))$, and hence the average number of background events lies in the interval $(\mu_b, \mu_b(1+\delta))$.

As we know nothing about possible values within this range, we consider the worst case. Eqs. (9) for the exclusion limit are modified to

$$\begin{cases} n_c = [\dfrac{\mu_{lim}}{ln(\mu_{lim} + \mu_b(1+\delta)) - ln(\mu_b(1+\delta))}], \\ \hat{\alpha} = \displaystyle\sum_{i=0}^{n_c} f(i; \mu_{lim} + \mu_b(1+\delta)) \\ \hat{\beta} = 1 - \displaystyle\sum_{i=0}^{n_c} f(i; \mu_b(1+\delta)) \\ \hat{\kappa} = \dfrac{\hat{\alpha} + \hat{\beta}}{2} \leq \epsilon, \end{cases} \quad (10)$$

where μ_{lim} is the exclusion limit at the worst background level $\mu_b \cdot (1+\delta)$.

5. CONCLUSIONS

We have considered a possible approach for the determination of exclusion limits for planned experiments. We propose to use the probability of making an incorrect decision in future hypothesis testing about the observation of the new phenomenon. We also propose an approach to take into account systematics of a theoretical origin.

Acknowledgments

We would like to thank Louis Lyons, V.A. Matveev and V.F. Obraztsov for their interest and useful comments. We are very grateful to A. Nikitenko for very fruitful discussions. We also wish to thank S.S. Bityukov, R.D. Cousins, Yu.P. Gouz, G. Kahrimanis, V.V. Smirnova, V.A. Taperechkina for discussions and E.A. Medvedeva for help in preparing the paper. S.B. is grateful to Arla LeCount and Jim Linnemann. This work has been supported by grant RFBR 03-02-16933.

References

[1] J.J.Hernandez, S.Navas and P.Rebecchi, "Estimating exclusion limits in prospective studies of searches", Nucl.Instr.&Meth. A **378** (1996) 301.

[2] T.Tabarelli de Fatis and A.Tonazzo, "Expectation values of exclusion limits in future experiments", Nucl.Instr.&Meth. A **403**(1998) 151.

[3] S.I. Bityukov and N.V. Krasnikov, "New physics discovery potential in future experiments," Modern Physics Letters **A13** (1998) 3235.

[4] L. Fleysher et al., "Exclusion regions and their power", physics/0308067, August, 2003.

[5] G. Punzi, "Sensitivity of searches for new signals and its optimization", these Proceedings.

[6] S.I. Bityukov and N.V. Krasnikov, Proc. of Workshop on Confidence limits, eds. F. James, L. Lyons, Y. Perrin, CERN-2000-005, 2000, p. 219.

[7] S.I.Bityukov and N.V. Krasnikov, Proc. of Conf. "Advanced statistical techniques in particle physics", eds. M.R. Whalley, L. Lyons, Durham, UK, 2002, p.77; hep-ph/0204326.

[8] J.Conrad et al., "Coverage of Confidence Intervals for Poisson Statistics in Presence of Systematic Uncertainties", Proc. of Conf. "Advanced statistical techniques in particle physics", eds. M.R. Whalley, L. Lyons, Durham, UK, 2002, p.58.

[9] S.I.Bityukov and N.V.Krasnikov, *On the observability of a signal above background*, Nucl.Instr.&Meth. **A452** (2000) 518.

[10] A.G.Frodesen, O.Skjeggestad and H.Toft, *Probability and Statistics in Particle Physics,* UNIVERSITETSFORLAGET, Bergen-Oslo-Tromso, 1979, p.408.

[11] V.Bartsch and G.Quast, "Expected signal observability at future experiments", CMS Internal Note 2003/039. August, 2003.