Advanced Statistical Techniques in Particle Physics
18 - 22 March 2002
Grey College, Durham

Uncertainties and Discovery Potential in Planned Experiments

Sergei I. Bityukov
Institute for High Energy Physics, Protvino, Russia
Nikolai V. Krasnikov
Institute for Nuclear Research RAS, Moscow, Russia

- **What do we mean in planned experiments by the statistical significance s of a signal ? In the case being considered there is a mean number n_b of background events and a mean number n_s of signal events is predicted.**
- **How to estimate the possibility of observation of predicted new phenomenon at presence of uncertainties in magnitudes n_s and n_b ?**

Signal and background events are considered as indistinguishable, independent and are taken from Poisson streams of events, i.e.

$$\mathbf{f}(\mathbf{n};\lambda) = \frac{\lambda^{\mathbf{n}}}{\mathbf{n}!}\mathbf{e}^{-\lambda}, \tag{1}$$

where the parameter of Poisson distribution λ is $\lambda = n_b + n_s$ in presence and $\lambda = n_b$ in absence of signal events in Nature.

The case of real experiment with n_{obs} observed events and n_b predicted background events.

Let us the statistical significance s of a signal is defined by [I.Narsky, NIM, A450 (2000) 444]

$$\frac{1}{\sqrt{2\pi}} \int_s^{\infty} e^{-\frac{x^2}{2}} dx = \sum_{k=n_{obs}}^{\infty} f(k; n_b). \tag{2}$$

Here s is a function on n_{obs} and n_b, i.e. $s = s(n_{obs}, n_b)$. The n_{obs} is the given number. The real number of background events is a random value with mean n_b.

Using the language of standard deviations one can say that the case $s \geq 5$ corresponds to new physics discovery, i.e. $s = 5$ corresponds to some critical value n_0 for hypotheses testing about discovery of new Phenomenon.
Let us define also the case $s \geq 2$ as weak evidence and the case $s \geq 3$ as strong evidence in favour of new physics.

The common used significance $S_1 = \dfrac{n_{obs} - n_b}{\sqrt{n_b}} = \dfrac{n_s}{\sqrt{n_b}}$ implies in this sense that the probability to observe $n \geq n_{obs}$ background events is

$$P(n \geq n_{obs}) = \frac{1}{\sqrt{2\pi}} \int_{S_1}^{\infty} e^{-\frac{x^2}{2}} dx. \tag{3}$$

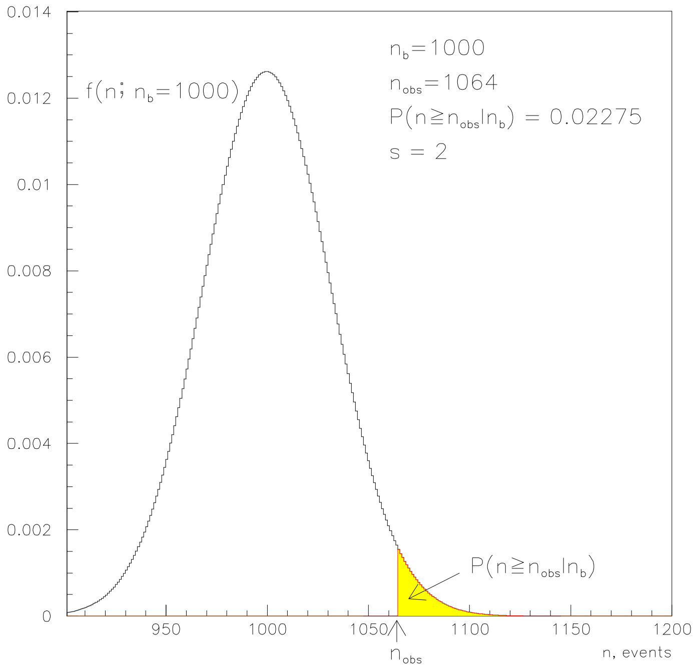

Figure 1: The case $n_b \gg 1$. Poisson distributions with parameters $\lambda = n_b = 1000$ (left). Here $P(n \geq n_{obs}|n_b) = 0.02275$ (i.e. $S_1 = 2$).

Note that to apply the hypotheses testing H_0 (null hypothesis: searched phenomenon exists in nature) versus H_1 (alternative hypothesis: new physics is absent) for comparison of the given number n_{obs} and Poisson distribution with mean n_b we must state null hypothesis.
It is obvious, that critical value to the given hypotheses testing, appropriate s, is n_{obs}. It allows to speak in the language of standard deviations.

Planned experiment. The case of predicted n_s and n_b.

Let us to compare two Poisson distributions with mean n_b and $n_s + n_b$. There are many approaches to choose the critical value of test (test with given significance, test with given power, equal-tailed test, equal-probability test, and so on).

Let us apply the notion of significance for comparison of two distributions.

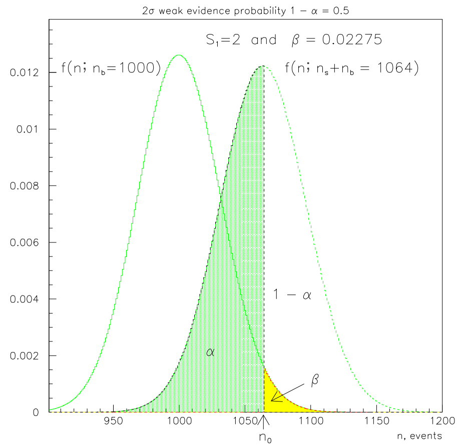

Figure 2: The case $n_b \gg 1$. Poisson distributions with parameters $\lambda = 1000$ (left) and $\lambda = 1064$ (right). Here $1 - \alpha = 0.5$ and $\beta = 0.02275$ (i.e. $S_1 = 2$).

In Figure the case when mean value of $n_0 = n_s + n_b$ correspond 2σ of a deviation from mean value of background n_b is shown. It means that Type I error $\beta = 0.02275$ and Type II error $\alpha = 0.5$.

Let $\Delta = P(n \geq n_0 | n_b)$, where n_0 is a critical value for hypotheses testing. As a result there is system

$$\beta(\mathbf{\Delta}) = \sum_{\mathbf{n}=\mathbf{n_0}(\mathbf{\Delta})+\mathbf{1}}^{\infty} \mathbf{f}(\mathbf{n};\mathbf{n_b}) \leq \mathbf{\Delta}, \tag{4}$$

$$\mathbf{1} - \alpha(\mathbf{\Delta}) = \sum_{\mathbf{n}=\mathbf{n_0}(\mathbf{\Delta})+\mathbf{1}}^{\infty} \mathbf{f}(\mathbf{n};\mathbf{n_s}+\mathbf{n_b}). \tag{5}$$

Acceptance $1-\alpha(\Delta)$ is a probability of discovery ($\Delta = 2.85\cdot 10^{-7}$), or a probability of strong evidence ($\Delta = 1.35\cdot 10^{-3}$), or a probability of weak evidence ($\Delta = 2.275\cdot 10^{-2}$). If $1-\alpha(\Delta) = 0.5$ then, correspondingly, $S_1 = 5$, $S_1 = 3$, and $S_1 = 2$.
If $1-\alpha(\Delta) > 0.5$ we have effective significance s determined by approximate formula (for given β)

$$\mathbf{s} = \mathbf{S_1} - \mathbf{k}(\alpha)\sqrt{\mathbf{1} + \frac{\mathbf{n_s}}{\mathbf{n_b}}} \tag{6}$$

where $k(\alpha)$ can be taken from PDG Table 28.1:
$k(0.5) = 0$; $k(0.25) = 0.66$; $k(0.1) = 1.28$; $k(0.05) = 1.64$.

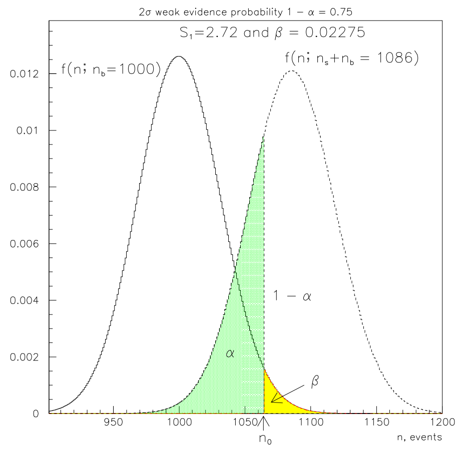

Figure 3: The case $n_b \gg 1$ and $S_1 = 2.72$. Poisson distributions with parameters $\lambda = 1000$ (left) and $\lambda = 1086$ (right). Here $1-\alpha = 0.75$ and $\beta = 0.02275$ (i.e. effective $s = 2$).

As it has been shown in [Bityukov, NIM A452 (2000) 518] the more proper of the significance in planned experiments is

$$\mathbf{2 \cdot S_{12} = 2 \cdot (\sqrt{n_s + n_b} - \sqrt{n_b})}. \tag{7}$$

The generalization of this significance to the case of $1 - \alpha > 0.5$ looks very attractive for approximate estimation of discovery potential

$$\mathbf{s = 2 \cdot (\sqrt{n_s + n_b} - \sqrt{n_b}) - k(\alpha)}. \tag{8}$$

The comparison of these approximate formulae is shown.

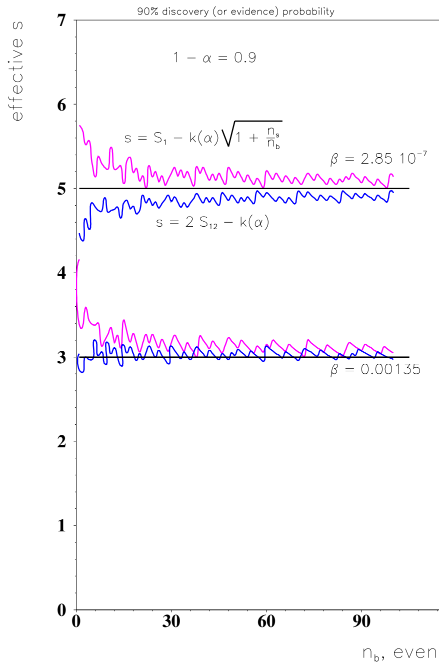

Figure 4: The estimation of effective significance s for given β and $1 - \alpha$.

In this Figure is shown the comparison of the results of the direct calculations the n_s for given n_b in accordance with formulae (4) and (5), i.e. 5σ discovery curve.

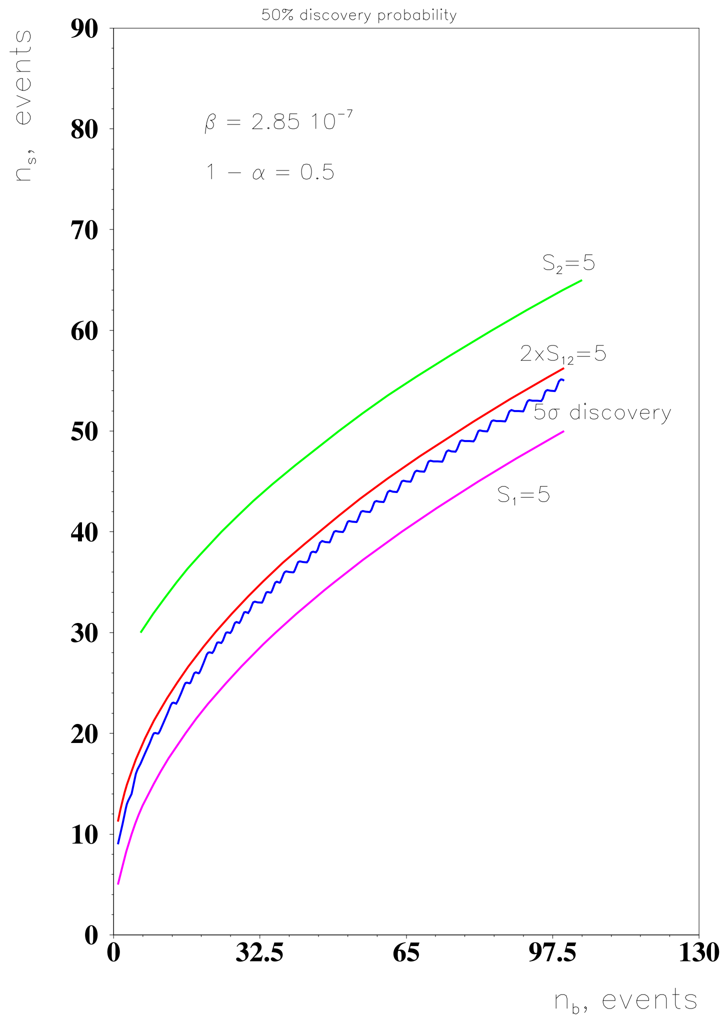

Figure 5: The 5σ discovery curve and dependences n_s versus n_b for $S_1 = 5$, $S_2 = 5$, $2 \cdot S_{12} = 5$. Here $1 - \alpha = 0.5$ and $\beta = 2.85 \cdot 10^{-7}$.

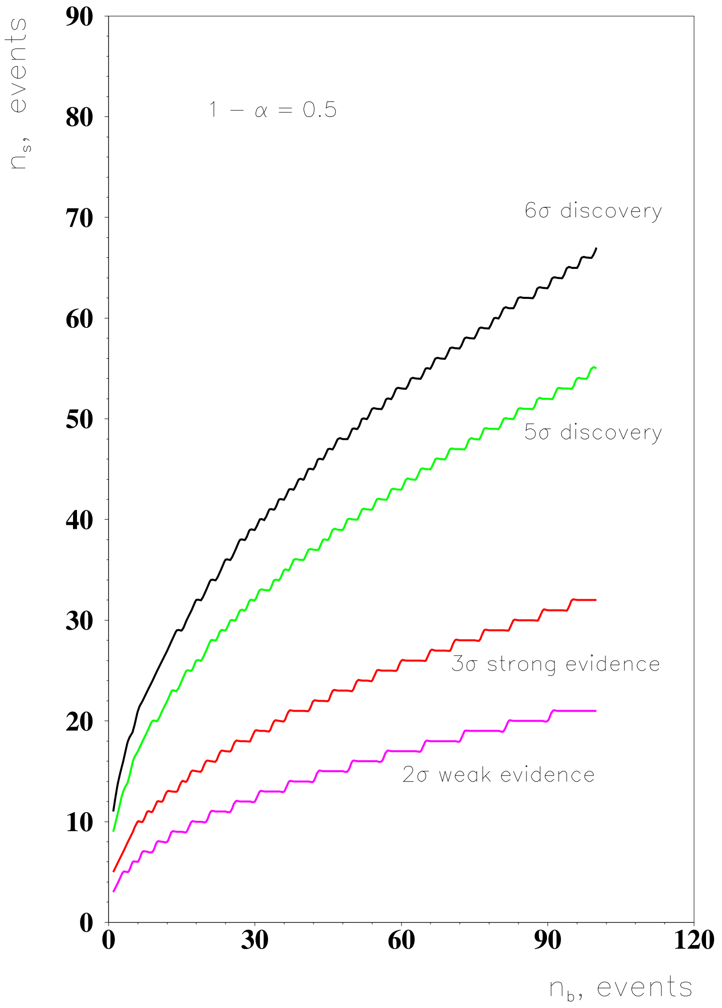

Figure 6: Dependences n_s versus n_b for $1 - \alpha = 0.5$ and for different values of β.

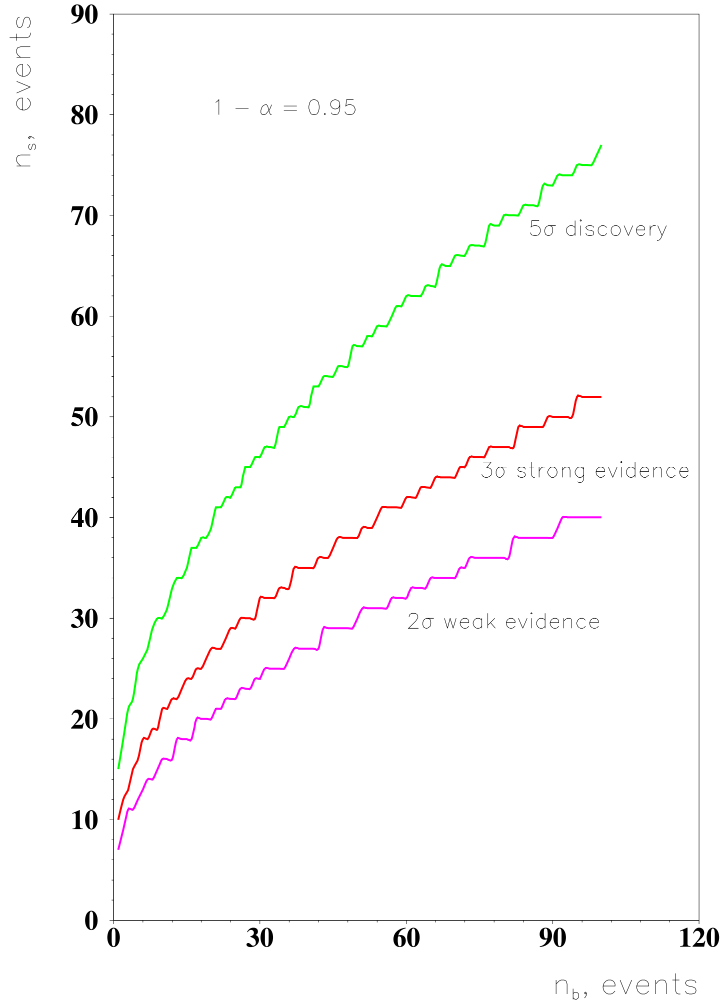

Figure 7: Dependences n_s versus n_b for $1 - \alpha = 0.95$ and for different values of β.

Exclusion limits.

Application of the test of equal-probability allows to estimate the minimal value of relative number of wrong decisions $\tilde{\kappa} = \dfrac{\alpha+\beta}{2-(\alpha+\beta)}$ of the future hypotheses testing for Poisson distributions. It is the uncertainty of observability of the phenomenon.
As a result the requirement $1-\tilde{\kappa} = 0.9$ allows to estimate an 90% exclusion limit.
In fact $1-\tilde{\kappa}$ there is a distance between two distributions (measure of distinguishability).

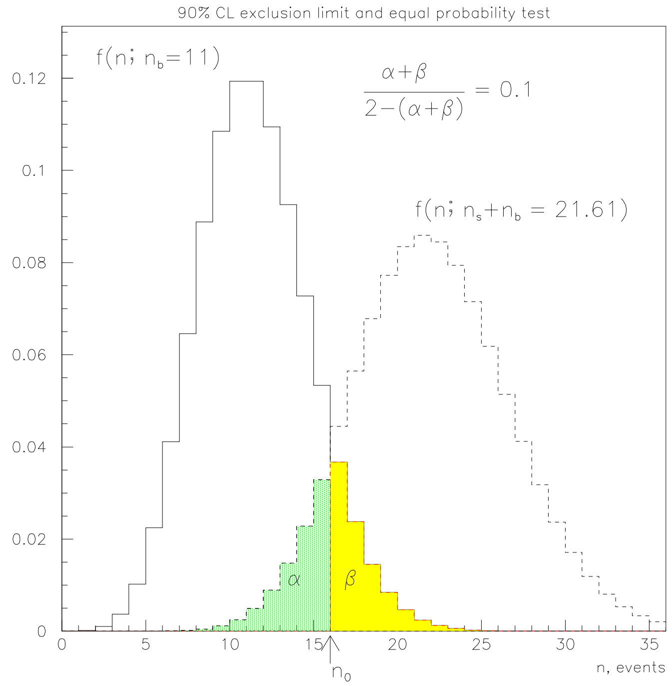

Figure 8: Equal-probability test for the case $n_b = 11$ and $n_s = 10.61$ gives the critical value $n_0 = 16$ and, correspondingly, the relative uncertainty $\tilde{\kappa} = 0.1$. It means if we observed $n_{obs} \le n_0$ events in planned experiment we exclude the signal with average $n_s = 10.61$ and above at 90%CL.

In [Hernandez, NIM A378 (1996) 301] the criterion $\beta < \Delta$ and $\dfrac{\alpha}{1-\beta} < \epsilon$ for determination of the exclusion limits have been applied. It means 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 this case two parameters $(\epsilon,\ \Delta)$ are required to construct exclusion limits.

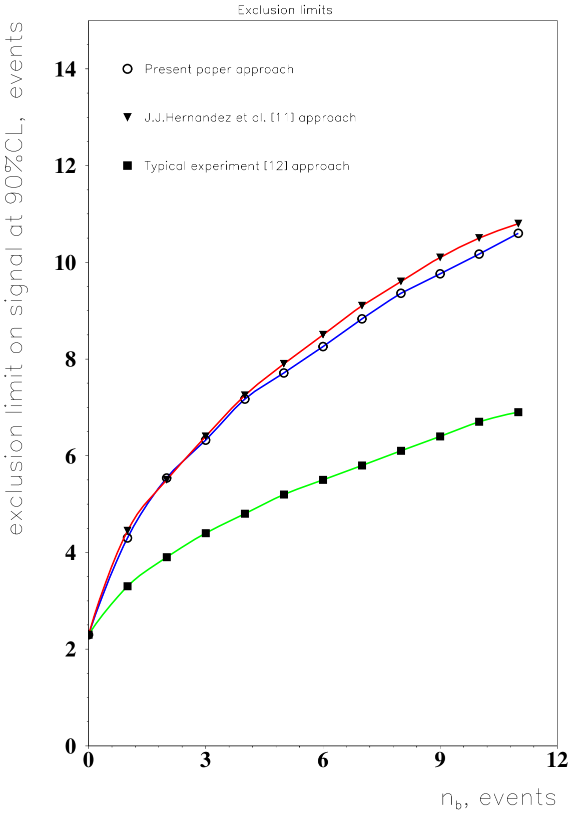

Figure 9: Estimations of the 90% CL exclusion limit on the signal in a future experiment as a function of the expected background.

An account of systematic uncertainties related to nonexact knowledge of background and signal cross sections.

There are considered two types of uncertainty in planned experiments: the systematic uncertainties due to imperfect knowledge of the background and signal cross sections and the statistical uncertainties due to statistical errors in determination of values n_b and n_s. Let us to consider forthcoming experiments to search for new physics. In this case the systematic uncertainties has theoretical origin without any statistical properties which can be measured.

Suppose uncertainty in the calculation of exact background cross section is determined by parameter δ, i.e. the exact cross section lies in the interval $(\sigma_b, \sigma_b(1+\delta))$ and the exact value of average number of background events lies in the interval $(n_b, n_b(1+\delta))$. Let us suppose $n_b \gg n_s$. In this instance the discovery potential most sensitive to the systematic uncertainties. Because we know nothing about possible values of average number of background events, we consider the worst case.

$$\beta(\mathbf{\Delta}) = \sum_{\mathbf{n}=\mathbf{n_0}(\mathbf{\Delta})+1}^{\infty} \mathbf{f}(\mathbf{n}; \mathbf{n_b}(\mathbf{1}+\delta)) \leq \mathbf{\Delta} \qquad (9)$$

$$\mathbf{1} - \alpha(\mathbf{\Delta}) = \sum_{\mathbf{n}=\mathbf{n_0}(\mathbf{\Delta})+1}^{\infty} \mathbf{f}(\mathbf{n}; \mathbf{n_s} + \mathbf{n_b}), \qquad (10)$$

These equations allow to estimate the scale of influence of background uncertainty to observability of signal.

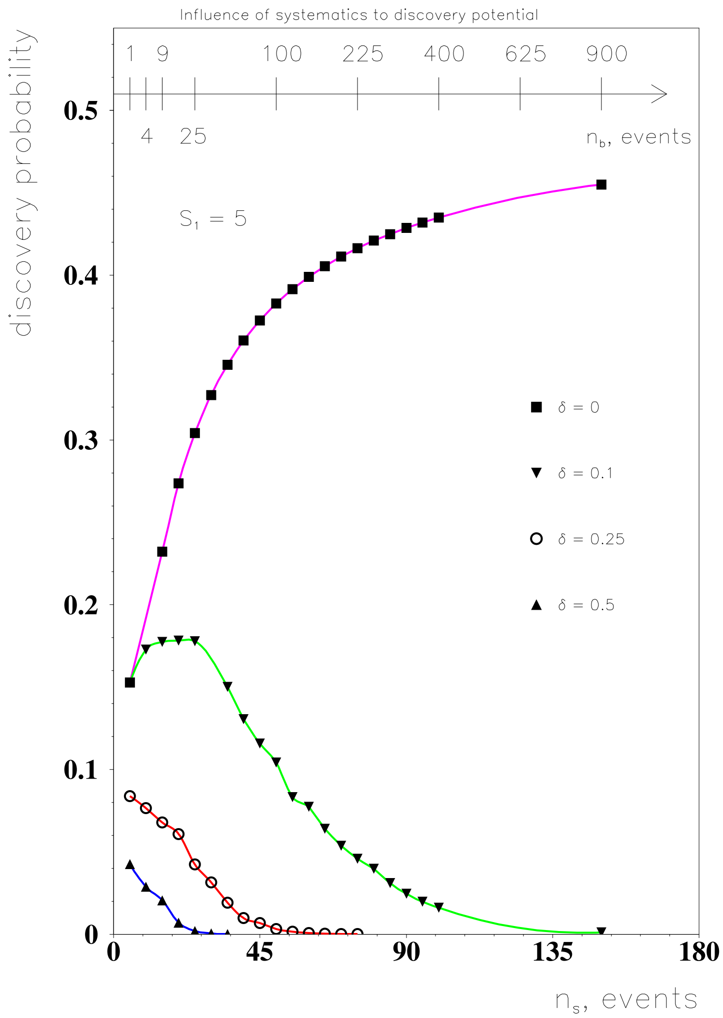

Figure 10: Discovery probability versus n_s for different values of systematic uncertainty δ for the case $S_1 = 5$.

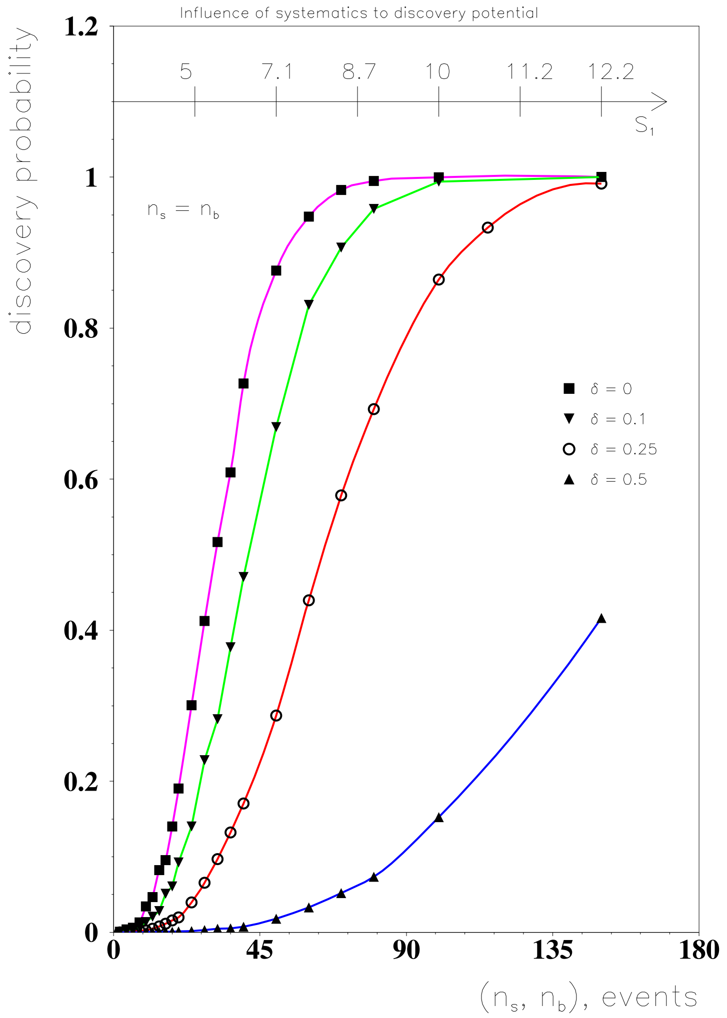

Figure 11: Discovery probability versus n_s for different values of systematic uncertainty δ for the case $n_s = n_b$.

The account for uncertainty in determination of n_b and n_s

Let us write down the density of Gamma distribution $\Gamma_{a,n+1}$ as

$$g_n(a,\lambda) = \frac{a^{n+1}}{\Gamma(n+1)} e^{-a\lambda} \lambda^n, \qquad (11)$$

where a is a scale parameter, $n+1 > 0$ is a shape parameter, $\lambda > 0$ is a random variable, and $\Gamma(n+1)$ is a Gamma function. Since the n is integer, then $n! = \Gamma(n+1)$.

Let us set $a = 1$, then for each n a continuous function

$$g_n(\lambda) = \frac{\lambda^n}{n!} e^{-\lambda}, \;\; \lambda > 0, \;\; n > -1 \qquad (12)$$

is the density of Gamma distribution $\Gamma_{1,n+1}$ with the scale parameter $a = 1$ (see Figure). The mean, mode, and variance of this distribution are given by $n+1$, n, and $n+1$, respectively.

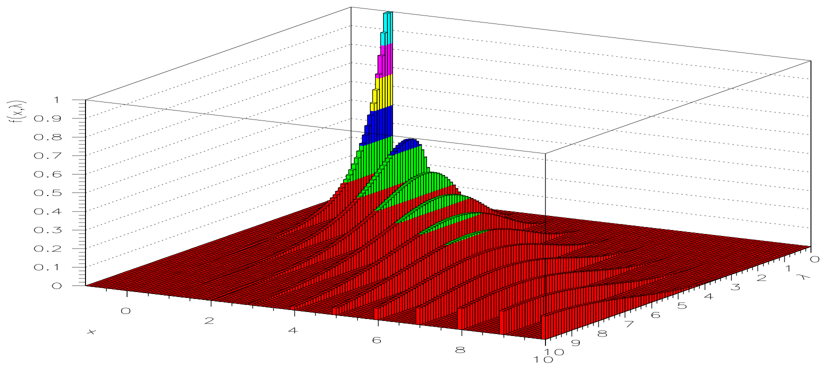

Figure 12: The behaviour of the probability density of the true value of parameter λ for the Poisson distribution in case of n observed events versus λ and n. Here $f(n, \lambda) = \dfrac{\lambda^n}{n!} e^{-\lambda}$ is both the Poisson distribution with the parameter λ along the axis n and the Gamma distribution with a shape parameter $n + 1$ and a scale parameter 1 along the axis λ.

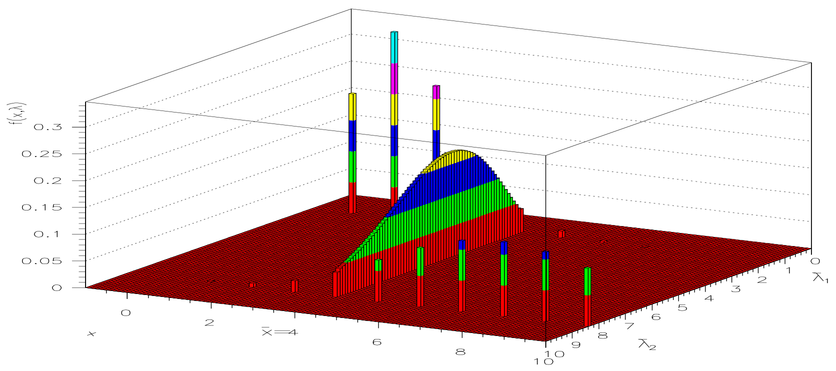

Figure 13: The Poisson distributions $f(n, \lambda)$ for λ's determined by the confidence limits $\hat{\lambda}_1 = 1.51$ and $\hat{\lambda}_2 = 8.36$ in case of the observed number of events $\hat{n} = 4$ are shown. The probability density of Gamma distribution with a scale parameter $a = 1$ and a shape parameter $n + 1 = \hat{n} + 1 = 5$ is shown within this confidence interval.

As it follows from the article [Cousins, Am.J.Phys 63 (1995) 398] and is clearly seen from the identity

$$\sum_{n=\hat{n}+1}^{\infty} f(n;\lambda_1) + \int_{\lambda_1}^{\lambda_2} g_{\hat{n}}(\lambda)d\lambda + \sum_{n=0}^{\hat{n}} f(n;\lambda_2) = 1 \ , \quad i.e. \tag{13}$$

$$\sum_{n=\hat{n}+1}^{\infty} \frac{\lambda_1^n e^{-\lambda_1}}{n!} + \int_{\lambda_1}^{\lambda_2} \frac{\lambda^{\hat{n}} e^{-\lambda}}{\hat{n}!} d\lambda + \sum_{n=0}^{\hat{n}} \frac{\lambda_2^n e^{-\lambda_2}}{n!} = 1$$

for any $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ the probability of true value of parameter of Poisson distribution to be equal to the value of λ in the case of one measurement $\hat{n}$ has probability density of Gamma distribution $\Gamma_{1,1+\hat{n}}$.

This allows to transform functions $f_0(n)$ and $f_1(n)$ accordingly to calculate the probability of discovery

$$1-\alpha = 1 - \int_0^{\infty} g_{n_s+n_b}(\lambda) \sum_{n=0}^{n_0} f(n)d\lambda = 1 - \sum_{n=0}^{n_0} \frac{C^n_{n_s+n_b+n}}{2^{n_s+n_b+n+1}}, \tag{14}$$

where the critical value n_0 under the future hypotheses testing about the observability is chosen so that the Type II error

$$\beta = \int_0^{\infty} g_{n_b}(\lambda) \sum_{n=n_0+1}^{\infty} f(n)d\lambda = \sum_{n=n_0+1}^{\infty} \frac{C^n_{n_b+n}}{2^{n_b+n+1}} \tag{15}$$

could be less or equal to $2.85 \cdot 10^{-7}$. Here C^k_m is $\dfrac{m!}{k!(m-k)!}$.

The Poisson distributed random values have a property: if $\xi \sim Pois(\lambda_1)$ and $\eta \sim Pois(\lambda_2)$ then $\xi + \eta \sim Pois(\lambda_1 + \lambda_2)$. It means that if we have two measurements $\hat{n}_1$ and $\hat{n}_2$ of the same random value $\xi \sim Pois(\lambda)$, we can consider these measurements as one measurement $\hat{n}_1 + \hat{n}_2$ of the random value $2 \cdot \xi \sim Pois(2 \cdot \lambda)$.

Correspondingly, in the case of m measurements $\hat{n}_1, \hat{n}_2, \ldots, \hat{n}_m$ of the random values $\xi_1, \xi_2, ..., \xi_m$, where $\xi_i \sim Pois(\lambda)$ for $i = 1, 2, \ldots, m$, the probability of true value of parameter of Poisson distribution to be concrete λ, has probability density of Gamma distribution

$$\Gamma_{m, 1+\sum_{i=1}^{m} \hat{n}_i}. \tag{16}$$

Conclusion

- **The method to estimate the discovery potential on new physics in planned experiments is discussed.**
- **The results of direct calculations and calculations by approximate formulae of dependences n_s versus n_b for different discovery probabilities and significances are presented.**
- **The method for determination of exclusion limits is described.**
- **The two approaches for taking into account the uncertainty in estimations of signal and background are proposed.**

Acknowledgments

We are grateful to Fred James, V.A. Matveev and V.F. Obraztsov for the interest to this work and for valuable comments. S.B. thanks Bob Cousins, George Kahrimanis, James Linnemann, Louis Lyons, Tony Vaiciulis, Alex Read, Byron Roy and Pekka Sinervo for very useful discussions. S.B. would like to thank James Stirling, Mike Whalley and Linda Wilkinson for having organized this interesting Conference which is the wonderful opportunity for exchange of ideas.