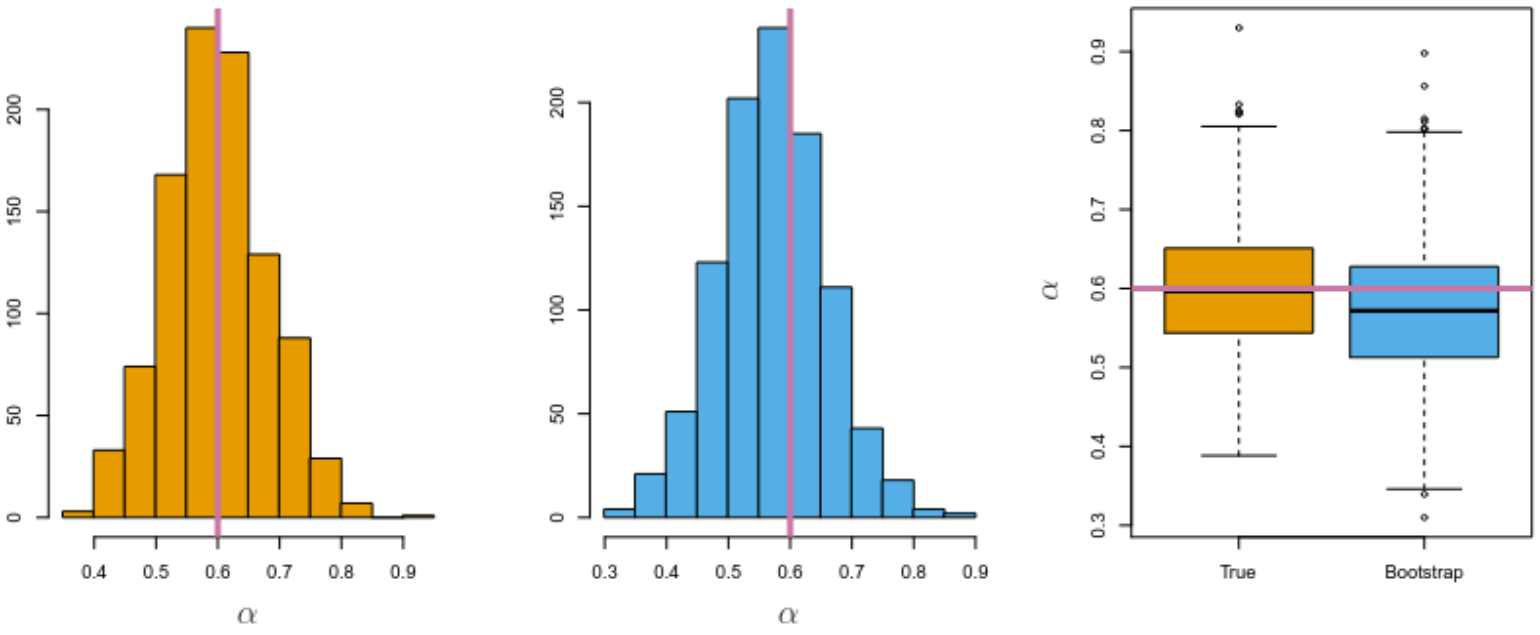
The Bootstrap

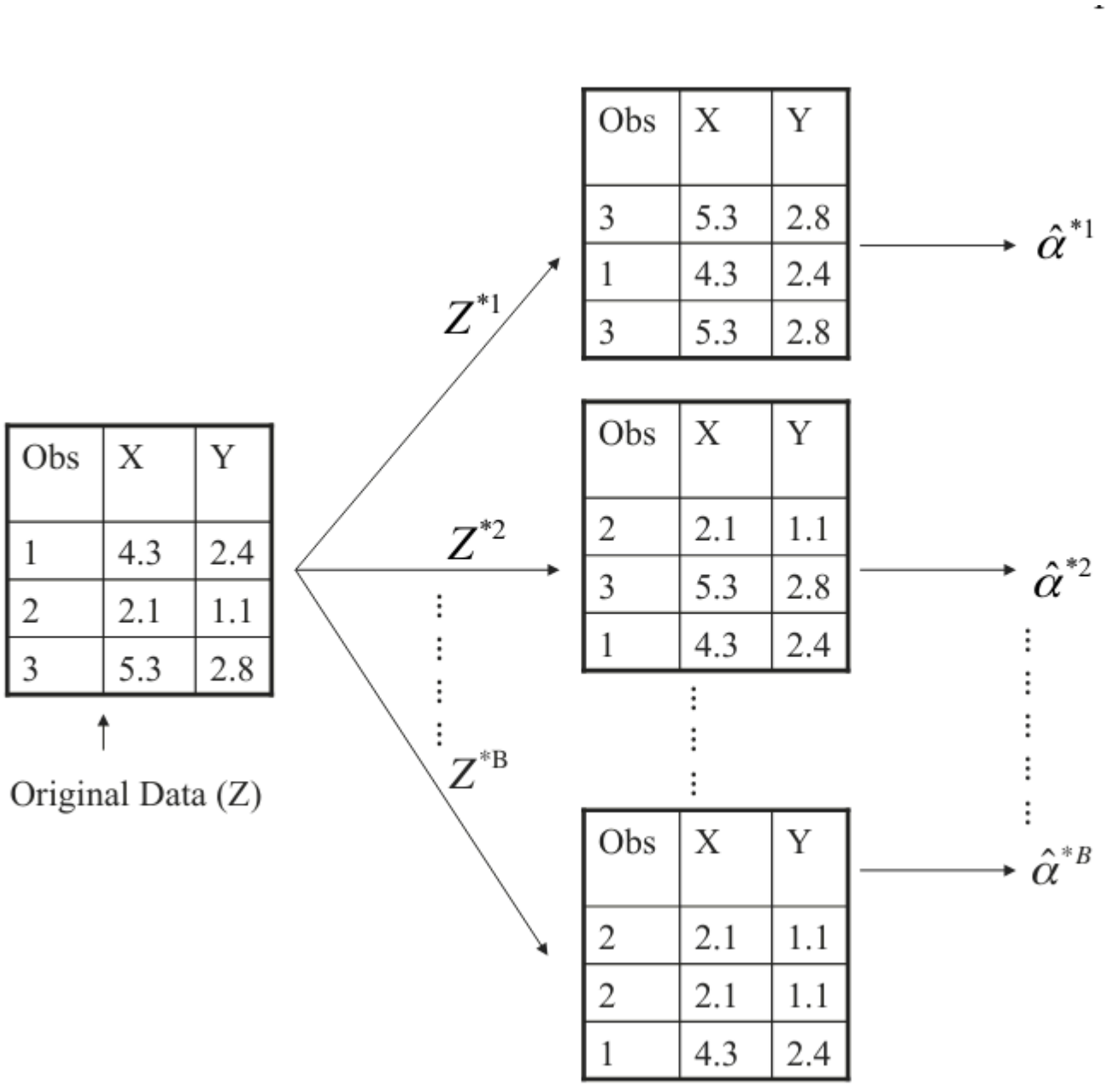
Its a statistical tool part of resampling techniques to estimate the uncertainty of :

- Mean μ
- Variance σ
- Model performance

By randomly Sampling with **replacement** many times to create multiple new datasets and these datasets are used to training and evaluate the learning models, Which is the same principle in statistics where estimated values are calculated based on multiple samples and each sample is called **Bootstrap sample**.



- Left : is a histogram of the estimates of α obtained by generating 1000 data from the true population
- Center : A histogram of α estimates from 1000 bootstrap samples from a single data set
- Right : Boxplot of both estimates which clearly shows how accurate bootstrap can be



- Graphical illustration of bootstrap on a data set with 3 observations which generate samples with **Replacement** (Which means same observation can occur more than once in the bootstrap)
- Each of the bootstrap sample generate an estimate α and by computing the standard error of these estimates