STATISTICALLY DUAL DISTRIBUTIONS IN STATISTICAL INFERENCE

S. I. BITYUKOV and V. V. SMIRNOVA

Institute for high energy physics, 142281 Protvino, Russia
E-mail: Serguei.Bitioukov@cern.ch, vera@cub.ihep.su

N. V. KRASNIKOV

Institute for nuclear research RAS, 117312 Moscow, Russia
E-mail: krasniko@ms2.inr.ac.ru

V. A. TAPERECHKINA

Moscow State Academy of Instrument Engineering and Computer Science, Moscow, Russia

The estimation of a parameter of a model by the measurement of a random variable whose distribution depends on this parameter is one of the main tasks of statistics. In this paper the notion of the statistically dual distributions is introduced. An approach, based on the properties of the statistically dual distributions, to resolve the given task is proposed.

1. Introduction

As shown in refs. [1, 2], in the framework of frequentist approach we can construct the probability distribution of the possible magnitudes of the Poisson distribution parameter to give the observed number of events $\hat{n}$ in a Poisson stream of events. This distribution, which can be called a confidence density function of a parameter, is described by a Gamma-distribution with the probability density function which looks like a Poisson distribution of probabilities. This is the reason for naming this pair of distributions as statistically dual distributions. Also, the interrelation between the Poisson and Gamma distributions was used in these papers to reconstruct the confidence density of the Poisson distribution parameter by a unique way and, correspondingly, to construct any confidence interval for the parameter.

According to B. Efron [3] the confidence density is the fiducial [4] distribution of the parameter. This distribution is considered as a genuine *a posteriori* density for the parameter without prior assumptions.

The same relation [5–7], which allows one to reconstruct the confidence density of a parameter in a unique way, exists between several pairs of statistically self-dual distributions (normal and normal, Laplace and Laplace and, as is shown below, Cauchy and Cauchy).

Note that the posterior distribution of the parameter also is used for the definition of conjugate families in the Bayesian approach. The interrelation between the statistically dual distributions and conjugate families is discussed in ref. [7].

2. Statistically dual distributions

Let us define statistically dual distributions.

Definition 1: Let $\phi(x, \theta)$ be a function of two variables. If the same function can be considered both as a family of the probability density functions (pdf) $f(x|\theta)$ of the random variable x with parameter θ and as another family of pdf's $\tilde{f}(\theta|x)$ of the random variable θ with parameter x (i.e. $\phi(x, \theta) = f(x|\theta) = \tilde{f}(\theta|x)$), then this pair of families of distributions can be named as **statistically dual distributions**.

The statistical duality of Poisson and Gamma-distributions follows from simple discourse.

Let us consider the Gamma-distribution with probability density

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

Changing the standard designations of the Gamma-distribution $\dfrac{1}{\beta}$, α and x for a, $n+1$ and μ, we get the following formula for the probability density of Gamma-distribution

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

where a is a scale parameter and $n+1 > 0$ is a shape parameter. Suppose $a = 1$, then the formula of the probability density of Gamma-distribution $\Gamma_{1,n+1}$ is

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

It is a common supposition that the probability of observing n events in the experiment is described by a Poisson distribution with parameter μ, i.e.

$$f(n|\mu) = \frac{\mu^n}{n!}e^{-\mu}, \ \mu > 0, \ n \geq 0. \tag{4}$$

One can see that if the parameter and variable in Eq. (3) and Eq. (4) are exchanged, in other respects the formulae are identical. As a result these distributions (Gamma and Poisson) are **statistically dual distributions**. These distributions are connected by the identity [1] (see, also, this identity in another form in refs. [8], [9], [10])

$$\sum_{i=\hat{n}+1}^{\infty} f(i|\mu_1) + \int_{\mu_1}^{\mu_2} g_{\hat{n}}(\mu)d\mu + \sum_{i=0}^{\hat{n}} f(i|\mu_2) = 1, \tag{5}$$

i.e.

$$\sum_{i=\hat{n}+1}^{\infty} \frac{\mu_1^i e^{-\mu_1}}{i!} + \int_{\mu_1}^{\mu_2} \frac{\mu^{\hat{n}} e^{-\mu}}{\hat{n}!}d\mu + \sum_{i=0}^{\hat{n}} \frac{\mu_2^i e^{-\mu_2}}{i!} = 1$$

for any real $\mu_1 \geq 0$ and $\mu_2 \geq 0$ and non-negative integer $\hat{n}$.

The definition of the confidence interval (μ_1, μ_2) for the Poisson distribution parameter μ using [1, 5]

$$P(\mu_1 \leq \mu \leq \mu_2|\hat{n}) = P(i \leq \hat{n}|\mu_1) - P(i \leq \hat{n}|\mu_2), \tag{6}$$

where $P(i \leq \hat{n}|\mu) = \displaystyle\sum_{i=0}^{\hat{n}} \frac{\mu^i e^{-\mu}}{i!}$, allows one to show that a Gamma-distribution $\Gamma_{1,1+\hat{n}}$ is the probability distribution of different values of μ parameter of Poisson distribution on condition that the observed value of the number of events is equal to $\hat{n}$, i.e. $\Gamma_{1,1+\hat{n}}$ is the confidence density of the parameter μ. This definition is consistent with the identity Eq. (5). Note, if we suppose in Eq. (5) that $\mu_1 = \mu_2$ we have a conservation of probability. The right-hand side of Eq. (6) determines the frequentist sense of this definition.

Another example of statistically dual distribution is the Cauchy distribution with unknown parameter θ and known parameter b. Here we also can exchange the parameter θ and variable x while conserving the same formula of the probability density.

The probability density of the Cauchy distribution is

$$C(x|\theta) = \frac{b}{\pi(b^2 + (x-\theta)^2)}. \tag{7}$$

The probability density of its statistically dual distribution is also the Cauchy distribution:

$$\tilde{C}(\theta|x) = \frac{b}{\pi(b^2 + (x-\theta)^2)}. \tag{8}$$

In such a way the Cauchy distribution can be named as **statistically self-dual distribution**. An identity like Eq. (5) also holds,

$$\int_{\hat{x}}^{\infty} C(x|\theta_1)dx + \int_{\theta_1}^{\theta_2} \tilde{C}(\theta|\hat{x})d\theta + \int_{-\infty}^{\hat{x}} C(x|\theta_2)dx = 1, \tag{9}$$

where $\hat{x}$ is the observed value of random variable x and $\tilde{C}(\theta|\hat{x})$ is the confidence density.

3. Statistical duality and estimation of the parameter of a distribution

It is easy to show that the reconstruction of the confidence density is unique if Eqs. (5) or (9) holds [5–7].

As a result we have the **Transform** (both for Poisson-Gamma pair of families of distributions and for statistically self-dual distributions)

$$\int_{\hat{x}}^{\infty} f(x|\theta_1)dx + \int_{\theta_1}^{\theta_2} \tilde{f}(\theta|\hat{x})d\theta + \int_{-\infty}^{\hat{x}} f(x|\theta_2)dx = 1 \tag{10}$$

between the space of the realizations $\hat{x}$ of random variable x and the space of the possible values of the parameter θ, i.e.

$$\tilde{f}(\theta|\hat{x}) = T_{cd}\hat{x}, \tag{11}$$

where T_{cd} is the operator of the Transform. Here θ_1 and θ_2 are the bounds of the confidence interval for location parameter θ. As is shown above in the case of Gamma- and Poisson distributions, the two integrals are replaced by sums and $-\infty$ is replaced by 0.

The Transform Eq. (10) allows one to use statistical inferences about the random variable for estimation of an unknown parameter.

The simplest examples of this are given by several infinitely divisible distributions.

Definition 2: A distribution F is **infinitely divisible** if for each n there exist a distribution function F_n such that F is the n-fold convolution of F_n.

As known the Poisson, Gamma-, normal and Cauchy distributions are infinitely divisible distributions. The sum of independent and identically distributed random variables, which obey one of the above families of distributions, also obeys the distribution from the same family. Applying the Transform Eq. (10) to this sum allows one to reconstruct the confidence density of the parameter in the case of several observation of the same random variable. It means that we construct the relation

$$\tilde{f}(n\theta|\hat{x}_1+\hat{x}_2+\cdots+\hat{x}_n) = T_{cd}(\hat{x}_1+\hat{x}_2+\cdots+\hat{x}_n), \quad (12)$$

where T_{cd} is the operator of the Transform Eq. (10), the set $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$ are the observed values. Thereafter we reconstruct the confidence density of θ, i.e. $\tilde{f}(\theta|\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)$.

The method for construction of the confidence density of the mean value of several random variables which obey the Poisson distribution (sample with size > 1) and the way how to take into account the statistical uncertainty is shown in refs. [2].

The use of the confidence density also can be formulated in Bayesian framework. Let us consider, as an example, the Cauchy distribution. We suppose in our approach that the parameter θ is not a random value and before the measurement we do not prefer any values of this parameter, i.e. possible values of the parameter have equal probability and a prior distribution of θ is $\pi(\theta) = const$. Suppose we observe $\hat{x}_1$ and update our prior via the Transform Eq. (10) to obtain $\tilde{C}(\theta|\hat{x}_1)$, which is the pdf of the Cauchy distribution. This becomes our new prior before observing $\hat{x}_2$. It is easy to show that in the case of observing $\hat{x}_2$ the reconstructed confidence density $\tilde{C}(2\theta|\hat{x}_1 + \hat{x}_2)$ also is the pdf of the Cauchy distribution and, correspondingly, $\tilde{C}(\theta|\hat{x}_1, \hat{x}_2)$ is our next new prior. By induction this argument extends to sequences of any number of observations, i.e. we use the iterative procedure

$$\tilde{C}(\theta|\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{n-1}, \hat{x}_n) = T_{pd}(\tilde{C}(\theta|\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{n-1}), \hat{x}_n), \quad (13)$$

where T_{pd} is the operator of the Transform between a priori density and a posteriori density of the parameter.

4. Conclusions

We have formulated the notion of statistically dual distributions in the framework of probabilistic (and, in this sense, frequentist) approach.

We have shown that the statistical duality allows one to connect the estimation of the parameter with the measurement of the random variable of the distribution due to the Transform Eq. (10).

All considered cases of statistically dual distributions belong to conjugate families (which are defined in the framework of Bayesian approach). For example [11], the distributions conjugate to Poisson distributions were built by a Monte Carlo method (i.e. in the frequentist approach). Hypothesis testing confirms that these distributions are Gamma-distributions as expected in this case.

This means that statistical duality gives a clear frequentist sense to the confidence density of the parameter and allows one to construct confidence intervals in an easy way.

Acknowledgments

We are grateful to Louis Lyons, Victor Matveev and Vladimir Obraztsov for interest and support of this work.

We also thank Glen Cowan, Bob Cousins, Yuri Gouz, George Kahrimanis, Alexandre Nikitenko and Claudia Wulz for very useful discussions. S.B. would like to thank the Organizing Committee of PhyStat05 for having organized this interesting Workshop which is the wonderful opportunity for exchange of ideas. This work has been partly supported by grants RFBR 04-01-97227 and RFBR 04-02-16020.

References

1. S.I. Bityukov, N.V. Krasnikov, V.A. Taperechkina, *Preprint IFVE* 2000-61, Protvino, 2000; also, *e-Print:* hep-ex/0108020, 2001.
2. S.I.Bityukov, *JHEP* 09 (2002) 060; also, *e-Print:* hep-ph/0207130, 2002; S.I. Bityukov, N.V. Krasnikov, *Nucl.Inst.&Meth.* **A502**, 795 (2003).

3. B. Efron, *Stat.Sci.* **13** 95 (1998).
4. R.A. Fisher, *Proc. of the Cambridge Philosophical Society* **26**, 528 (1930).
5. S.I. Bityukov, V.A. Taperechkina, V.V. Smirnova, *e-Print:* math.ST/0411462, 2004.
6. E.A. Barkova, S.I. Bityukov, V.A. Taperechkina, *e-Print:* math.ST/0507452, 2005.
7. S.I. Bityukov, N.V. Krasnikov, in Proc. of Max-Ent'05, August 2005, San Jose, CA, USA, AIP Conference Proceedings **803** 398 (2005).
8. E.T. Jaynes: Papers on probability, statistics and statistical physics, Ed. by R.D. Rosenkrantz, D.Reidel Publishing Company, Dordrecht, Holland, 1983.
9. A.G.Frodesen, O.Skjeggestad, H.Toft, *Probability and Statistics in Particle Physics,* UNIVERSITETS-FORLAGET, Bergen-Oslo-Tromso, 1979.
10. R.D. Cousins, *Am.J.Phys.* **63** 398 (1995) 398.
11. S.I. Bityukov et al., *Nucl.Instr.&Meth.* **A534**, 228 (2004); also, *e-Print:* physics/0403069, 2004.