

Estimating the mean of a small sample under the two parameter lognormal distribution

Peter Hingley
European Patent Office
phingley@epo.org

Keywords: Estimation model, Human birth weights, Lognormal distribution, Maximum Likelihood Estimate, Probability distribution, Simulation.

When making statistical inferences about the means of small samples, the confidence limits for the mean are often calculated assuming a normal distribution. But many biological variables follow the lognormal distribution [1], for example the birth weights of babies (EG data at [2]). Here, sampling distributions (probability density functions) are found for the maximum likelihood estimates (MLE) of sample mean and variance when data are lognormally distributed. They are derived analytically, making some use of the Technique for Estimator Densities (TED) [3], and then checked by using simulations with random numbers.

For an I.I.D. sample of size n with lognormal estimation, the sample mean has a lognormal distribution that is conditional on the variance. The distribution of the sample variance of the logarithms follows the usual transform of the central chi-squared distribution. The joint distribution of the sample mean and variance shows the extent to which the mean is affected by the variance.

When a normal distribution is wrongly used for estimation on lognormally distributed data, the sample mean still has a lognormal distribution. But the distribution for the MLE of the variance differs. From the distribution for one observation, that for larger sample sizes can be approached by using convolutions. The assumption of a normal estimation model biases the confidence interval for the mean. There is a discussion of the extent to which this is of practical importance when estimating means for small samples of birth weights and other lognormally distributed data sets.

References

- [1] N. Johnson, S. Kotz, S. Balakrishnan, *Continuous Univariate Distributions*, Vol. 1, 2nd edition, Wiley, 1994.
- [2] V. Iannelli, *Baby Birth Weight Statistics - Growth and Development Basics*, www.verywell.com/baby-birth-weight-statistics-2633630 , 2016.
- [3] P. Hingley, *Applications and Extensions of a Technique for Estimator Densities*, IAENG Journal of Applied Mathematics, 39:1, 2009.