Advanced Computing and Analysis Techniques in Physics Research
December 1-5, 2003
KEK, Tsukuba, Ibaraki, Japan

Experimental test of the probability density function of true value of Poisson distribution parameter for observed number of events

S.I. Bityukov, V.A. Medvedev, V.V. Smirnova
Institute for High Energy Physics, Protvino, Russia
Yu.V. Zernii
MSA IECS, Moscow, Russia

The probability density function of true value of Poisson distribution parameter for observed number of events is constructed by computer experiment. The analysis of the PDF confirms that this distribution is gamma-distribution.

Scope

- Introduction
- The computer experiment
- The analysis of results
- Conclusion

Introduction

Let us consider the Gamma-distribution with the density of probability

$$g_x(\beta, \alpha) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)}. \tag{1}$$

At change of standard designations of Gamma-distribution

$$\frac{1}{\beta}, \ \alpha \ \textbf{and} \ x \quad \textbf{on} \quad a, \ n+1 \ \textbf{and} \ \lambda$$

the following formula for density of probability of Gamma-distribution takes place

$$g_n(a, \lambda) = \frac{a^{n+1}}{\Gamma(n+1)} e^{-a\lambda} \lambda^n, \tag{2}$$

where a is a scale parameter and $n+1 > 0$ is a shape parameter.

Suppose $a = 1$, then the density of probability of Gamma-distribution $\Gamma_{1,n+1}$ looks like Poisson distribution of probabilities:

$$g_n(\lambda) = \frac{\lambda^n}{n!} e^{-\lambda}, \ \lambda > 0, \ n > -1. \tag{3}$$

As it follows from the article [1] (see also [2]) and is clearly seen from the identity [3]

$$\sum_{k=n+1}^{\infty} f(k;\lambda_1) + \int_{\lambda_1}^{\lambda_2} g_n(\lambda)d\lambda + \sum_{k=0}^{n} f(k;\lambda_2) = 1 \ , \quad i.e. \qquad (4)$$

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

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 observation n has probability density of Gamma distribution $\Gamma_{1,1+n}$.

The mean, mode, and variance of this distribution are given by $n+1$, n, and $n+1$, respectively.

It means, the observed value n is associated with the most probable value of parameter of Poisson distribution, and the mean value of the parameter of Poisson distribution should correspond to $n+1$, i.e. the estimation of parameter of Poisson distribution by one observation is displaced on 1 from the measured value of number of events.

The equation (4) also shows that we can mix Bayesian and frequentist probabilities.

As a result, we can easily construct the confidence intervals, to take into account systematics and statistical uncertainties of measurements at statistical conclusions about the quality of planned experiments, to estimate the value of the parameter of Poisson distribution on several observations [3, 4].

Nevertheless there are works in which the approaches based on other assumptions of distribution of true value of parameter of Poisson distribution at presence of its estimation on one observation, for example [5], develop.

On the other side the works using methods Monte Carlo for construction of confidence intervals and for estimations of Type I and Type II errors in the testing of hypotheses recently have appeared [6, 7].

Therefore the experimental test with the purpose to confirm, that the true value of parameter of Poison distribution at one observation has density of probability of Gamma-distribution, and with the purpose to check up applicability of methods Monte Carlo to such tasks was carried out.

The computer experiment

From the eq.(4) follows, that any prior, except the uniform, on value of parameter of Poisson distribution in distribution of true value of this parameter at presence of the measured estimation n is excluded by existence of the boundary conditions determined by the sums of appropriate Poisson distributions.

Therefore we carried out the uniform scanning in parameter of Poisson distribution with step 0.1 from value $\lambda = 0.1$ up to value $\lambda = 20$, playing the Poisson distribution of 30000 trials for each value λ (Fig.1) with the using of function RNPSSN [8].

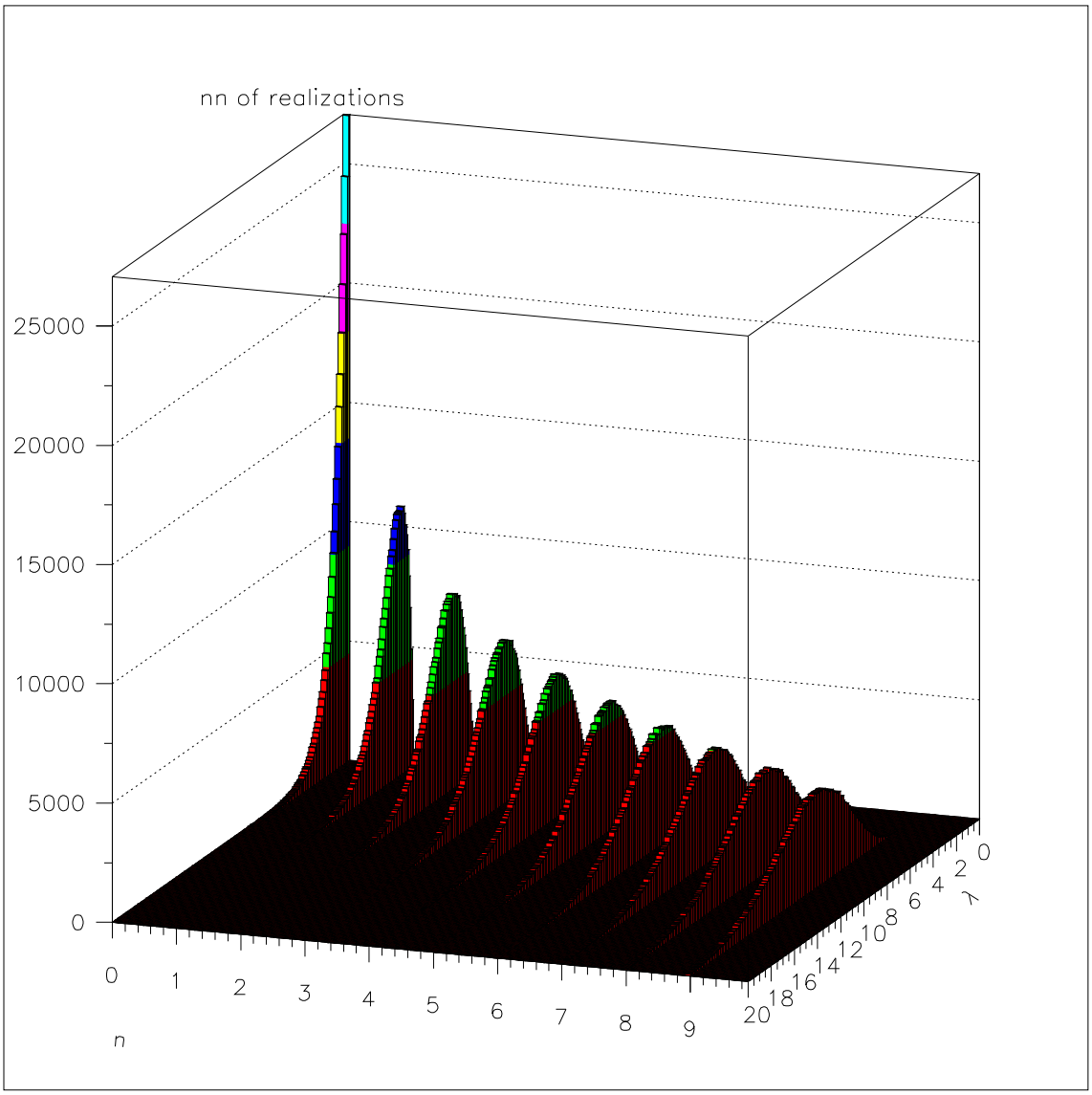

Figure 1: Amount of occurrences of n in the interval from 0 up to 9. Scannings in parameter of the Poisson distribution (30000 trials at each value of parameter λ) was carried out with step 0.1 in the interval of λ from 0.1 up to 20

After scanning for each value of number of the dropped out events n the empirical density of probability of true value of parameter of Poisson distribution be λ if the observation is equal n was obtained.

The analysis of results

In Fig.2 are shown the distribution (a), obtained at scanning in parameter λ with the selection of number of the dropped out events $n = 6$, and distribution (b) $\Gamma_{1,7}$, the appropriate area.

One can see that the average value of parameter $\lambda \approx 7$.

It means, the number of observed events is the estimation of the most probable value of the parameter of Poisson distribution (the mean value has bias).

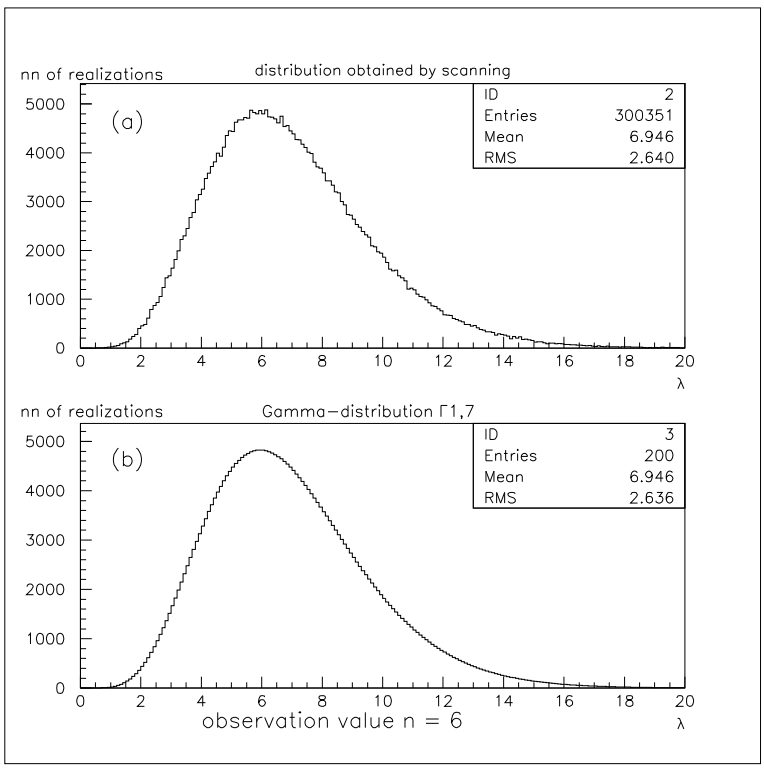

Figure 2: Distributions of occurrences of value $n = 6$ depending on value of parameter λ. The distribution (a) is obtained at Monte Carlo scanning in parameter λ. The distribution (b) is obtained by direct construction of Gamma-distribution $\Gamma_{1,7}$

The same distributions obtained by the selection of number of dropped out events $n = 0$ (Fig.3) and $n = 8$ (Fig.4) in logarithmic scale also are shown.

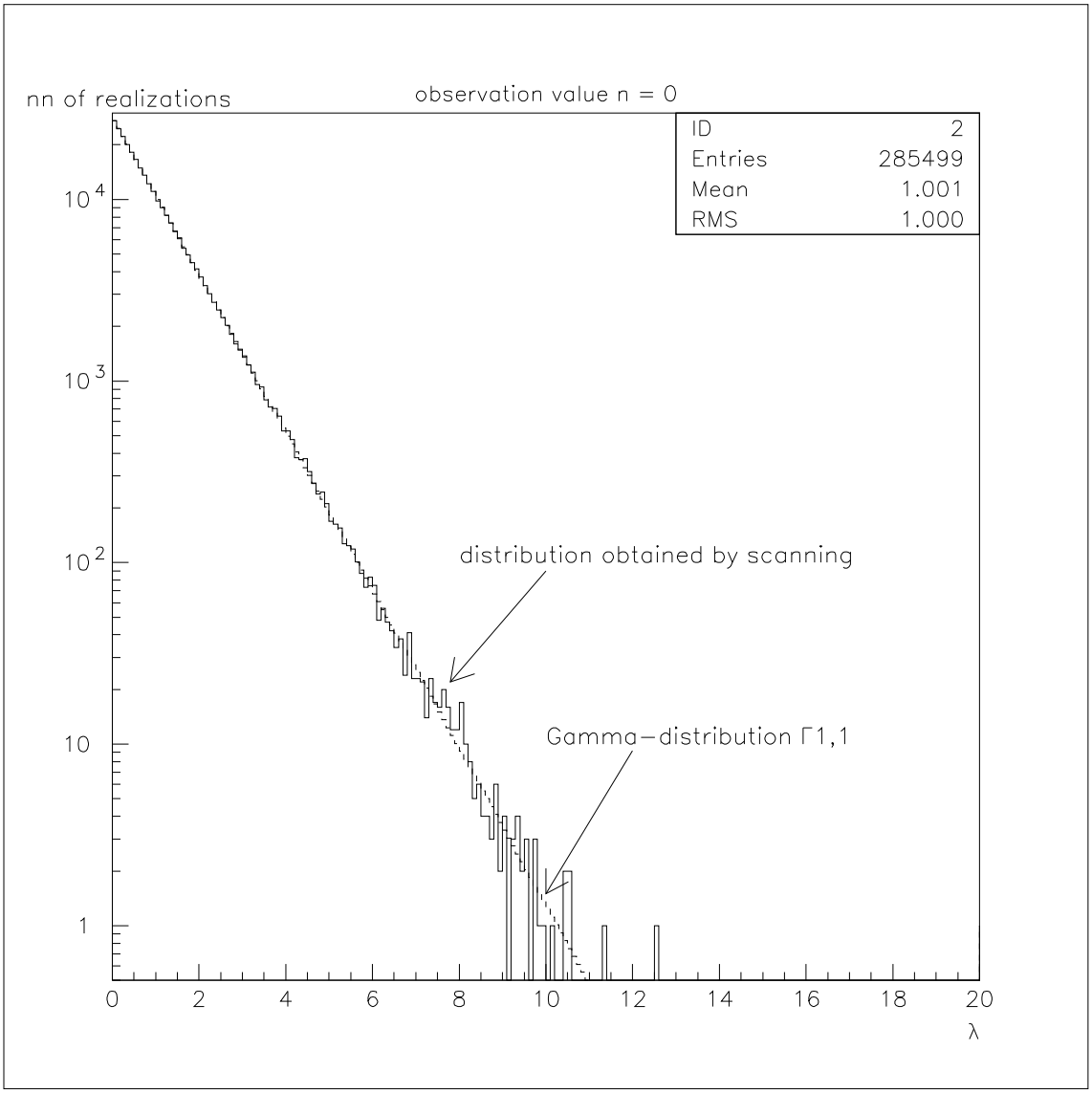

Figure 3: Distributions of occurrences of value $n = 0$ depending on value of parameter λ. The distribution (a) is obtained at Monte Carlo scanning in parameter λ. The distribution (b) is obtained by direct construction of Gamma-distribution $\Gamma_{1,1}$

Table 1: The probability of compartibility

n	probability
0	1.000000
1	0.999646
2	0.992521
3	0.999986
4	0.999969
5	0.999084
6	0.999986
7	0.999892
8	0.752075
9	0.974236

In Tab.1 are presented the values of probabilities of compatibility of the empirical distribution, obtained Monte Carlo by the scanning in parameter λ, and the appropriate Gamma-distribution for values n from 0 up to 9.

The calculations are based on the Kolmogorov Test (the function HDIFF of the package HBOOK [8]).

The authors of a package as criterion of coincidence of two distributions recommend to use the requirement of value of probability of compatibility more than 0.05.

In Fig.4 the least conterminous distributions are given.

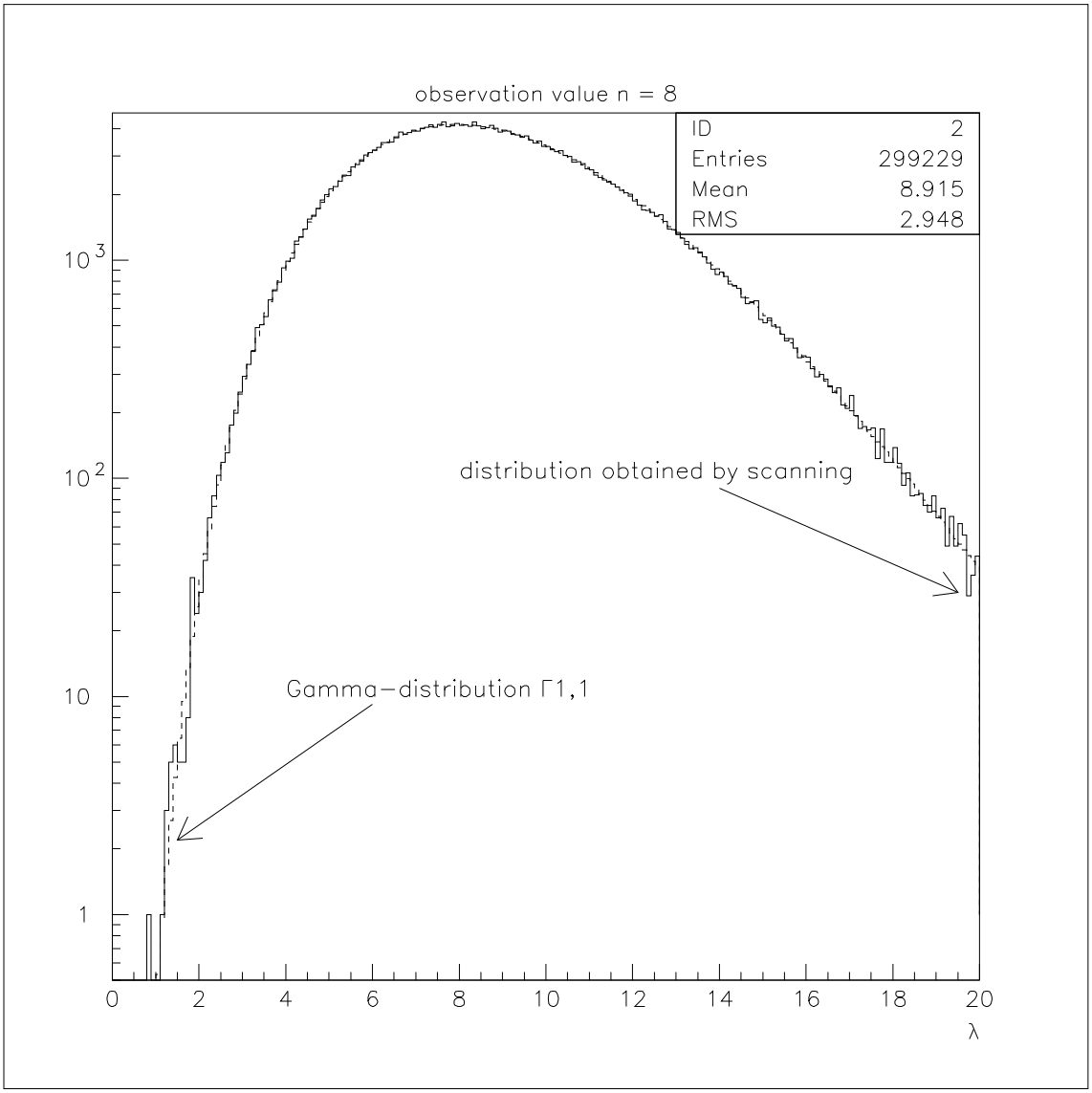

Figure 4: Distributions of occurrences of value $n = 8$ depending on value of parameter λ. The distribution (a) is obtained at Monte Carlo scanning in parameter λ. The distribution (b) is obtained by direct construction of Gamma-distribution $\Gamma_{1,9}$

Thus, the obtained results do not contradict the statement that conditional distribution of true value of parameter of Poisson distribution at single observation has a Gamma-distribution.

Conclusion

In the report the results of the computer experiment on the check of the statement, that true value of parameter of Poisson distribution at an estimation of this parameter on one observation n is the Gamma-distribution $\Gamma_{1.n+1}$, are presented.

The obtained results confirm the conclusions of the paper [3, 4] about a kind of conditional distribution of true value of parameter of Poisson distribution at single observation.

Note, that the given results also specify the applicability of method Monte Carlo for construction of conditional distribution of the true value of parameters of various distributions.

Acknowledgments

The authors thank N.V. Krasnikov, V.F. Obraztsov, V.A. Petukhov and M.N. Ukhanova for support of the given work. The authors also are grateful to S.S. Bityukov for fruitful discussions and constructive criticism. S.B. would like to thank Toshiaki Kaneko.

References

[1] **R.D. Cousins, Why isn't every physicist a Bayesian ?** Am.J.Phys **63 (1995) 398.**

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

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

[3] **S.I. Bityukov, N.V. Krasnikov, V.A. Taperechkina,** Confidence intervals for Poisson distribution parameter, **Preprint IFVE 2000-61, Protvino, 2000; also, e-Print: hep-ex/0108020, 2001.**

[4] **S.I.Bityukov,** On the Signal Significance in the Presence of Systematic and Statistical Uncertainties, **JHEP 09 (2002) 060, http://www.iop.org/EJ/abstract/1126-6708/2002/09/060; e-Print: hep-ph/0207130.**

S.I. Bityukov and N.V. Krasnikov, Signal significance in the presence of systematic and statistical uncertainties, **NIM A, 502 (2003) 795.**

[5] **G.J. Feldman and R.D. Cousins, Unified approach to the classical statistical analysis of small signal,** Phys.Rev. **D 57 (1998) 3873-3889**

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

[7] **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.**

[8] **CERNLIB, CERN PROGRAM LIBRARY, Short Writeups, Entry V136 and Y250, (CERN, Geneva, Switzerland, Edition - June 1996)**