Journal of the Nigerian Association of Mathematical Physics
***Volume* 22 (November, 2012), pp 175 – 178**

The Permutation Likelihood

Ogbonmwan S. M.

Department of Mathematics,
University of Benin, Benin City, Nigeria

Abstract

A major attribute of a Statistical principle especially for the frequentists is explored in this work. The basic and fundamental principle of the permutation test is briefly reviewed and used to generate likelihood for parameters. The promise of getting exact results is inspired by the advent of the computer. The theory and methodology of the permutation likelihood generation are described.

Keywords: bootstrap, permutation test, exact likelihood, kernel density likelihood, rank order statistics.

1.0 Introduction

In its basic form, the likelihood is a function of the parameters of statistical models which plays a key role in statistical inference. Thus, for a continuous random variable X, the likelihood function depending on the parameter Θ is expressed as

$$L(\theta|x) = f_\theta(x).$$

Several approaches have been adopted in generating likelihoods. For example, the bootstrap which was introduced by Efron [1-3] is an intensive computer-based Monte Carlo resampling method.

The bootstrap has been a useful statistical tool for the estimation of sampling distribution of random variables. Ogbonmwan and Wynn [4, 5] gave incisive discussions on the bootstrap generated likelihood. The papers sketched the theory of likelihood generation. Ogbonmwan [6] used the bootstrap technique to obtain simulated likelihoods for k ≥ 2 contracts of parameters. Furthermore, Owen [7, 8] provided related technique for empirical likelihood. Chen [9-11] engaged empirical likelihood in alliance with kernel density estimation to construct confidence interval for the value of a probability density at a given point. Holmes and Adams [12] incorporated the conventional k~ nearest neighbor algorithm to explore classification models.

Nearly in every paper cited in this text and of course more in the literature, the likelihoods generated are not in many cases exact. For example, the bootstrap configurations (samples) used to generate the bootstrap empirical likelihood is not a total permutation of all the elements in the original sample. It eventually turns out that the bootstrap estimates are based on a sample from the configurations that would be needed in the permutation approach. This is a limitation of the bootstrap approach. However, this limitation of the bootstrap does not rule out its correctness and reliability. It suffices to say that the bootstrap provides very good estimates of the permutation approach which actually provides exact results.

The purpose of this paper is to provide the theoretical framework that will take advantage of the advent of the computer, generate all the permutations necessary for the likelihood in a two sample problem. The procedures will be described in sections 3 and 4. General remarks and conclusions are made in section 5. Section 2 reviews the permutation approach.

2.0 The Permutation Approach

The permutation approach has been known to provide exact result in statistical inference (see Agresti [13] and Good [14]). The difficulty of making use of the permutation test has been the difficulty of performing intensive looping in the computer programming required for a complete enumeration for unconditional permutation of all elements in a statistical experiment. Odiase and Ogbonmwan [15] has sufficiently addressed this difficulty. Several alternatives have been put in place of the computationally intensive unconditional exact permutation but none of these alternatives provides accurate results. They all provide estimates to the exact permutation method.

Corresponding author**:** E-mail: -, Tel. +234 8056752390

In this article, the special case of a 2 x n tables in a two sample experiment with row and column totals allowed to vary with each permutation is considered. A complete enumeration of all the permutations yields maximum power than the alternatives, see Opdyke [16].

Good [14] initiated a five step approach for the permutation test. In line with the bootstrap resampling method for generating likelihood the following steps would be considered.

1. For a given set of data, choose a statistic of interest.
2. Compute the test statistic for the original data set.
3. Do a complete enumeration of all the distinct permutations of the data sets and for every permutation compute the test statistic.
4. Construct the distribution for the test statistic using the results of Step 3

Step 3 which gives a complete enumeration of all distinct permutations of the experiment is the hub of the generation of the permutation likelihood that is considered in this article. Doing a complete permutation of the observations is not without a cost. In addition to good computing skill, serious price is paid for time and space complexities. For example, a two sample experiment can be considered as follows:

Let $X_N = (X_1, X_2)$, where $N = n_1 + n_2$, $X_i = (X_{i1}, X_{i2}, X_{i3}, \cdots, X_{in})^T$, $i = 1, 2$, and n_i is the sample size for sample i. Assume that X_N is composed of N independent and identically distributed random variables. This would yield: $\frac{(n_1+n_2)!}{n_1!n_2!} = \frac{N!}{n_1!n_2!}$ possible permutations of N variates of the two samples of sizes n_1 and n_2 respectively. Each permutation is equally likely with a probability $\left(\frac{N!}{n_1!n_2!}\right)^{-1}$. For the simple case of equal sample sizes, $n_1 = n_2 = n$, the total number of complete distinct permutations is $\frac{(2n)!}{(n!)^2} = \frac{N!}{(n!)^2}$ with equal probability of $\frac{(n!)^2}{N!}$.

Odiase and Ogbonmwan [15] developed a systematic algorithm for generating all the distinct permutations for a two sample experiments. The paper adopted two approaches viz: the use of the actual observations and by making use of the ranks of the observations.

3.0 The Permutation Likelihood

Consider that the parameter being estimated by a statistic T emanates from a family of transformations of the data set $X = (X_1, X_2, X_3, \cdots, X_n)$ to Y_θ such that

$$Y_\theta = (y_1, y_2, \cdots, y_n)_\theta = g_\theta(x) \tag{3.1}$$

Assume that the transformed data set $y_1, y_2, \cdots, y_n$ are independent and identically distributed with a distribution function F, not depending on θ. Also, assume that $y_{0i} = X_i, i = 1, 2, \cdots, n$. If the statistic of interest of interest X on is defined by $T = t(x)$,then define $T_\theta = t(y_\theta)$ and take the partial likelihood for θ to be the density T_θ of at the observed value of t_θ. The permutation likelihood is obtained by considering the values of the total enumeration of all the permutations of the statistic $T(y_\theta)$ and then let the values be listed in some order say $T_{i1}^0, T_2^0, T_3^0, \cdots, T_{N^0}^0$ which will have the empirical cumulative density function (cdf) of the form:

$$P_T(t|\theta) = \frac{1}{N^0} \# \left\{T_i^0 \middle| T_i^0 \le t\right\} \tag{3.2}$$

The empirical cdf in (3.2) is smoothed to obtain a continuous density $\hat{f}_T(t|\theta)$ by using any of the density estimators. The kernel density estimator is the choice in this work. Thus, the permutation likelihood is the density of the value of the permutation T_θ^0 at t_θ, where $T_\theta^0 = t(y_\theta)$ realized from the transformed data set $(y_1, y_2, \cdots, y_n)_\theta$. An alternative and simplified version that gives the exact empirical likelihood is to do a count of the proportion of T_j^0 values that lie in some interval around j .Thus, for some $\varepsilon > 0$ the exact empirical permutation likelihood is defined as

$$L(\theta) = \frac{1}{N^0} \# \left\{T_j^0 \middle| t - \varepsilon \le T_j^0 \le t + \varepsilon\right\} \tag{3.3}$$

3.1 The Two Sample Case

Consider a two sample experiment in which every individual belongs to one and only one of two distinct samples with means μ_1 and μ_2 (say). Suppose interest lies in the estimation or test of hypothesis about the means. In particular, consider two samples drawn from two distributions whose means differ only by a shift parameter. If we let $X_1, X_2, X_3, \cdots, X_m$ and $Z_1, Z_2, Z_3, \cdots, Z_n$ be two samples, then the parameter can be estimated by the difference in means, $\overline{X} - \overline{Z}$. Interest will be to generate the likelihood for the parameter θ. If $X_{m+i} = Z_i$, $i = 1, 2, \cdots, N$ and $N = m + n$, then the transformed data set would be

$$y_\theta = \left(X_1, X_2, X_3, \cdots, X_m,\ Z_1 - \theta,\ Z_2 - \theta, Z_3 - \theta, \cdots, Z_n - \theta\right) \tag{3.4}$$

Observe that (3.4) is simply $y_\theta = \left(X_1, X_2, X_3, \cdots, X_m,\ X_{m+1} - \theta,\ X_{m+2} - \theta,\ \cdots, X_N - \theta\right)$ which could be expressed as:

$$y_{y_\alpha} = \begin{cases} X_i,\ (i = 1, 2, \cdots, n) \\ X_i - \theta,\ (i = m+1,\ m+2,\ \cdots, N) \end{cases} \tag{3.5}$$

(3.5) is true since $X_{m+i} = Z_i$. Thus the estimate T^0_θ can be defined as:

$$T_\theta = \frac{1}{n}\sum_{i=m+1}^{N} Y_{\theta i} - \frac{1}{m}\sum_{i=1}^{m} Y_{\theta i} \tag{3.6}$$

And hence the permutation likelihood of θ is the density approximation

$$L^0(\theta) = L(\theta | x, Z) = \hat{f}_{T^0_\theta}(t_\theta) \tag{3.7}$$

where

$$T^0_\theta = \frac{1}{n}\sum_{i=m+1}^{N} Y_{\theta i} - \frac{1}{m}\sum_{i=1}^{m} Y_{\theta i} \tag{3.8}$$

With $y^0_{\theta i}$ as ith permutation from $(y_{\theta 1}, y_{\theta 2}, \cdots, y_{\theta N})$ and t_θ is the observed value of T_θ, i.e. $t_\theta = t - \theta$. By adopting the kernel density estimator, we have

$$L^0(\theta) = \hat{f}_{T^0_\theta} = \hat{f}(t) = \frac{1}{N^0 h}\sum_{i=1}^{N^0} k\left(\frac{t - T^0_i}{h}\right) \tag{3.9}$$

where $k(.)$ is the Guassian kernel and h is the approximate window width (of order 2) Then, the exact permutation likelihood is defined as the distribution:

$$L^0(\theta) = \frac{1}{N^0}\#\left\{T^0_j \middle| t - \varepsilon \le T^0_j \le t + \varepsilon\right\} \tag{3.10}$$

for some $\varepsilon > 0$.

3.2 Higher Order Forms:

Osemwenhkae [17] obtained the optimal values for any order m of the window width to be

$$h_{m,opt} \approx \left\{\frac{(m!)^2}{2m}\frac{\pi^{3/2}}{2^m \Gamma\left(\frac{2m+1}{2}\right)\left(\Gamma\left(\frac{m+1}{2}\right)\right)^2}\right\}^{\frac{1}{2m+1}} \delta n - \frac{1}{2m+1} \tag{3.11}$$

Thus, by substituting $h_{m,opt}$ in place of h in (3.9) we generate higher order forms of the permutation kernel likelihood of the form:

$$L^0_m(\theta) = \hat{f}_{T^0_\theta}(t) = \hat{f}(t) = \frac{1}{N^0 h_{m,opt}}\sum_{i=1}^{N^0} k\left(\frac{t - T^0_i}{h_{m,opt}}\right) \tag{3.12}$$

where m=2, 2, 6, 8, ...

4.0 Discussion

The exact permutation likelihood has been examined in this work. The beauty of the work is that the likelihoods are realized unconditionally. In this article, consideration is given to the special case of 2xn tables (i.e. the two sample experiments) with row and column totals allowed to vary with each permutation. Large sample approximations are commonly adopted in several nonparametric tests as alternatives to the tabulated exact tests. The requirement for such alternative approximations to be reliable is that the sample size should be sufficiently large. This requirement is not even necessary in the generation of the permutation likelihood since all the permutations are to be generated to produce the likelihood. All other alternatives such as the bootstrap likelihood, provide approximate likelihood to the permutation likelihood considered in this work. This work has sufficiently provided all the theoretical background needed for the generation of empirical permutation likelihood.

References

[1] Efron, B. (1979a) Bootstrap Methods: Another Look at the Jacknife.

[2] Efron, B. (1979b): Computers and the Theory of Statistics: Thinking the Unthinkable SIAM Rev. 21, 460-480.

[3] Efron B, (1982): The Jacknife, the Bootstrap and Other Resampling plan. SIAM Regional Conference Series in Applied Mathematics.

[4] Ogbonmwan, S.M., and Wynn, H. P. (1987): Resampling Generated Likelihood, Statistical Decision Theory and Related Topics IV; Purdue University, Vol.1 pp. 133-147.

[5] Ogbonmwan, S. M. (1992): Simulation Codes and Likelihood Generation in Scientific Computing. Edited by Fatunla. S. O. pp. 144-152.

[6] Ogbonmwan, S. M. (2003), "Empirical Bootstrap Likelihood for Contrasts of Parameters". Journal of the Nigerian Association Mathematical Physics., Vol. 7, pp. 97 – 104.

[7] Owen, A. B. (1988): Empirical Likelihood Ratio Confidence Intervals for a Single Functional, Biometrika, 75, 237-249.

[8] Owen, A. B. (1990): Empirical Likelihood Ratio Confidence Regions. Ann. Statist. 18, 90-120.

[9] Chen, S. X. (1993). On the coverage accuracy of empirical likelihood confidence interval regions for linear regression model. Ann. Inst. Statist. Maths. 45, 621-637.

[10] Chen, S. X. (1994). Empirical likelihood confidence intervals for linear regression coefficients. J. Mult. Anal. 49, 24-40.

[11] Chen, S. X. (1996). Empirical likelihood confidence intervals for nonparametric density estimation. Biometrika, 83(2), 329-341.

[12] Holmes, C. C. and Adams N. M. (2003): Likelihood inference in nearest-neighbour classification models. Biometrika 90, 99 – 102.

[13] Agresti, A. (1992): A Survey of Exact Inference for Contigency Tables, Statistical Science, 7,131-177.

[14] Good, P. (2000): Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypothesis (2nd Edition). N.Y Springer Verlag.

[15] Odiase, J. I., and Ogbonmwan, S. M. (2005): An Algorithm for GeneratingUnconditional Exact Permutation Distribution for a Two-Sample Experiment, Journal of Modern Applied Statistical Methods. 4(1), 319-332.

[16] Opdyke, J. D. (2003). Fast Permutation Tests that Maximize Power under Conventional Monte Carlo Sampling for Pairwise and Multiple Comparisons. Journal of Modern Applied Statistical Methods, 2, 27-49.

[17] Osemwenkhae, J. E. (2003): Higher Order Forms in Kernel Density Estimation, Ph.D. Thesis, Department of Mathematics, University of Benin, Benin City, Nigeria.