

BOOTSTRAP FOR COMPARING TWO POPULATIONS

Ilir Palla

“Fan Noli” University, Korçë, Albania
email: pallailir@yahoo.com

Abstract

Statistical inference is based on the sampling distributions of sample statistics. The bootstrap is first of all a way of finding the sampling distribution, at least approximately, from just one sample. A sampling distribution is based on many random samples from the population. In place of many samples from the population, create many resamples by repeatedly sampling with replacement from this one random sample. The sampling distribution of a statistic collects the values of the statistics from many samples. The bootstrap distribution of a statistic collects its values from many resamples. The bootstrap distribution gives information about the sampling distribution. In this paper we have compared means of two populations, such as weight of fish in natural habitat and weight of fish in tanks. We have used bootstrap method to estimate their mean and standard deviation. We have estimated the difference of population means $\mu_1 - \mu_2$. When both populations are roughly normal the two-sample t procedures compare the two population means. The bootstrap can also compare two populations, without the normality condition. Samples are taken from both populations are independent and with different size. We have estimated the difference of population means where the samples are small sized and are skewed. For further accuracy we have used the permutation test even when the sampling distribution is skewed. The data of this case are simulated. To get results we have used R program, which has automated proper bootstrap procedures. In practice the population usually is not a normal distributions, we may suggest using the bootstrap method which does not require this condition of normality to get statistical inferences.

Keywords: *sampling distribution, bootstrap method, two-sample, permutation test.*