Uncertainties in estimation of quality of planned experiments

S.I. Bityukov[1,2], N.V. Krasnikov

Institute for nuclear research RAS,

Prospect 60-letiya Octyabrya, 117312 Moscow, Russia

[1] *Institute for high energy physics, 142281 Protvino, Russia*

[2] *http://home.cern.ch/bityukov*

October 18, 2004

Abstract

In the paper the approach to estimation of quality of planned experiments is proposed. This approach is based on the analysis of uncertainty, which will take place under the future hypotheses testing about the existence of a new phenomenon in Nature (H0: new physics is present in Nature versus H1: new physics is absent). The signal significance is considered in terms of Type I and Type II errors. Approximate formulae for calculation of signal significance in the language of standard deviations are obtained. The probability of making a correct decision in hypotheses testing is proposed as one of the estimators of the quality of planned experiment. Also the possible measure of the distinguishability of hypotheses is discussed. Incorporation of uncertainties to signal significance and the probability of making a correct decision in planned experiment is an topical problem. The proposed estimators allow to take into account for systematics and statistical uncertainties in determination of signal and background rates.

1 Introduction

One of the common goals in the forthcoming experiments is the search for new phenomena. In the forthcoming high energy physics experiments (LHC, TEV22, NLC, ...) the main goal is the search for physics beyond the Standard Model (supersymmetry, Z'-, W'-bosons, ...) and the Higgs boson discovery as a final confirmation of the Standard Model. In estimation of the discovery potential of the planned experiments (to be specific in this paper we shall use as an example CMS experiment at LHC [1]) the background cross section (the Standard Model cross section) is calculated and for the given integrated luminosity $\mathcal{L}$ the average number of background events is $n_b = \dot{\sigma}_b \cdot \mathcal{L}$. Suppose the existence of a new physics leads to additional nonzero signal cross section $\dot{\sigma}_s$ with the same signature as for the background cross section that results in the prediction of the additional average number of signal events [1] $n_s = \dot{\sigma}_s \cdot \mathcal{L}$ for the integrated luminosity $\mathcal{L}$.

The total average number of the events is $< n >= n_s + n_b = (\dot{\sigma}_s + \dot{\sigma}_b) \cdot \mathcal{L}$. So, as a result of new physics existence, we expect an excess of the average number of events. In real experiments the probability of the realization of n events is described by Poisson distribution

$$f(n;\mu) = \frac{\mu^n}{n!} e^{-\mu}. \tag{1}$$

Here $\mu =< n >$ is the average number of events. Remember that the Poisson distribution $f(n;\mu)$ gives [2] the probability of finding exactly n events in the given interval of (e.g. space and time) when the events occur independently of one another at an average rate of μ per the given interval. For the Poisson distribution the variance σ^2 equals to μ. So, to estimate the probability of the new physics discovery we have to compare the Poisson statistics with $\mu = n_b$ and $\mu = n_b + n_s$. Usually, high energy physicists use the following "significances" for testing the possibility to discover new physics in an experiment:

(a) "significance" $S_1 = \dfrac{n_s}{\sqrt{n_b}}$ [1, 3],

[1]It should be noted that the existence of new physics can also lead to the decrease of the cross section due to destructive interference or some nonlocal formfactors. In this paper we consider the case when the new physics existence leads to additional positive contribution to the background cross section. The consideration of the opposite case is straightforward.

(b) "significance" $S_2 = \dfrac{n_s}{\sqrt{n_s + n_b}}$ [4],

(c) "significance" $S_{c12} = 2(\sqrt{n_s + n_b} - \sqrt{n_b})$ [5] (see, also, [6]).

A conventional claim is that for S_1 (S_2) ≥ 5 we shall discover new physics (here, of course, the systematic uncertainties are ignored). For $n_b \gg n_s$ the significances S_1 and S_2 coincide (the search for Higgs boson through the $h \to \gamma\gamma$ signature). For the case when $n_s \sim n_b$, S_1 and S_2 differ. Therefore, a natural question arises: what is the correct definition for the significance S_1, S_2 or anything else ?

It should be noted that there is a crucial difference between the planned experiment and the real experiment. In the real experiment the total number of events n_{obs} is a given number (already has been measured) and we compare it with n_b when we test the validity of the standard physics. So, the number of possible signal events is determined as $n_s = n_{obs} - n_b$ and it is compared with the average number of background events n_b. The fluctuation of the background is $\sigma_{fb} = \sqrt{n_b}$, therefore, we come to the S_1 significance as the measure of the distinction from the standard physics. In the conditions of the planned experiment when we want to search for new physics, we know only the average number of the background events and the average number of the signal events, so we have to compare the Poisson distributions $f(n; n_b)$ and $f(n; n_s + n_b)$ to determine the probability to find new physics in planned experiment.

In this paper we describe a method for estimation of the discovery potential on new physics in planned experiments. We also estimate the influence of systematic uncertainties related to nonexact knowledge of signal and background cross sections on the probability to discover new physics in planned experiments. An account of such systematics is very essential in the search for supersymmetry at LHC. We find that the more proper definition of the significance of 50% discovery probability in planned experiments is $S_{c12} = 2(\sqrt{n_s + n_b} - \sqrt{n_b})$ in comparison with often used significances $S_1 = \dfrac{n_s}{\sqrt{n_b}}$ and $S_2 = \dfrac{n_s}{\sqrt{n_s + n_b}}$, where n_s and n_b are the average numbers of signal and background events. For $1 - \alpha > 0.5$, i.e. for discovery probability more than 50%, there is additional additive contribution $-k(\alpha)$ in formula for the significance. Here α is a Type I error in hypotheses testing about observability of new physics. We propose a method for taking into account statistical uncertainties caused by inexact determination the number of signal and background events.

2 Estimators of quality of planned experiments

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)

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

2.1 Signal significance

As it has been mentioned in the introduction the crucial difference between planned experiment and real experiment is that in real experiment we know the number of observed events, therefore we can compare the Standard Model with experimental data directly, whereas in the case of planned experiment we know only the average number of background events n_b and the average number of signal events (for the case when we have new physics in addition to the Standard Model). Therefore in the case of planned experiment an additional "input" parameter is the probability of the discovery. Suppose we test two models: the Standard Model with the average number of events $\mu = n_b$ and the model with new physics and the average number of events $\mu = n_s + n_b$.

To discover new physics we have to require that the probability $\beta(\Delta)$ of the background fluctuations for $n > n_0(\Delta)$ is less than Δ, namely

$$\beta(\Delta) = \sum_{n=n_0(\Delta)+1}^{\infty} f(n; n_b) \le \Delta \qquad (2)$$

The probability $1-\alpha(\Delta)$ that the number of events in a model with new physics will be bigger than $n_0(\Delta)$ is equal to

$$1-\alpha(\Delta) = \sum_{n=n_0(\Delta)+1}^{\infty} f(n; n_s + n_b) \tag{3}$$

It should be stressed that if Δ is a given number then $\alpha(\Delta)$ is a function of Δ or, vice versa, we can fix the value of α in Eq.3 then Δ is a function of α. The meaning of *the probability of the discovery* $1-\alpha$ is the probability that in the case of new physics an experiment will measure the number of events bigger than n_0 such that the probability that the Standard Model can reproduce such number of events is rather small (β).

In other words we choose the critical value n_0 for hypotheses testing[2] about observability of new physics requiring that Type II error $\beta \leq \Delta$. Then we calculate the Type I error α and the probability of discovery (or *the probability of the evidence*) $1-\alpha$.

For fixed value of α and known values of n_s, n_b we can calculate β using Eqs.2,3. In our numerical calculations we take $\alpha = 0.5$ and 0.1. Consider now the limiting case $n_b \gg 1$ when Poisson distribution approaches Gaussian distribution. Eqs.2,3 take the form

$$\beta \approx \int_{n_0}^{\infty} P_G(x; n_b, n_b)dx \tag{4}$$

$$1-\alpha \approx \int_{n_0}^{\infty} P_G(x; n_s + n_b, n_s + n_b)dx \tag{5}$$

Consider at first the most simple case when $\alpha = 0.5$ (see Figs.1-2 for an illustration). For $\alpha = 0.5$ parameter n_0 in Eq.5 is equal to $n_0 = n_s + n_b$. Eq.4 takes the form

$$\beta \approx \int_{S_1}^{\infty} P_G(x; 0, 1)dx, \tag{6}$$

where

$$S_1 = \frac{n_s}{\sqrt{n_b}} \tag{7}$$

The significance S_1 is determined by Eq.7 and it is often used in experiment proposals [1, 3].

[2]A simple statistical hypothesis H_0 (new physics is present, i.e. $\mu = n_s + n_b$) against a simple alternative hypothesis H_1 (new physics is absent, i.e. $\mu = n_b$) [2].

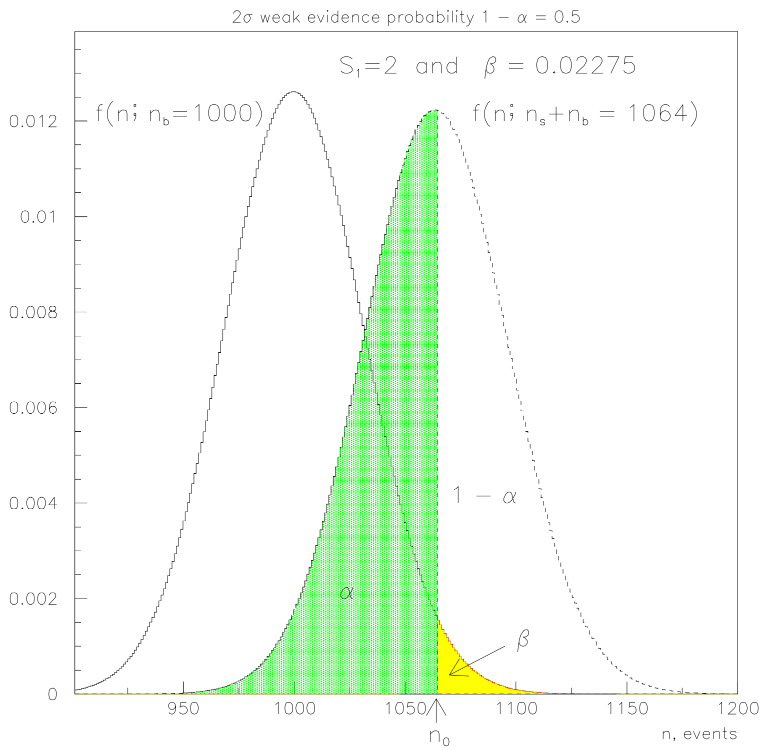

Figure 1: The case $n_b \gg 1$. Poisson distributions with parameters $\mu = 1000$ and $\mu = 1064$. Here $1-\alpha = 0.5$ and $\beta = 0.02275$ (i.e. $S_1 = 2$).

For $1-\alpha > 0.5$ (see Fig.3 for an illustration) the parameter n_0 in Eq.4 is equal to

$$n_0 = n_s + n_b - k(\alpha)\sqrt{n_s + n_b}, \tag{8}$$

where $k(\alpha)$: $k(0.5) = 0$; $k(0.25) = 0.66$; $k(0.1) = 1.28$; $k(0.05) = 1.64$ (as an example, Tab.31.1 [7]). We can define the *effective significance* $s(\alpha)$ (instead of S_1 in Eq.6), i.e. corrected significance S_1, corresponding the discovery probability $1-\alpha$, as

$$s(\alpha) = \frac{n_s}{\sqrt{n_b}} - k(\alpha)\sqrt{1 + \frac{n_s}{n_b}}. \tag{9}$$

So, we see that the asymptotic formula (Eq.7) for the significance $s(\alpha)$ is valid only for $1-\alpha = 0.5$.

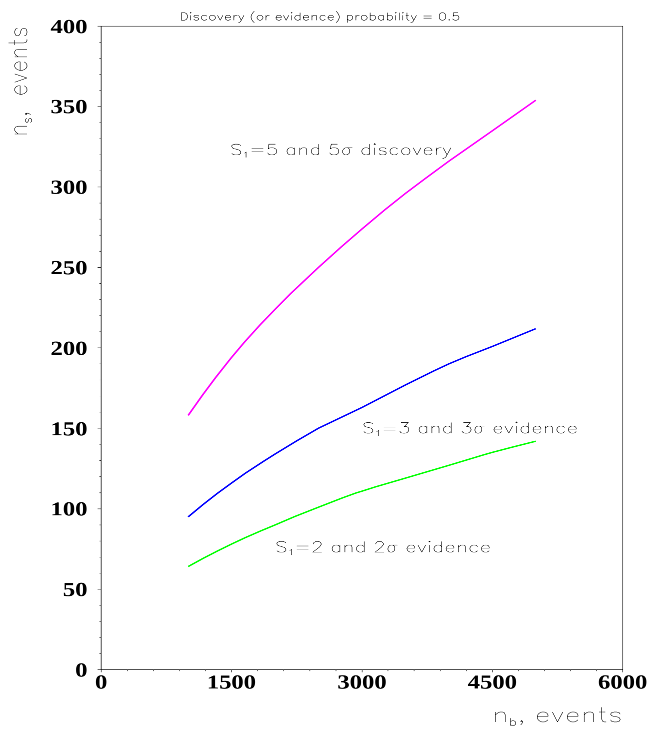

Figure 2: The case $n_b \gg 1$. Dependences n_s versus n_b for $S_1 = 5$, $S_1 = 3$ and $S_1 = 2$ coincide with 5σ discovery, 3σ strong evidence, and 2σ weak evidence curves, correspondingly. The probability of discovery $1 \ - \ \alpha = 0.5$.

As it has been shown in Ref. [8] the more proper of the significance in planned experiments is S_{c12}. The generalization of this significance to the case of $1 - \alpha > 0.5$ looks very attractive for approximate estimation of discovery potential

$$s(\alpha) = 2 \cdot (\sqrt{n_s + n_b} - \sqrt{n_b}) - k(\alpha). \tag{10}$$

The comparison of formulae (Eq.9,10) is shown in Fig.4.

It should be stressed that very often in the conditions of planned experiment the average numbers of background and real events are not very big and we have to solve Eqs.2,3 directly to construct 5σ *discovery*, 3σ *strong evidence* and 2σ *weak evidence* curves. Our numerical results are presented in Figs. (5 - 6).

As an example consider the search for standard Higgs boson with a mass

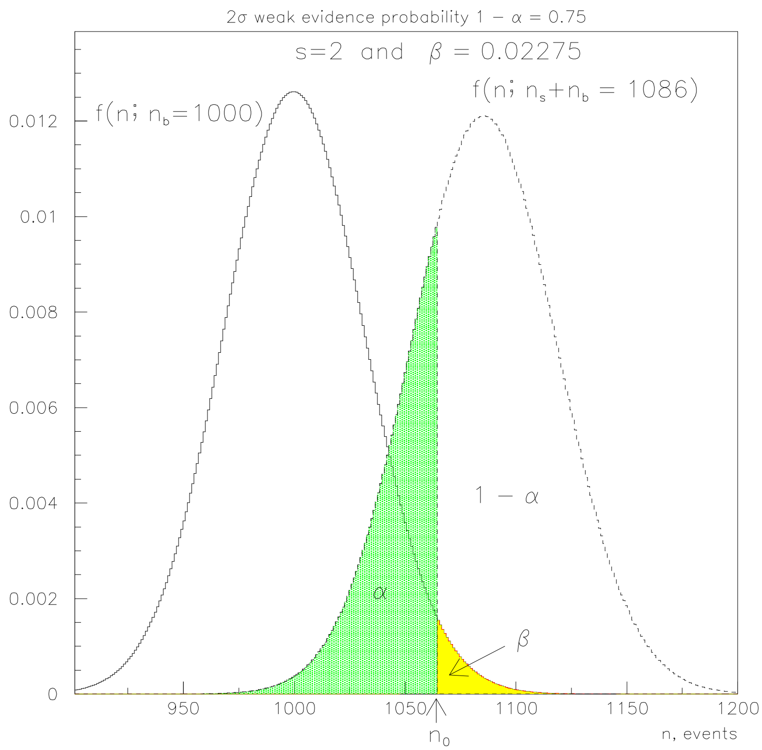

Figure 3: The case $n_b \gg 1$ and $S_1 = 2.72$. Poisson distributions with parameters $\mu = 1000$ and $\mu = 1086$. Here $1 - \alpha = 0.75$ and $\beta = 0.02275$ (i.e. effective $s = 2$).

$m_h = 110\ GeV$ using the $h \to \gamma\gamma$ decay mode at the CMS detector. For total luminosity $\mathcal{L} = 3 \cdot 10^4 pb^{-1}(2 \cdot 10^4 pb^{-1})$ one can find [1] that $n_b = 2893(1929), n_s = 357(238)$, $S_1 = \dfrac{n_s}{\sqrt{n_b}} = 6.6(5.4)$. Using Eq.9 and Table of the standard normal probability density function [2] we find that $1 - \alpha(\Delta_{dis}) = 0.93(0.60)$. It means that for total luminosity $\mathcal{L} = 3 \cdot 10^4 pb^{-1}(2 \cdot 10^4 pb^{-1})$ the CMS experiment will discover at $\geq 5\sigma$ level standard Higgs boson with a mass $m_h = 110\ GeV$ with a probability 93(60) percents [3].

For the case when we are interested in estimation of the lower bound on

[3] In other words let us suppose that we have constructed 100 identical CMS detectors. At $\geq 5\sigma$ level the Higgs boson will be discovered at 93(60) CMS detectors

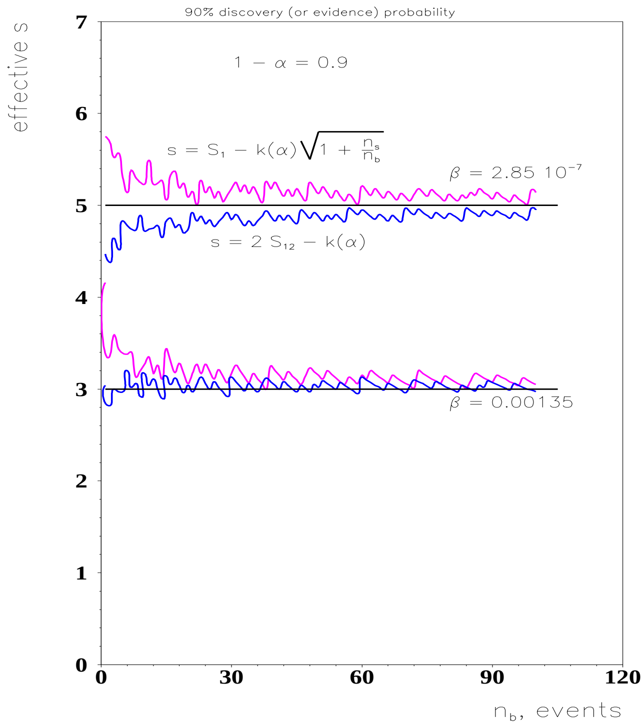

Figure 4: The estimation of effective significance $s(\alpha)$ for given β and $1 - \alpha$.

number n_s of signal events (bound on new physics) we can use the equations

$$1 - \alpha(\Delta) = \sum_{n=n_0(\Delta)+1}^{\infty} f(n; n_b + n_s) \tag{11}$$

$$\beta(\Delta) = \sum_{n=n_0(\Delta)+1}^{\infty} f(n; n_b) \geq \Delta \tag{12}$$

2.2 The probability of making a correct decision

Let us introduce the estimator

$$\hat{\kappa} = \frac{\hat{\alpha} + \hat{\beta}}{2} \tag{13}$$

of the uncertainty [5] $\kappa = \alpha + \beta$, when testing H_0 versus H_1 with an equal-tailed test. It is the probability of making an incorrect choice in favour 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.

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* [9] 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 favour of one of the hypotheses is independent of whether H_0 or H_1 is true. We consider two conditional distributions of probabilities

$$\begin{cases} f_0(n) = f(n;\mu_s + \mu_b), \\ f_1(n) = f(n;\mu_b) \end{cases} \tag{14}$$

We suppose that any prior suppositions about H_0 and H_1 can be included in $f_0(n)$ and $f_1(n)$.

2.2.1 Equal-tailed test

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 [10] ($\hat{\alpha} = \hat{\beta}$), their combination Eq.13 is the probability of making incorrect choice in favour of one of the hypotheses [9].

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 realized 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.13 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 $\quad(\hat{\kappa}-\hat{\alpha}=\hat{\kappa}-\hat{\beta}=0)$. As a result the estimator Eq.13 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.

The advantages of this estimator are:

- in case of continuous distributions this probability is independent of which hypothesis is chosen as H_0, which is H_1, and which is true,
- in case of discrete distributions the error $\hat{e}$ can be taken into account,
- this estimator allows the comparison of planned experiments.

2.2.2 Equal probability test

The equal probability test [8] gives results close to the equal-tailed test in the case of Poisson distributions, i.e. we can consider $\hat{\kappa}$ under equal probability test as an approximation of the probability of making incorrect decision.

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

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

$$\begin{cases} \hat{\alpha}=\sum\limits_{i=0}^{n_c} f(i;\mu_s+\mu_b)=\sum\limits_{i=0}^{n_c} f_0(i), \\ \hat{\beta}=1-\sum\limits_{i=0}^{n_c} f(i;\mu_b)=1-\sum\limits_{i=0}^{n_c} f_1(i), \end{cases} \tag{15}$$

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 $f_0(i) \le f_1(i)$). This follows directly from

$$\hat{\kappa}=\frac{\hat{\alpha}+\hat{\beta}}{2}=\frac{1}{2}(1-\sum_{i=0}^{n_c}(f_1(i)-f_0(i))). \tag{16}$$

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, and

$$n_c = [\frac{\mu_s}{ln(\mu_s + \mu_b) - ln(\mu_b)}], \quad (17)$$

where square brackets mean the integer part of a number.

2.2.3 On the signal significance

$\hat{\kappa}$ plays the role of Δ in the definition of the confidence level and, correspondingly, of the significance S of an excess of signal events above background [11] in planned experiments. In the case of Poisson distributions the definition of significance as

$$\hat{\kappa} = \frac{1}{\sqrt{2\pi}} \int_{S_{12}}^{\infty} e^{-\frac{x^2}{2}} dx. \quad (18)$$

leads to the formula [5, 8]

$$S_{c12} = 2 \cdot S_{12} = 2 \cdot (\sqrt{\mu_s + \mu_b} - \sqrt{\mu_b}). \quad (19)$$

A factor two is needed to correspond with common practice. As shown in [12] this approximation has good statistical properties as the significance for Poisson distributions.

2.3 Distinguishability of hypotheses

The probability of making a correct decision in hypotheses testing has disadvantages to be *a measure of the distinguishability of hypotheses*, namely,

- the equal-tailed test gives non-minimal magnitudes of sum of Type I and Type II errors [9, 13],
- the $\hat{\kappa}$ mainly assumes values in $[0, \frac{1}{2}]$ (the desirable region for probabilistic measure is $[0, 1]$ [14]),
- the equal-tailed test does not always give a single-valued critical region for complicated distributions.

2.3.1 The possible measure of the distinguishability of hypotheses

The value [4]

$$1 - \tilde{\kappa} = 1 - \frac{\hat{\alpha} + \hat{\beta}}{2 - (\hat{\alpha} + \hat{\beta})}, \tag{21}$$

is devoid of these disadvantages if we use the equal probability test [15]
There are 3 possibilities.

a) Distributions $f_0(n)$ and $f_1(n)$ have no overlapping, hence, the distributions are completely distinguishable and any result of the experiment will give the correct choice between hypotheses, i.e. $1 - \tilde{\kappa} = 1$.

b) Distributions $f_0(n)$ and $f_1(n)$ coincide completely. It means, that it is impossible to get a correct answer, i.e. $f_0(n)$ and $f_1(n)$ are not distinguishable, i.e. $1 - \tilde{\kappa} = 0$.

c) Distributions $f_0(n)$ and $f_1(n)$ do not coincide, but they have an overlapping, i.e. $\tilde{\kappa}$ is the fraction of incorrect decisions under the equal probability test. If in this case we hold the designation $\hat{\kappa} = \dfrac{\hat{\alpha} + \hat{\beta}}{2}$ then

$$\tilde{\kappa} = \frac{\hat{\kappa}}{1 - \hat{\kappa}}. \tag{22}$$

2.3.2 The case of Poisson distributions

Let the probability of observing n events in an experiment be described by a Poisson distribution with parameter μ. Then the Type I and II errors and, correspondingly, the measure of the distinguishability of hypotheses can be written as

---

[4]If we will use the geometric approach (let us the A is a set of possible realizations of the result of the planned experiment if the hypothesis H_0 takes place in Nature and the B is a set of possible realizations of the result of the planned experiment if the hypothesis H_1 takes place) then we have the total number of the possibilities for decision equals to $A \bigcup B$ and the fraction of incorrect decisions will be

$$\tilde{\kappa} = \frac{A \bigcap B}{A \bigcup B} = \frac{\hat{\alpha} + \hat{\beta}}{2 - (\hat{\alpha} + \hat{\beta})}. \tag{20}$$

$$\begin{cases} \hat{\alpha} = \sum_{i=0}^{n_c} f(i; \mu_s + \mu_b) = \sum_{i=0}^{n_c} f_0(i), \\ \hat{\beta} = 1 - \sum_{i=0}^{n_c} f(i; \mu_b) = 1 - \sum_{i=0}^{n_c} f_1(i), \\ 1 - \tilde{\kappa} = 1 - \dfrac{\hat{\alpha} + \hat{\beta}}{2 - (\hat{\alpha} + \hat{\beta})}, \end{cases} \tag{23}$$

with critical value n_c determined by Eq.17 (see, Fig.7).

Notice, that this critical value in case of the process of Poisson with parameter $\mu \cdot t$ practically conserves the linearity with respect to time

$$n_c \cdot t = [\frac{\mu_s \cdot t}{ln(\mu_s + \mu_b) - ln(\mu_b)}]. \tag{24}$$

3 Incorporating systematic uncertainties

There is considered the systematics and statistical uncertainties in planned experiments due to imperfect knowledge of the background and signal cross sections.

In Ref. [16] (R.D. Cousins and V.L. Highland) the systematic uncertainty is the uncertainty in the sensitivity factor (inefficiencies in the registration and/or the reconstruction of events and so on). This uncertainty has statistical properties which can be measured or estimated in real or Monte Carlo experiment. The systematic effects in Ref. [17] (G. D'Agostini and M. Raso) as supposed has stochastic behavior too and these effects are taken into account in frame of Bayesian approach. In review [19] P. Sinervo has motivated three classes of systematic uncertainties in measurements. Class 1 systematics are uncertainties that can be constrained by ancillary measurements and can therefore be treated as statistical uncertainties. Class 2 systematics arised from model assumption in the measurement of from poorly understood features of the data or analysis technique that introduce a potential bias in the experimental outcome. Class 3 systematics arise from uncertainties in the underlying theoretical paradigm used to make inferences using the data. This classification can be applied in planning of experiments too. Below we consider the way to take into account for statistical errors in determination of values n_b and n_s [18] (class 1 systematics) and the systematic uncertainty

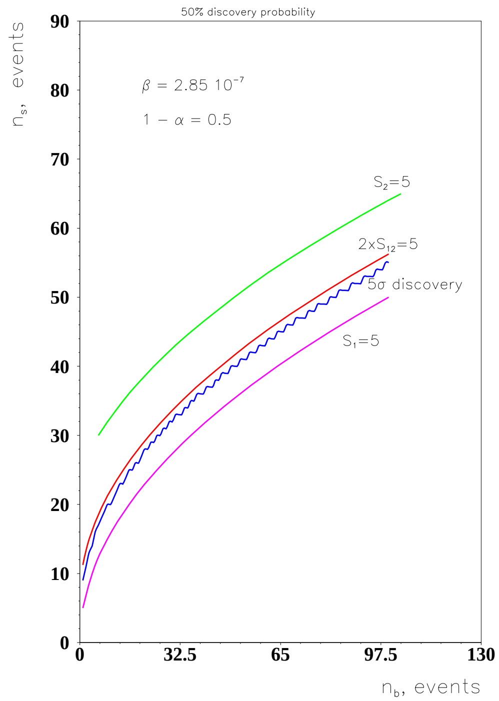

Figure 5: The 5σ discovery curve and dependences n_s versus n_b for $S_1 = 5$, $S_2 = 5$, $2 \cdot S_{12} = 5$. Here $1 - \alpha = 0.5$ and $\beta = 2.85 \cdot 10^{-7}$.

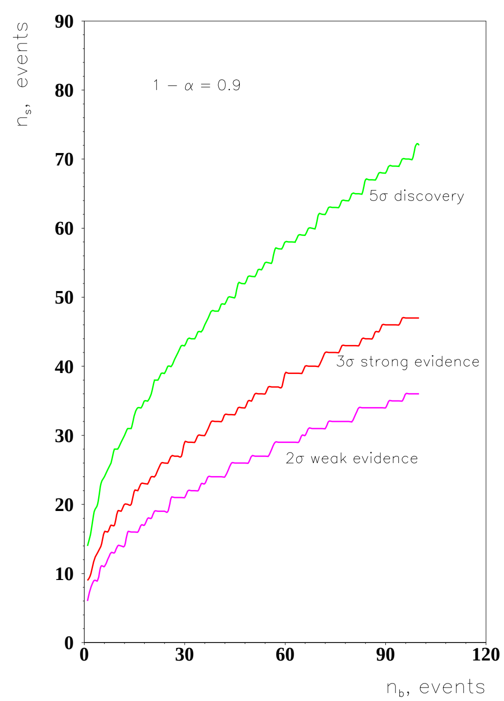

Figure 6: Dependences n_s versus n_b for $1 - \alpha = 0.9$ and for different values of β.

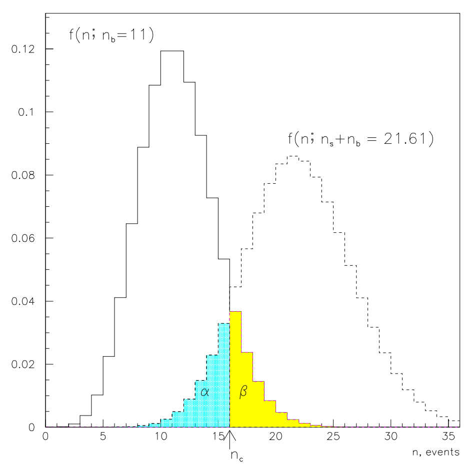

Figure 7: Equal probability test for the case $n_b = 11$ and $n_s = 10.61$ gives the critical value $n_c = 16$ and, correspondingly, the probability of incorrect decision $\hat{\kappa} = 0.09$ and the measure of distinguishability of hypotheses $\tilde{\kappa} = 0.1$.

which has theoretical origin without any statistical properties [5, 18] (class 3 systematics) in frame of frequentist approach.

Let the values $n_s = \hat{n}_s$ and $n_b = \hat{n}_b$ be known from Monte Carlo calculations. In this case they are random variables. These values can be considered as estimators of unknown parameters. Consequently, the values n_c, α and β are also random variables. It means that $1 - \tilde{\kappa}$ is the estimator of the distinguishability of hypotheses. Let us consider how the uncertainties in the knowledge of n_s and n_b influence on the measure of distinguishability of hypotheses $1 - \tilde{\kappa}$. The cases of the signal significance and the probability of making a correct decision are considered in Refs. [18] and [9], correspondingly.

Suppose, as before, that the streams of signal and background events are Poisson's.

3.1 The uncertainty in determination of the signal and background rates

As shown in ref. [20] the Gamma-distribution $\Gamma_{1,n+1}$ (with probability density

$$g_n(\mu) = \frac{\mu^n}{n!} e^{-\mu}, \;\; \mu > 0, \;\; n > -1) \tag{25}$$

and the Poisson distribution with parameter μ are statistically dual distributions. The identity [21] (see, also, [22, 11, 23])

$$\sum_{k=n+1}^{\infty} f(k;\mu_1) + \int_{\mu_1}^{\mu_2} g_n(\mu) d\mu + \sum_{k=0}^{n} f(k;\mu_2) = 1 \tag{26}$$

for any $\mu_1 \geq 0$ and $\mu_2 \geq 0$ is true for these distributions. It allows estimate the parameter μ of the Poisson distribution by the measurement of the random variable n, because the parameter value of this distribution (in case of single observation of the number of events) is described by Gamma-distribution $\Gamma_{1,1+n}$ with mean, mode, and variance $n + 1$, n, and $n + 1$, respectively [21, 18]. This statement was checked by Monte Carlo experiment [24].

The identity Eq.26 shows that conditional distribution of the probability of true value of parameter of Poisson distribution is a Gamma-distribution $\Gamma_{1,1+n}$ on condition that the measured value of the number of events is equal to n. As a result we can mix Bayesian and frequentist probabilities in frame of frequentist approach.

It allows to transform the probability distributions $f(i; n_s + n_b)$ and $f(i; n_b)$ accordingly to calculate the measure of distinguishability of hypotheses

$$\begin{cases} \hat{\alpha} = \int_0^{\infty} g_{n_s+n_b}(\mu) \sum_{i=0}^{n_c} f(i;\mu) d\mu = \sum_{i=0}^{n_c} \frac{C^i_{n_s+n_b+i}}{2^{n_s+n_b+i+1}}, \\ \hat{\beta} = 1 - \int_0^{\infty} g_{n_b}(\mu) \sum_{i=0}^{n_c} f(i;\mu) d\mu = 1 - \sum_{i=0}^{n_c} \frac{C^i_{n_b+i}}{2^{n_b+i+1}}, \\ 1 - \tilde{\kappa} = 1 - \frac{\hat{\alpha} + \hat{\beta}}{2 - (\hat{\alpha} + \hat{\beta})}. \end{cases} \tag{27}$$

Here the critical value n_c under the future hypotheses testing about the observability is chosen in accordance with test of equal probability and C^i_N is $\dfrac{N!}{i!(N-i)!}$. Also we suppose that the Monte Carlo luminosity is exactly the same as the data luminosity later in the experiment.

The Poisson distributed random values have a property: if $\xi_i \sim Pois(\mu_i)$, $i = 1, 2, \ldots, m$ then $\sum\limits_{i=1}^{m} \xi_i \sim Pois(\sum\limits_{i=1}^{m} \mu_i)$. It means that if we have m observations $\hat{n}_1$, $\hat{n}_2$, ..., $\hat{n}_m$ of the same random value $\xi \sim Pois(\mu)$, we can consider these observations as one observation $\sum\limits_{i=1}^{m} \hat{n}_i$ of the Poisson distributed random value with parameter $m \cdot \mu$. According to Eq.26 the probability of true value of parameter of this Poisson distribution has probability density of Gamma distribution $\Gamma_{1,1+\sum_{i=1}^{m} \hat{n}_i}$. Using the scale parameter m one can show that the probability of true value of parameter of Poisson distribution in the case of m observations of the random value $\xi \sim Pois(\mu)$ has probability density of Gamma distribution $\Gamma_{m,1+\sum_{i=1}^{m} \hat{n}_i}$ [5], i.e.

$$\begin{aligned} G(\sum \hat{n}_i, m, \mu) &= g_{(\sum_{i=1}^{m} \hat{n}_i)}(m, \mu) \\ &= \frac{m^{(1+\sum_{i=1}^{m} \hat{n}_i)}}{(\sum_{i=1}^{m} \hat{n}_i)!} e^{-m\mu} \mu^{(\sum_{i=1}^{m} \hat{n}_i)}. \end{aligned} \tag{28}$$

Let us assume that the integrated luminosity of planned experiment is

[5] The probability density of the $\Gamma_{a,n+1}$-distribution (in our notation) is $g_n(a,\mu) = \dfrac{a^{n+1}}{\Gamma(n+1)} e^{-a\mu} \mu^n$, where a is a scale parameter and $n+1$ is a shape parameter.

$\mathcal{L}$ and the integrated luminosity of Monte Carlo data is $m \cdot \mathcal{L}$. For instance, we can divide the Monte Carlo data into m parts with luminosity corresponding to the planned experiment. The result of Monte Carlo experiment in this case looks as set of m pairs of numbers ($(n_b)_i, \ (n_b)_i + (n_s)_i$), where $(n_b)_i$ and $(n_s)_i$ are the numbers of background and signal events observed in each part of Monte Carlo data. Let us denote $N_b = \sum\limits_{i=1}^{m} (n_b)_i$ and $N_{s+b} = \sum\limits_{i=1}^{m} ((n_s)_i + (n_b)_i)$. Correspondingly (see page 98, [11]),

$$\begin{cases} \hat{\alpha} = \displaystyle\int_0^{\infty} G(N_{b+s}, m, \mu) \sum_{i=0}^{n_c} f(i;\mu) d\mu = \sum_{i=0}^{n_c} C^i_{N_{s+b}+i} \frac{m^{1+N_{s+b}}}{(m+1)^{1+N_{s+b}+i}}, \\ \hat{\beta} = 1 - \displaystyle\int_0^{\infty} G(N_b, m, \mu) \sum_{i=0}^{n_c} f(i;\mu) d\mu = 1 - \sum_{i=0}^{n_c} C^i_{N_b+i} \frac{m^{1+N_b}}{(m+1)^{1+N_b+i}}. \end{cases} \tag{29}$$

As a result, we have a generalized system of equations for the case of different luminosity in planned data and Monte Carlo data to calculate the measure of distinguishability of hypotheses $1 - \tilde{\kappa} = 1 - \dfrac{\hat{\alpha} + \hat{\beta}}{2 - (\hat{\alpha} + \hat{\beta})}$. The set of values $C^i_{N+i} \dfrac{m^{1+N}}{(m+1)^{N+i+1}}, \ i = 0, 1, \ldots$ is a negative binomial (Pascal) distribution with real parameters $N + 1$ and $\dfrac{1}{m+1}$, mean value $\dfrac{1+N}{m}$ and variance $\dfrac{(1+m)(1+N)}{m^2}$.

3.2 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 any 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 point the admissible level of uncertainty in theoretical calculations for given experiment proposal. In principle, it is “reproducible inaccuracy introduced by faulty technique” [25] and according to [26] it contains the sense of “incompetence”.

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 $(\dot{\sigma}_b, \dot{\sigma}_b(1+\delta))$ 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 is the most sensitive to the systematic uncertainties. As we know nothing about possible values of average number of background events, we consider the worst case [5]. Taking into account Eqs.23 we have the formulae [6]

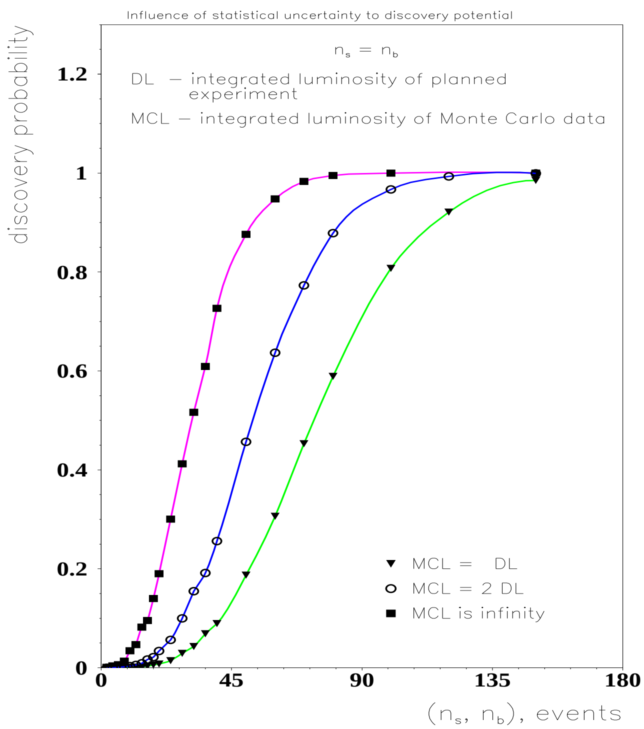

Figure 8: Discovery probability versus n_s with and without account for statistical uncertainty in determination of n_s and n_b. The case $n_s = n_b$. The curves are constructed under condition $\beta = 2.85 \cdot 10^{-7}$.

[6]Eqs.30 realize the worst case when the background cross section $\dot{\sigma}_b(1+\delta)$ is the maximal one, but we think that both the signal and the background cross sections are minimal. Also, we suppose that $n_b(1+\delta) < n_s + n_b$.

$$\begin{cases} \hat{\alpha} = \sum_{i=0}^{n_c} f(i; n_b + n_s) \\ \hat{\beta} = 1 - \sum_{i=0}^{n_c} f(i; n_b(1+\delta)) \\ 1 - \tilde{\kappa} = 1 - \dfrac{\hat{\alpha} + \hat{\beta}}{2 - (\hat{\alpha} + \hat{\beta})}, \end{cases} \tag{30}$$

where n_c is

$$n_c = [\frac{n_s - n_b \cdot \delta}{ln(n_s + n_b) - ln(n_b \cdot (1+\delta))}]. \tag{31}$$

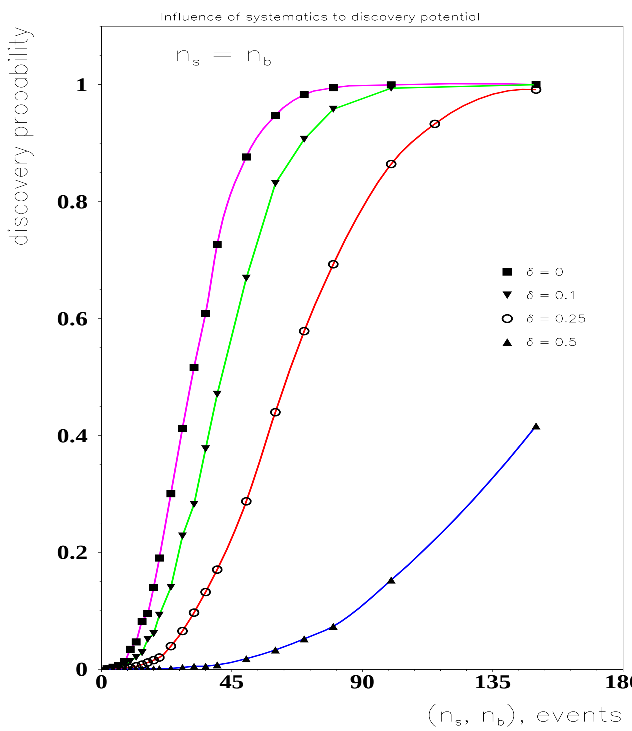

Figure 9: Discovery probability versus n_s for different values of systematic uncertainty δ for the case $n_s = n_b$. The curves are constructed under condition $\beta = 2.85 \cdot 10^{-7}$.

4 Conclusions

In this paper we have considered several estimators of the quality of planned experiments. These estimators allow to compare the discovery potential of different experiment proposals. We estimate the influence of statistical uncertainty in determination of mean numbers of signal and background events and propose a possible way to take into account effects of theoretical origin systematics. The scale of this influence to the discovery probability is shown in Fig.11 for statistical uncertainty and in Fig.12 for systematics.

The authors would like to thank the organizers of the Quarks2004 Conference for an interesting and enjoyable meeting. The author is grateful to Louis Lyons, V.A. Kachanov, V.A. Matveev and
V.F. Obraztsov for the interest and useful comments, S.S. Bityukov, R.D. Cousins, Yu.P. Gouz, G. Kahrimanis, A. Nikitenko, V.V. Smirnova, V.A. Taperechkina for fruitful discussions. This work has been supported by grants RFBR 03-02-16933 and RFBR 04-02-16020.

References

[1] *The Compact Muon Solenoid.* Technical Proposal, CERN/LHCC 94 -38, 1994.

[2] See as an example, W.T. Eadie, D. Drijard, F.E. James, M. Roos, and B. Sadoulet, *Statistical Methods in Experimental Physics,* North Holland, Amsterdam, 1971.

[3] ATLAS Detector and Physics Performance, TDR, CERN/LHCC/99-15, p.467, CERN, May, 1999.

[4] See as an example:

D. Denegri, L. Rurua and N. Stepanov, CMS Note CMS TN/96-059, October 1996.

S. Abdullin, *Search for SUSY at LHC: Discovery and Inclusive Studies,* Presented at International Europhysics Conference on High Energy Physics, Jerusalem, Israel, August 19-26, 1997, CMS Conference Report 97/019, November 1997.

[5] S.I. Bityukov and N.V. Krasnikov, Modern Physics Letters **A13** (1998) 3235;

[6] N. Brown, Z.Phys., **C49**, 1991, p.657.

H. Baer, M. Bisset, C. Kao and X. Tata, Phys.Rev., **D46**, 1992, p.1067.

[7] K. Hagiwara et al., Phys.Rev. D**66** (2002) 010001-1.

[8] S.I. Bityukov and N.V. Krasnikov, Nucl.Instr.&Meth. **A452** (2000) 518;

[9] S.I. Bityukov and N.V. Krasnikov, *The probability of making a correct decision in hypotheses testing as estimator of quality of planned experiments*, in G. Erickson and Y. Zhai (eds.), "Bayesian Inference and Maximum Entropy Methods in Science and Engineering", 23-th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Jackson Hole, Wyoming, 3-8 August 2003, AIP Conference Proceedings, vol.**707**:455-464, Melville, NY, 2004; also e-Print: physics/0309031.

[10] J.O. Berger, B. Boukai, Y. Wang, Stat.Sci.12 (1997) 133.

[11] A.G. Frodesen, O. Skjeggestad, H. Tøft, *Probability and Statistics in Particle Physics,* UNIVERSITETSFORLAGET, Bergen-Oslo-Tromsø, 1979.

[12] V.Bartsch and G.Quast, CMS Internal Note 2003/039. August, 2003.

[13] S.I. Bityukov and N.V. Krasnikov, *Incorporating systematics and statistical uncertainties into exclusion limits*, in Proceedings of the Conference Statistical Problems in Particle Physics, Astrophysics, and Cosmology (PhyStat'2003), Menlo Park, California, 8-11 Sep 2003, eds L. Lyons, R. Mount, R. Reitmeyer, p.318; eConf C030908:THNT002,2003.

[14] A.L. Gibbs, F.E. Su, e-Print: math.PR/0209021.

[15] S.I. Bityukov and N.V. Krasnikov, *Distinguishability of Hypotheses*, presented in the IX International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT'03), KEK, Tsukuba, Japan, December 1-5, 2003; NIM A (2004), in press.

[16] R.D. Cousins and V.L. Highland, Nucl.Instr.&Meth. **A320** (1992) 331-335.

[17] G. D'Agostini and M. Raso, CERN-EP/2000-026, 1 February, 2000; also, hep-ex/0002056. 22 Feb. 2000.

[18] S.I. Bityukov and N.V. Krasnikov, Nucl.Instr.&Meth. **A502** (2003) 795; S.I. Bityukov, JHEP 09 (2002) 060; e-Print: hep-ph/0207130.

[19] P. Sinervo *Definition and Treatment of Systematic Uncertainties in High Energy Physics and Astrophysics*, in Proceedings of the Conference Statistical Problems in Particle Physics, Astrophysics, and Cosmology (PhyStat'2003), Menlo Park, California, 8-11 Sep 2003, eds L. Lyons, R. Mount, R. Reitmeyer, p.318; eConf C030908:THNT002,2003.

[20] S.I. Bityukov, V.V. Smirnova, V.A. Taperechkina, "Statistically dual distributions and estimation of the parameters", submitted to Applied Mathematics Research eXpress.

[21] S.I. Bityukov, N.V. Krasnikov, V.A. Taperechkina, Preprint IFVE 2000-61, Protvino, 2000; e-Print: hep-ex/0108020, 2001.

[22] E.T. Jaynes: Papers on probability, statistics and statistical physics, Ed. by R.D. Rosenkrantz, D.Reidel Publishing Company, Dordrecht, Holland, 1983, p.165.

[23] R.D. Cousins, Am.J.Phys **63** (1995) 398-410.

[24] S.I. Bityukov, V.A. Medvedev, V.V. Smirnova, Yu.V. Zernii, *Experimental test of the probability density function of true value of Poisson distribution parameter by single observation of number of events*, presented in the IX International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT'03), KEK, Tsukuba, Japan, December 1-5, 2003; NIM A (2004), in press.

[25] R. Bevington, *Data reduction and Analysis for the Physical Sciences*, McGraw Hill 1969.

[26] R. Barlow, *Systematic errors: facts and fictions,* IPPP/02/39, DCPT/02/78, Proceedings of International Conference "Advanced Statistical Techniques in Particle Physics", March 18-22, 2002, Durham, UK, p.134. http://www.ippp.dur.ac.uk/statistics/