# A Databased Measure for Statistical Inference

by
Walid A. Nuri, Ph.D.
Independent Statistician
walid.nuri@gmail.com

**Abstract**
This paper presents an alternative to the probabilistic approach in measuring events of a sample space and in defining a method making a decision based on a specific statistical decision rule. It starts by introducing the concept of a population image that is created using the measurements of the provided random sample (selected for inference about a population parameter). . Using the population image, this approach employs a computer simulation technique to select (with replacement) a large number m, say, of random samples from the measurements Then, for an event E in a sample space, the function $\eta_m(E)$ is defined as the number of cases within these samples that support the event E. The measure $\eta$ is defined as the limiting value of $\eta_m(E)$ as $m \rightarrow \infty$. It is called a measure of favorability. Statistical inference problems frequently encountered in applications are discussed using this new approach to measuring events. Actual examples are given to show how to apply this approach to these inference problems.

**Key Words:** simulation, supporting event, population image, measure of favorability, optimal estimator

## 1. Introduction

Statistics emerged as an important field of science more than a hundred years ago. A great number of statisticians have contributed to the advancement of statistical theory and methods that are used in many fields of application.

From the beginning scientists began to think of a subjective approach in building a decision rule that would provide strong evidence that their decision about a parameter of a population is sound. Such a decision is based on information gathered from a random sample together with some prior knowledge of the characteristics of the population (through a history of observations) such as the nature of the measurements and of the sampling units. That evidence was thought of in terms of probability. Thus, it was necessary to construct a probability model suitable to the measurements of the population under study. The majority of statistical methods used in applications fall in the area of parametric statistics. A second major branch is nonparametric statistics, in which statisticians do not assume a probability model for the population under study. However, these methods are limited in their applications and in some cases can only be applicable under specific conditions, such as the assumption of a large sample size. These specific conditions used in nonparametric statistics result in approximating a decision ule in order to acquire a probability structure.

Assuming a probability model suitable to the measurements of a random sample is essentially idealizing the real world. The task of giving data a realistic probability structure becomes complicated when one encounters a fairly complicated real-world inference issue. One may need to deal with an issue in which multiple measurements must be observed to support an inference about multiple parameters in a study. In this case statisticians may attempt a multivariate distribution, and in using that approach, they fall into a vast probability and mathematical endeavor leading to a purely theoretical argument far removed from the realistic situation. Thus, statisticians who deal withstatistical methods in the applied sciences have found that not many scientists have an in-depth understanding of probabilistic thinking when it comes to the statistical interpretation of the analysis of their research results. It is really hard for those scientists to make the connection between their real-world science and the concepts of probability.

Uppermost in our minds should be the fact that a *well*-selected random sample to support an inference about a parameter of a population must reflect all characteristics of that population. The information gathered from that random sample should lead to the creation of a statistical decision rule about the parameter under study. This fact has led statisticians over the past 50 years to employ *simulation techniques*, with the aid of computers, to reach decision rules about population parameters. During all these years statisticians have argued and discussed the validity in assuming a probability structure for data of any given population. Since the time of Fisher (1956) statisticians have not been totally convinced that introducing an idealized probability approach would represent truly the behavior of population data.

Many statisticians have tried to apply computer simulation techniques to interpret the results of a random sample, and many attempts have been made to validate numerically the application of estimation and testing hypothesis methods (created in the probability approach). The *bootstrap* and *jackknife* methods [two resampling techniques introduced by Efron (1979, 1981)] are two early methods which show that a variety of statistical questions can be answered numerically. These methods are used to repeatedly select (with replacement) random samples from a given sample. Resampling in the bootstrap method is confined to the total measurements of the random samples, whereas the jackknife method selects a number of subsets from a given random sample on which to perform the resampling.

These resampling methods have been applied by many statisticians to a variety of fields, such as economics, business, psychology, and education (Ang, 1998; Fan, 1996). Moreover, developers (including companies like SAS and SPSS) have produced software to facilitate using computers to simulate and apply these new methods.

In reviewing all these distinguished accomplishments, it is evident that statisticians generally regard these resampling techniques as mere aids to validate the applicability of the wide range of statistical methods that have been created using the probability measure to describe events of the sample space. However, to do justice to the power of these computer-aided techniques, one should define a new measure (instead of the probability measure) to describe the events of a sample space. With this new measure, these resampling methods could then be adopted as a general method for statistical inference.

The conclusion one may draw from this discussion is that there must be an alternative to our classical probabilistic approach in statistical inference and in creating decision rules. The alternative approach presented here introduces a measure, called a *favorability measure*, to describe events of a sample space. This measure is also linked to what is called here a *supporting event* that will provide strong evidence to support the choice of a decision rule.

When one makes a decision, there must be some events that support that particular decision. If one is observing a series of events and the majority of events all point to that same decision, then we have a supporting event for our decision. This is exactly the way we deal with decision-making in everyday life.

In statistics, by observing, for example, a value of 52 (in some units) for a sample mean of a random sample (which happens to be a good estimate of the population mean in many applications), one may make the decision that the true population mean is more than 50. Suppose that we have an extensive series of random samples that are selected from the particular population (in the same manner as the first sample) and they all result in means greater than 50. Then the event, which is the union of all those events, is a *supporting event* which provides strong evidence for the validity of our decision

In practice we usually do not have an extensive series of random samples. Instead, as in the present approach, a computer simulation technique is used. In this process a random numbers generator is employed to select (with replacement) a series of random samples from the measurements of a provided random sample (selected for inference about a parameter for a population). This type of statistical analysis was certainly the basis for motivating statisticians to use computer-aided techniques such as resampling techniques. The next section introduces a new measure for this approach and provides details of random samples. The random samples are selected from not only the given random sample, but also from what is going to be called a *population image*. This image is defined as the collection of sample sets (plus the given sample) that is created by shifting the measurements of the given sample by some increments.

## 2. The Method

When analyzing data collected in a random sample, it is assumed that this sample was selected according to the rules of sampling techniques. This data comprises a finite number of measurements representing all plausible measurements of the particular population. The measurements of a population can be either countable or uncountable, but in practice a population usually contains a countable number of plausible measurements. This is how the measurements of a random sample and its population are envisioned in this paper.

Also, in this paper prior information and data obtained in previous samples can be added to the data at hand so long as those previous samples were collected under conditions similar to those of the current sample. In this case the total information gathered from all samples would give us a sufficient source of information to draw whatever inferences we are focusing on about the parameters of the particular population. Thus, when *a random sample* is mentioned later on in this paper, it is understood as having these characteristics.

Since the inference approach proposed in this paper employs simulation techniques, some definitions are needed. The process of providing a series of random samples is carried out by first selecting a random sample, $S_0$, of, say, n measurements (from the provided sample that was selected from a population). From this sample, other samples are created by shifting the measurements by an amount -d or +d, $d=jt_1$, where $j=1, \ldots, r$ ($rt_1=5SD$), and $t_1$ is a fraction of the measurements units. Let $I_{mg}$ be the collection of all the created samples plus the given sample. $I_{mg} = \{S_0, S_1, S_2, \ldots, S_{2r}\}$. We call this $I_{mg}$ an image of that population.

Consider an event E in a sample space. We randomly select, from each of the s(=2r+1) sets, a large number of samples (m, say, >1000), with replacement and with sizes <n. Then, the following function is defined for the event E:

$$\square\eta_{m,s}(E) = (1/ms) \sum_{i=1}^{m} \sum_{j=1}^{s} g_{i,j} \ ,$$

where $g_{i,j}$ indicates the proportion of cases that favors the event E for the ith sanmple within the jth set. The proposed function for measuring an event E is η, defined by

$$\eta(E) = \text{Limit } \{\eta_{ms}(E)\}, \text{ as } m \to \infty \ .$$

It is clear that $0 \le \eta(E) \le 1$ for every event E in a sample space. Using these definitions, one can prove the following theorem:

Theorem 1. The function η is a *measure* defined on a sample space such that $0 \le \eta(E) \le 1$ for every event E in the sample space. We call this measure, a measure of favorability.

To make the connection between this measure and the probability measure, consider a random variable Y representing the measurement of a random sample that is collected for some study and let the event G = {y : Y ≤ x} for some real value x. Then, by applying the process of selecting m random samples from the population image as described above, one can show that

$$\eta_{m,s}(G) = (1/ms) \sum_i \sum_j F_{i,j}(x),$$

for every value x. We know that $F_{i,j}(x)$, in the probabilistic approach, is the *empirical distribution* of X for the ith sample within the jth set. One can prove that

$$E[\eta_{ms}(G)] = F(x)$$

and

$$Var[\eta_{ms}(G)] = (1/ms) F(x)[1-F(x)]$$

Therefore, the following theorem can be proved.

Theorem 2: $\eta_{ms}(G) \to F(x)$ in probability as $m \to \infty$, where F(x) is the cumulative distribution function of x and G is an event defined as G = {y : Y ≤ x} for a random variable Y representing the measurement of a sample.

This theorem proves that one can stop here at using the measure η and need go no further to assume a probability structure for the data of a population. Having this measure for events that cover specific measurements in a population becomes the basis for creating methods for inference about parameters of that population. This theorem also proves that η provides a sufficient tool for solving inference problems that otherwise may not easily be solved with confidence when one uses a specific probability model for the measurements of a population.

The above results confirm the fact that a *well-selected* random sample for a specific inference goal would contain all the information necessary for creating decision rules to serve the particular inference objectives.

Because of the nature of the measure, $\eta$, as the limiting value of $\eta_m(E)$ (as $m \to \infty$) for an event E of a sample space, for the rest of this discussion the following convention is used. The term $\eta(E)$ shall refer to a very large number of simulations, M, say, such that any increase beyond M would not significantly affect the value of $\eta$.

From experience we know that when the number of random samples M is more than 5,000 we see no significant difference in the values of $\eta$ as M increases. However, one may choose to select any number of random samples beyond 5,000. Needless to say, it would take only a fraction of a second to generate that large a number of samples using a PC. Also, it takes only a few seconds to run and get results for each of the above-mentioned statistical applications. PASCAL programming language was used to generate random samples and to apply this method to a number of inference problems, as shown later in this paper.

## 3. The Method in Practice

The method proposed in this paper is demonstrated by describing in detail its application to three statistical inference issues. The first inference issue is parameter estimation; the second issue is estimating the sample size to ensure a specific relationship between a parameter and its estimate; and the third issue is testing hypotheses about population parameters.

### 3.1 Parameter Estimation

A different approach from the classical probabilistic approach is presented in this paper for estimating a parameter of a population. As mentioned above, when selecting a random sample for an inference about a parameter in a population under study, we know that there are only a countable number of plausible measurements, whether we are dealing with discrete or continuous measurements.

Consider a population of size N plausible measurements. Let $\theta$ be a population parameter and $\theta_n$ be its estimate provided by a sample of size n that is randomly selected from this population. Suppose that in a simulation process we select M random samples (with replacements), each of a size <n. Let the event $A_n$ be defined as

$$A_n = [\ \{ \theta_n \}_M : \mathrm{l}\ \theta_n - \theta\ \mathrm{l} \leq \delta\ ],$$

where $\{ \theta_n \}_M$ indicates the estimates from the M random samples and $\delta$ is a real positive number. Then $\theta_n$ is said to be an *optimal estimate* of $\theta$ if Limit $\eta(A_n) \to 1$, as $n \to \infty$.

### 3.2 Sample Size Estimation

In conjunction with the present approach of providing estimates by employing simulation techniques for selecting random samples, the population image needs to be enlarged by increasing r and by decreasing the frction of unt of measurements ($t_1$) to provide a large set of measurements for the population image described above. This extension makes sense when one remembers that the characteristics and statistical properties of a population are learned through a history of observations and a series of random samples. One advantage in creating this enlarged population image is to be able to select random samples with larger sample size than the provided sample, The important question now is how to determine the size of the sample that needs to be selected for a particular inference goal. This is explained below.

Suppose that $\theta_n$ is an optimal estimate of a population parameter $\theta$ and we need to find the sample size such that *with high favorability* $\theta_n$ will be at a distance d from the parameter $\theta$. Here, the high favorability is measured by a value for $\eta$ of at least 0.95. Select a large number, say, M, of random samples from the created *population image*, with different sample sizes $n_i$ , $i = 1,2,\dots, s$, where s is an integer. Let the event B be defined as $B = [\ \{ \theta_n \}_M : | \theta_n - \theta | \le d\ ]$. Then, we want to try different values of the $n_i$'s until one of them satisfies $\eta(B) = \psi$, where $\psi$ is the specified level of high favorability which is limited, in this paper and for convenience, to two values, either 0.95 or 0.99.

### 3.3 Testing Hypotheses

This paper's inference approach uses terminology for testing hypotheses similar to that used in the probabilistic approach. A null hypothesis is denoted by $H_0$ and an alternative hypothesis will be denoted by $H_1$ . The hypothesis statement for a parameter is given by:

$$H_0: \theta \in \omega \ \text{ vs. } \ H_1: \theta \in \Omega - \omega,$$

where $\Omega$ is the parameter space and $\omega$ is a subset of $\Omega$. Note that $\Omega$ is defined on the real line and that all subsets of $\Omega$ are in the form of intervals, e.g. $\omega = (\theta_1, \theta_2)$.

Suppose that from a random sample, $\theta_n$ is an optimal estimator of a parameter $\theta$. Then, we need to find a rejection region R, say, to establish the following decision rule, according to the above hypothesis statement. Reject $H_o$ if $\theta_n \in R$ and accept $H_o$ if $\theta_n \in S - R$ , where S is a sample space.

Two types of errors occur in testing hypotheses. The first error is termed Error 1, $e_1(\theta)$, say, and Error 2 is $e_2(\theta)$. They are defined as: $e_1(\theta) = \eta$ [rejecting $H_0 | H_0$ is true, $(\theta \in \omega)$], and $e_2(\theta) = \eta$ [accepting $H_0 | H_0$ is not true, $(\theta \in \Omega\text{-}\omega)$].

In practice three hypotheses are commonly used:

$$H_0 : \theta \le \theta_0 \ \text{ vs. } \ H_1 : \theta > \theta_0 \ ,$$

$$\text{or } H_0 : \theta > \theta_0 \ \text{ vs. } \ H_1 : \theta \le \theta_0 \,,$$

$$\text{or } H_0 : \theta_1 < \theta \le \theta_2)], \text{ vs. } \ H_1 : \theta \le \theta_1 \text{, or } \theta > \theta_2$$

Then $e_1(\theta)$ is maximized at some point $\theta$ in the specified interval. In the first two hypotheses the maximum is at $\theta_0$ . Let $e_1(\theta_0) = a_1$, which is usually given the values between 0.01 and 0.05. On the other hand, error 2 is a decreasing function of $\theta$ as $\theta$ moves away from $\theta_0$ and $e_2(\theta) \to 0$ as $|\theta| \to \infty$.

The rejection region is defined by a critical value, c, which is determined as described here. Following the simulation process in applying the measure $\eta$, a large number of random samples, m, are selected from each one of the data sets of the population image. Then, the estimates $\theta_n$ for $\theta$ are calculated, after adjusting for the null hypothesis, and then arranged in ascending order within each of the m samples. The rejection regions for the above three hypotheses include the estimates that are: $>c$, $\leq c$ , and ($\leq c_1$ ,or $> c_2$), respectively.

Thus, one can see that if, for instance, $a_1 = 0.05$, the critical values for the above hypotheses are the averages of the specific percentiles over the sample sets: the 95th percentile first hypothesis, the 5th percentile for the second, and (the 2.5th percentile or the 97.5th percentile) for the third.

## 4. Examples of Inference Issues

In this section some inference issues that are frequently dealt with in statistical applications are discussed using the method introduced in this paper.

### 4.1 Estimating a Population Mean μ

The following example shows that a sample mean $\bar{x}$ is an optimal estimate, as defined above, of the population mean μ.

*4.1.1 Example*

A population of size N = 7380 measurements is created. The population mean is $\mu = 97.6$ and variance $\sigma^2 = 270.9$. From this population, m = 1000 random samples were selected for different sample sizes and with different values for δ that were attempted. The attempted values for δ are $0.1\sigma$, $0.05\sigma$, and $0.03\sigma$ The following results show that the identity

$$\eta[\ \{\bar{x}\}_m : |\bar{x} - \mu| \leq \delta\ ] = 1$$

was satisfied when the sample sizes, corresponding to the above δ values, reached 1000, 2500, and 6000, respectively. This simplified example shows that $\bar{x}$ is an optimal estimate of the true population mean μ.

### 4.2 Estimating the Sample Size n so that the Sample Mean is Some Specified Distance from the Population Mean μ.

Suppose that we need to find the random sample size n, so that with *high favorability* the sample mean would be at a distance, d, say, from the true population mean μ. This favorability, as explained above, is determined by the measure η, which is given here a value of 0.95. Then, for the event E

$$E = [\{\bar{x}\}_M\ : |\bar{x} - \mu| \leq d\ ], \text{ with } \eta(E) = 0.95.$$

This procedure starts by characterizing the *population image*, which is determined from prior information using previous samples from the population under study. The sample mean and variance of this sample is treated as the population mean μ and variance $\sigma^2$, respectively. Then, the process of simulation is carried out by selecting a large number of random samples of different sizes from the population image. The mean is calculated for each of these generated random samples. This process of simulation continues until one particular sample size satisfies the goal value of the measure η, defined above. The process of finding the sample size discussed above can be illustrated by the following example.

*4.2.1 Example*

The average diastolic blood pressure (DBP) in a random sample of 140 patients was 104.2 mmHg with variance equal to 227.01 $(mmHg)^2$. The population image was constructed¸ as described above, to determine the sample size needed so that the sample mean would fall within 2 mmHg of the population mean. Different sample sizes were tried, and for each size 10,000 random samples were selected from the population image. It was found that the size reaches 215 patients when 94.9% of the means were within that specified interval, i.e. when $\eta(E)$ equals 0.949. Note that assuming a normal population for the blood pressure measurements, this sample size comes out to be 217, for which one would expect the sample average to be within that specified interval with a probability of 95%.

### 4.3 Testing Hypotheses About the Mean μ of a Population

Three cases of testing hypotheses are discussed in this section plus a case of a binomial population, all with examples. The method described below can also be applied for other population parameters such as variance, standard deviation, and range.

*4.3.1 Testing a one-sided hypothesis*

This testing problem is one of many examples in which the sample mean is an optimal estimate of the population mean $\mu$.

$$H_o : \mu \leq \mu_o \quad \text{vs.} \quad H_1 : \mu > \mu_o \, .$$

Here, $e_1(\mu_0)$ is a maximum for all $\mu < \mu_0$, as stated above. This will lead us to establish a rejection region R by following the same process of simulation given above. Again, in this paper the traditional error values 0.05 (or 0.01) are also used to determine the critical value for rejecting or accepting the null hypothesis. The example below shows the results of applying the proposed method.

*4.3.1.1 Example*

The average score of 200 county high school students on a state-wide test is 70.4 (with variance equal to 78.38). From the results of this sample, one would like to see if the performance of the students in that county reaches the state minimum requirement of 70 or above on that test. If we let $\mu$ be the true state average score on that test, then we can test the following hypothesis:

$$H_o : \mu \leq 70 \quad \text{vs.} \quad H_1 : \mu > 70$$

Twenty-seven sets of data are created from the sample measurements (after adjusting for the null value, as described above), each of which contains $3 \cdot 200 = 600$ measurements. Next, 1000 random samples, each containing 200 measurements, are selected (with replacement) from each of the 27 sets of data. The average of the $95^{th}$ percentiles of all the 27 sets is calculated to get the critical value, which is found to be c = 71.00. Since the sample average is less than c, $H_o$ cannot be rejected.

*4.3.2 Testing a two-sided hypothesis*

For this two-sided hypothesis:

$$H_o : \mu = \mu_o..\text{vs.} \quad H_1 : \mu \neq \mu_o$$

we need to consider two critical values, $c_1$ and $c_2$, that make up the rejection region which is defined by all the sample means (in the simulation process) that fall below $c_1$ and all those that fall above $c_2$. This means that the first error, $e_1$, will be divided into two parts. The measure $\eta$ measures all the events in which the means fall below $c_1$, plus all the events in which the means fall above $c_2$. For practical purposes this error ($e_1$) is divided into two equal parts for the two critical values. The value of this error is usually at least 0.01. Then, the decision rule is to reject $H_o$ if the sample mean is less than $c_1$ or greater than $c_2$, and accept otherwise.

*4.3.2.1 Example*
Using the example shown in section 4.3.1, we can test the following hypothesis:

$$H_o\text{: } \mu = 68 \text{ vs. } H_1\text{: } \mu \neq 68$$

If one assumes the error to be 0.05, then $c_1$ corresponds to the 2.5$^{th}$ percentile and $c_2$ corresponds to the 97.5$^{th}$ percentile. In the simulation process it is found that $c_1$ equals 65.88 and $c_2$ equals 69.62. Since the average is more than 69.62, the above hypothesis is rejected. This means that the average score for this high school is more than 68. Assuming a normal distribution for the test score measurements, then the t-value for this example comes out to be 3.87.

*4.3.3 Testing for two population means*
The following information represents the systolic blood pressure data for two randomly selected samples of patients. The sample sizes are 42 and 43 with means of 141.14 and 140.06 and variances of 80.95 and 56.07, respectively. We then test the following hypothesis:

$$H_0 : \mu_1 = \mu_2 \text{ vs. } H_1 : \mu_1 \neq \mu_2$$

When we apply the simulation process, we get the following results. The sample sizes selected for simulation are 38 and 39, and the average of the 2.5 and the 97.5 percentiles for the means are -3.74 and 3.71, respectively. Assuming a normal distribution, the t-value is 0.597. Thus, the above hypothesis cannot be rejected.

*4.3.4 Binomial population*
In a random sample of size 100 from a binomial population, it is found that p=0.20. We want to test the following hypothesis:

$$H_o : p = 0.14 \text{ vs. } H_1 : p \neq 0.14$$

The average of the 2.5 and the 97.5 percentiles for the estimated values of p are 0.06 and 0.22, respectively, and the t-value is 1.50. Since p falls in between 0.06 and 0.22, we cannot reject the null hypothesis.

## 5. Discussion

Statistical methods were developed mainly through probabilistic thinking in measuring events and the *strength* of decision rules. Establishing a probability model suitable to the measurements of a random sample is creating an idealization of the real world. Moreover, the task of giving data a realistic probability structure becomes complex when one encounters a fairly complicated real-world inference issue. On the other hand, nonparametric statistical methods are limited in their applications. Some of these methods can only be applied under specific conditions, and they may require a large sample size in an application (sometimes larger than is practical to collect) to establish a decision rule that is defined in the framework of the probability measure. Adding to that, when applying statistical methods to various fields of research, statisticians have the experience that very few researchers have an in-depth understanding of the statistical interpretation of the results of their research. Most of them have difficulty in making the connection between their real-world research and the probability concept.

Over at least the past five decades statisticians have been assisted by computers to easily and quickly use techniques to simulate and study numerically the characteristics of various statistical methods. The *bootstrap* method is one of the first methods to show that many statistical estimates can be provided numerically and without assuming a particular probability model by generating random samples from a given sample.

The above facts have led the author to seek an alternative approach to our classical probabilistic thinking in creating a statistical decision rule for an inference about a population parameter. This approach is mainly motivated by the fact that a *well-selected* random sample designed to draw an inference about a parameter of a population should reflect all the properties and characteristics of that population. The information gathered from that random sample should lead to the creation of a statistical decision rule about the parameter under study.

The concept of an event that may support a specific decision rule is put forth. Suppose that we have an extensive series of randomly selected samples from a population and an event (which is observed in any of these samples) that seems to support that decision rule. If that event occurs within the majority of these samples, then this event, which supports that particular initial decision, is a *supporting event*. This is exactly the way we deal with everyday issues. Certainly, this concept can more easily be understood and accepted by mainstream scientific researchers than the probability concept. Thus, it is necessary to introduce an alternative measure to the probability measure. Practically speaking, one cannot provide a series of random samples that are selected from a given population for a specific inference goal. Therefore, this paper introduces a measure (denoted by $\eta$) that is defined as the proportion of cases, in a series of randomly selected samples from what is called here a *population image*, that support a particular event of a sample space.

A population image is defined as a collection of sample sets created from a given random sample by shifting all the measurements of this sample in some defined increments. A population image will result in giving us a much larger number of measurements with a wider range than found in the original random sample. This new approach employs a computer simulation technique which will generate a large number of random samples that will be selected from the population image. A random numbers generator is employed to select a large number of random samples (with replacement) from all those measurements of the created population image.

The *main objective* of this paper is to introduce the measure $\eta$ (called a *measure of favorability*) for events in a sample space. Theorem 2 of this paper supports the main finding that one can use the measure $\eta$ for statistical inference about parameters of a population and with no need to assume a probability structure for the data of that population.

Note: The findings of this paper are applied to real-life examples to support the results described for estimation and testing hypothesis. The author created a simulation program (using PASCAL language) to create and define each example's population image from real random samples, and to select a large number of random samples from these images.

## 6. References


Ang, R. P. 1998. Use of jackknife statistic to evaluate result replicability. *Journal of General Psychology* 125, 218-228.

Billingsley, P. 1995. Probability and Measure, 3rd Ed. New York: Wiley & Sons.

Diaconis, P. and B. Efron. 1983. Computer-intensive methods in statistics. *Scientific American* May, 116-130.

Efron, B. 1979. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* 7, 1-26.

Efron, B. 1981. Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika* 63, 589-599.

Efron, B. 1994. An Introduction to the bootstrap. R.J. Tibshirani. Monographs on Statistics and Applied Probability, Vol. 57. New York: Chapman & Hall.

Efron, B. 2009. Empirical Bayes estimates for large-scale prediction problems. *Journal of the American Statistical Association* 104, 1015-1028.

Fan, X. and L. Wang. 1996. Comparability of jackknife and bootstrap results: An investigation for a case of canonical correlation analysis. *Journal of Experimental Education* 64, 173-189.

Fisher, R. A. 1956. Statistical methods and scientific inference. Edinburgh: Oliver and Boyd.

Gibbons, J. D. 2003. Nonparametric Statistical Inference, $4^{th}$ Ed. New York: CRC Press, 645 pp. Subhabrata Chakraborti. A Series of Textbooks and Monographs, Vol. 168. Chapman & Hall.

Tukey, J. W. 1958. Bias and confidence in not quite large samples. *Annals of Mathematical Statistics* 29, 614.