

Statistically dual distributions and conjugate families

Sergei I. Bityukov

Institute for High Energy Physics, Protvino, Russia

Nikolai V. Krasnikov

Institute for Nuclear Research RAS, Moscow, Russia

Plan

- Introduction
- Statistically dual distributions
- Confidence density and confidence interval
- Reconstruction of confidence density
- Conjugate families
- Conjugate families and confidence density
- Conclusion

Introduction

The notion of **statistically dual distributions** was recently introduced.

As shown ([Bityukov, 2004](#); [Barkova, 2005](#)), several pairs of distributions ([Poisson and Gamma](#), [normal and normal](#), [Cauchy and Cauchy](#), [Laplace and Laplace](#)) are statistically dual distributions

These distributions allow to exchange the parameter and the random variable, conserving the same formula for the distribution of probabilities.

The usage of statistically dual distributions in statistical inference often allows to reconstruct **the confidence density** (and, hence, confidence intervals and limits) **of the parameter** for many distributions by the easy way.

According to [B. Efron \(B. Efron, 1996\)](#) the confidence density is **the fiducial (R.A. Fisher, 1930) distribution**, which is considered as a genuine a posteriori density for parameter without prior assumptions, **of the parameter**.

Note, that the posterior distribution of the parameter also is used for the definition of **the conjugate families**.

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**.

All pairs of considered here statistically dual distributions (Poisson and Gamma, normal and normal, Cauchy and Cauchy, Laplace and Laplace) obey the identity which looks like (Bityukov, 2004; also for Gamma-Poisson pair, Jaynes, 1975; Bityukov, 2000)

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

where \hat{x} is the observed value of random variable x , θ_1 and θ_2 are bounds of confidence interval for location parameter θ . In case of Gamma- and Poisson distributions the some of integrals are replaced by sums and $-\infty$ is replaced by 0.

This identity allows to reconstruct the unique confidence density $\tilde{f}(\theta|\hat{x})$ of the parameter θ in the case of single observation \hat{x} of the random variable x .

Confidence density and confidence interval

The identity (**Eq.1**) leaves no place for any other construction of the confidence intervals (θ_1, θ_2) in our case, except (**Bitjukov, 2000, 2004**)

$$P(\theta_1 \leq \theta \leq \theta_2 | \hat{x}) = P(x \leq \hat{x} | \theta_1) - P(x \leq \hat{x} | \theta_2), \quad (2)$$

where \hat{x} is a result of single observation of random variable x , $\theta_1 \leq \theta_2$ are real values and $P(x \leq \hat{x} | \theta) = \int_{-\infty}^{\hat{x}} f(x|\theta) dx$ (in the case of **Poisson** distribution instead of integral must be sum).

Note, that the right part of **this definition of confidence interval**, which determines the confidence density, has an evident frequentist meaning.

Let us show that if **Eq.1** takes place we can reconstruct only single confidence density of the parameter θ .

Reconstruction of confidence density (**I**)

Let us suppose that $\tilde{f}(\theta|\hat{x})$ is the confidence density of parameter of the distribution if observed value of random variable x is equal to \hat{x} . As it follows from **Definition 1**, the $\tilde{f}(\theta|\hat{x})$ is the pdf of the variable θ by definition. As a result we may consider $\tilde{f}(\theta|\hat{x})$ as a posteriori density of the parameter θ in case of the observation \hat{x} of random variable x .

Reconstruction of confidence density (II)

On the other hand: if $\tilde{f}(\theta|\hat{x})$ is not equal to this confidence density and the confidence density of the parameter of our distribution is the other function $h(\theta|\hat{x})$ then there takes place another identity

$$\int_{\hat{x}}^{\infty} f(x|\theta_1)dx + \int_{\theta_1}^{\theta_2} h(\theta|\hat{x})d\theta + \int_{-\infty}^{\hat{x}} f(x|\theta_2)dx = 1 \quad (3)$$

This identity is correct for any real $\theta_1 \leq \theta_2$ and \hat{x} too. The first and third terms in the left part of this identity determine the boundary conditions of the confidence interval.

If we subtract **Eq.3** from **Eq.1** then we have

$$\int_{\theta_1}^{\theta_2} \tilde{f}(\theta|\hat{x}) - h(\theta|\hat{x})d\theta = 0. \quad (4)$$

We can choose the θ_1 and θ_2 by the arbitrary way. Let us make this choice so that $\tilde{f}(\theta|\hat{x})$ is not equal $h(\theta|\hat{x})$ in the interval (θ_1, θ_2) and, for example, $\tilde{f}(\theta|\hat{x}) > h(\theta|\hat{x})$ and $\theta_2 > \theta_1$. In this case we have

$$\int_{\theta_1}^{\theta_2} \tilde{f}(\theta|\hat{x}) - h(\theta|\hat{x})d\theta > 0, \quad (5)$$

and we have contradiction. Hence $\tilde{f}(\theta|\hat{x}) = h(\theta|\hat{x})$ everywhere except, may be, a finite set of points.

As a consequence, the reconstruction of the confidence density $\tilde{f}(\theta|\hat{x})$ of the parameter θ for given distribution is an unique.

Conjugate families

The fundamental difference between Bayesian and classical statistics is that in Bayesian statistics unknown parameter are treated as random variable, and that the use of Bayes' theorem requires the specification of prior distribution for this parameter. Correspondingly, the definition of conjugate families is based on using of probability density of parameter distribution (Casella, 2001).

Definition 2: Given a family F of pdf's $f(x|\theta)$ indexed by a parameter θ , then a family, Π of prior distributions is said to be **conjugate** for the family F if the posterior distribution of θ is in the family Π for all $f \in F$, all priors $\pi(\theta) \in \Pi$ and all possible data sets x .

Let us rephrase the given definition.

A family of distributions of the parameter θ is said to be a conjugate family of distributions if it is closed under sampling, i.e. that if the prior distribution of θ belongs to this family, then for any sample size and for any value of the observation of the sample, the posterior distribution of θ belongs to the same family.

Conjugate families & confidence density (I)

We can apply this definition to sample size 1, which is taken from one of considered here distributions.

As it follows from identity (Eq.1), the posterior density of the parameter θ in case of single observation \hat{x} is the density of statistically dual distribution to distribution of the random variable x and, consequently, this posterior density is an unique confidence density.

It is due to the fact that the identity (Eq.1) takes place for any real values $\theta_1 \leq \theta_2$ and the reconstruction of the confidence density is unique.

As known, the pairs of distributions **Gamma-Poisson** and **normal-normal** belong to conjugate families. The reconstruction of a posteriori density of parameter as an unique confidence density gives the evidence that the pairs of distributions **Cauchy-Cauchy** and **Laplace-Laplace** also belong to conjugate families.

Conjugate families & confidence density (II)

Let us consider, as an example, the **Cauchy** distribution. We suppose in our approach that the parameter θ is not random value and before measurement we can not prefer any of values of this parameter, i.e. possible values of the parameter have equal probability and $\pi(\theta) = \text{const}$. Suppose we observe \hat{x}_1 and update our prior via **Eq.1** to obtain $\tilde{f}(\theta|\hat{x}_1)$, which is the pdf of **Cauchy** distribution. This becomes our new prior before observing \hat{x}_2 . It is easy to show that in the case of the observing \hat{x}_2 the reconstructed confidence density (or our next new prior) $\tilde{f}(\theta|\hat{x}_1, \hat{x}_2)$ also is the pdf of **Cauchy** distribution. By induction this argument extends to sequences of any number of observations.

The way for construction of confidence density of mean value of several random variables which obey the **Poisson** distribution (sample with size > 1) is shown in refs. (**Bityukov, 2002**).

Conclusion

We formulate the notion of the statistically dual distributions in frame of the probabilistic (and, in this sense, in frequentist) approach.

We show that the statistical duality often allows to connect the estimation of the parameter with the measurement of the random variable of the distribution.

All considered cases of statistically dual distributions belong to conjugate families (which are defined in frame of Bayesian approach). For example (Bityukov, 2004B), the distributions conjugate to Poisson distributions were built by Monte Carlo method (i.e. in frame of frequentist approach). The hypotheses testing confirms that these distributions are Gamma-distributions as expected in this case.

By this means the statistical duality gives the clear frequentist sense to **the confidence density** of the parameter and it allows to construct the confidence intervals by the easy way.

It seems likely that the statistical duality can be considered as one of the possible indicators of conjugate families.

Acknowledgments

We are grateful to **Kevin Knuth**, **Victor Matveev** and **Vladimir Obraztsov** for interest and support of this work.

We also thank **Serge Bityukov**, **Yuri Gouz**, **George Kahrmanis**, **Alexandre Nikitenko**, **Vera Smirnova** and **Vera Taperechkina** for very useful discussions.

S.B. would like to thank the **Organizing Committee of MaxEnt'05** for having organized this interesting Workshop which is the wonderful opportunity for exchange of ideas.

The participation of **S.B.** in the Workshop has been supported by the **E.T. Jaynes Center**, the **Organized Committee of MaxEnt 2005** and by the grant **RFBR 05-01-10769**.

References

- E.A. Barkova, S.I. Bityukov, V.A. Taperechkina (2005), Statistical duality of the Laplace distribution, e-Print: math.ST/0507452, 2005.
- S.I. Bityukov, N.V. Krasnikov, V.A. Taperechkina (2000), Confidence intervals for Poisson distribution parameter, Preprint IFVE 2000-61, Protvino, 2000; also, e-Print: hep-ex/0108020, 2001.
- S.I. Bityukov (2002), On the Signal Significance in the Presence of Systematic and Statistical Uncertainties, JHEP 09 (2002) 060; e-Print: hep-ph/0207130.
- S.I. Bityukov, N.V. Krasnikov, Signal Significance in the Presence of Systematic and Statistical Uncertainties, Nucl.Inst.&Meth. A502 (2003) 795-798
- S.I. Bityukov, V.A. Taperechkina, V.V. Smirnova (2004), Statistically dual distributions and estimation of the parameters, e-Print: math.ST/0411462, 2004.
- S.I. Bityukov, N.V. Krasnikov, V.A. Taperechkina, V.V. Smirnova, Statistically dual distributions in statistical inference, in preparing.
- S.I. Bityukov, V.A. Medvedev, V.V. Smirnova, Yu.V. Zernii (2004B), Experimental test of the probability density function of true value of Poisson distribution parameter by single observation of number of events, Nucl.Instr.&Meth. A534 (2004) 228-231; also e-Print: physics/0403069, 2004.
- G. Casella and R. L. Berger (2001), Statistical Inference (2nd Edition). Duxbury Press, 2001.
- B. Efron (1996), R.A. Fisher in the 21st Century. Stat.Sci. 13 (1998) 95-122.
- R.A. Fisher (1930), Inverse probability. Proc. of the Cambridge Philosophical Society 26 (1930) 528-535.
- E.T. Jaynes (1975), E.T. Jaynes: Papers on probability, statistics and statistical physics, Ed. by R.D. Rosenkrantz, D.Reidel Publishing Company, Dordrecht, Holland, 1983, p.165.