5-1-2003

JMASM6: An Algorithm For Generating Exact Critical Values For the Kruskal-Wallis One-Way ANOVA

Todd C. Headrick
Southern Illinois University at Carbondale, headrick@siu.edu

JMASM Algorithms and Code
JMASM6: An Algorithm For Generating Exact Critical Values For the Kruskal-Wallis One-Way ANOVA

Todd C. Headrick
Southern Illinois University, Carbondale

A Fortran 77 subroutine is provided for computing exact critical values for the Kruskal-Wallis test on k independent groups with equal or unequal samples sizes. The subroutine requires the user to provide sorting and ranking routines and a uniform pseudo-random number generator. The program is available from the author on request.

Key words: Kruskal-Wallis test, nonparametric statistics, one-way ANOVA

Introduction

Kruskal and Wallis (1952) derived a rank-based nonparametric test as an alternative to the one-way analysis of variance (ANOVA) on k independent groups. It has been demonstrated that the Kruskal-Wallis (K-W) test can have considerable power advantages over the ANOVA F test when the assumption of normality is violated (e.g., Aman & Headrick, 2003).

The null distribution of the K-W statistic is derived under the assumption that all N observations are from the same population. Because the number of ways N ranks may be divided into groups of $n_1,\ldots,n_k$ grows rapidly, most statistics textbooks (e.g., Conover, 1999; Gibbons, 1992; Siegel & Castellan, 1989) limit the reporting of exact critical values for the K-W statistic to no more than $k=3$ groups and with $n_j \leq 5$ observations per group.

Todd C. Headrick is Assistant Professor of Statistics, 222-J Wham Building, Mail Code 4618, Southern Illinois University-Carbondale, IL, 62901. His areas of research interest are statistical computing, nonparametric statistics, and optimization. Email him at headrick@siu.edu.

The asymptotic null distribution of the K-W statistic is chi-squared with $k-1$ degrees of freedom. As such, for $k>3$ and $n_j>5$ observations per group, the K-W asymptotic null distribution is recommended as the reference for making the decision to reject or fail to reject the null hypothesis that all k population distribution functions are identical (Conover, 1999, p. 289).

Most commonly used statistical software packages (e.g., Minitab; SPSS) that compute the K-W statistic only provide the asymptotic p-value. This may present a problem to an applied researcher because this p-value can be conservative relative to the exact p-value when both k are n_j are small. For example, for $k = 5$ and $n_j = 5$ for all $j = 1,\ldots,5$, the chi-squared critical value associated with $a = .05$ is 9.4876 whereas the exact critical value is 8.8985. Thus, using the asymptotic critical value for this design has the effect of lowering the Type I error rate from .050 to approximately .0363.

In view of the above, the purpose of this paper is to provide a subroutine that computes the exact critical values for the K-W statistic. The subroutine will compute critical values for any number of k populations with equal or unequal sample sizes.

Methodology

The subroutine initially generates N uniform pseudo-random numbers on the interval $(0,1)$. It is assumed that the probability of obtaining any tied scores is zero. The uniform deviates are then ranked to form a permutation of the numbers from $1,\ldots, N$. The algorithm then sequentially splits the permutation of ranks into k groups in accordance to the user's specified sample sizes of $n_1,\ldots,n_k$. The K-W statistic is then computed as

$$H = \left(\frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} \right) - 3(N+1),$$

where R_j is the sum of the ranks in the jth group. This process is repeated until a sufficient number of H statistics are generated to adequately model the null distribution.

The algorithm then selects the critical values associated with the alpha levels of .01, .05, and .10. In general, the critical values returned by the subroutine are associated with a range of p-values. For example, for $k=3$ and $n_j=5$ for all $j=1,2,3$, the subroutine will return the exact critical value of 5.659997 for $a=.05$. However, this critical value is associated with p-values ranging from approximately .0483 to .0537. As such, the program informs the user that the null hypothesis may be rejected if the computed K-W statistic is *strictly greater* than the critical value of 5.659997.

The method used in the subroutine for selecting critical values yields the same values reported in Conover (1999, Table A8, p. 539). It should be noted that this method is different from the method that was used for selecting the critical values reported in Gibbons (1992, Table K, p. 503) and Siegel & Castellan (1989, Table O, p. 356). Specifically, for the example above, these texts report a critical value of 5.78. This method indicates to the reader that the null hypothesis may be rejected if the computed K-W statistic is *greater than or equal to* the critical value of 5.78.

Conclusion

The program leaves it to the user to specify the number of K-W statistics to generate. The larger the value of N requires a larger number of K-W statistics to be generated to adequately model the null distribution. Thus, it is recommended that trials be repeated with an increasing number (e.g., 100,000; 500,000; 1,000,000, etc.) of K-W statistics generated in each trial run. It is suggested that the user terminate this process when changes in the critical values are less than 10^{-4}.

References

Aman, S. Y., & Headrick, T. C. (2003). An empirical investigation of Type I error and power between data transformations to normality and nonparametric analysis in the context of one-way ANOVA. Annual meeting of the American Educational Research Association, SIG Educational Statisticians, Chicago, IL.

Conover, W. J. (1999). *Practical nonparametric statisics* (3rd ed.). New York: Wiley.

Gibbons, J. D. (1992). *Nonparametric statistical inference* (3rd ed.). New York: Marcel Dekker.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion analysis of variance. *Journal of the American Statistical Association*, 47, 583-621; errata, ibid., 48, 907-911.

Minitab (2000). *Minitab for Windows*, *release* 13.3, Minitab, Inc., State College, PA.

Siegel, S., & Castellan, N. J. (1989). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

SPSS (2002). *SPSS for Windows*, *version* 11.0, SPSS, Inc., Chicago, IL.

```
      SUBROUTINE KW(K, N, M, ISAMP, CRIT)
C***********************************************************************
C K is the specified number of populations in the one-way ANOVA.
C N is the specified total sample size.
C M is an array with K specified sample sizes n_1,...,n_k .
C ISAMP is the specified number of K-W
C sample statistics to be generated. CRIT is an array of critical
C values for the K-W test to be returned.
C***********************************************************************
      REAL X(N), RKX(N), SMRKX(K), STAT(ISAMP), CRIT(3)
      INTEGER M(K)
      DOUBLE PRECISION DSEED
C***********************************************************************
C Generate the specified number of K-W sample statistics.
C***********************************************************************
      DO 10 I = 1, ISAMP
C***********************************************************************
C Call the uniform random number generator routine. Generate an array X
C with N uniform deviates. Call the ranking routine. Rank the N uniform
C deviates in X and place these ranks into the array RKX.
C***********************************************************************
      CALL UNIFORM (DSEED, N, X)
      CALL RANK (X, N, RKX)
C***********************************************************************
C Compute the K-W statistic denoted below as H.
C***********************************************************************
      IE = 0
      SUM1 = 0
      DO 20 J = 1, K
      IB = IE + 1
      IE = IE + M(J)
      SUM2 = 0
      DO 30 L = IB, IE
      SUM2 = SUM2 + RKX(L)
      SMRKX(J) = SUM2**2 / FLOAT(M(J))
 30 CONTINUE
      SUM1 = SUM1 + SMRKX(J)
 20 CONTINUE
      H = (12 / (FLOAT(N)*(FLOAT(N) + 1)))*SUM1 − 3*(FLOAT(N) + 1)
C***********************************************************************
C Store the H statistic in the array STAT.
C***********************************************************************
      STAT(I) = H
   10 CONTINUE
C***********************************************************************
C Sort the array STAT.
C***********************************************************************
      CALL SORT (STAT, ISAMP)
```

```
C************************************************************************
C Obtain the critical values associated with the alpha levels
C .01, .05, and .10.
C************************************************************************
      DO 40 I = ISAMP, 1, —1
      PR = (FLOAT(ISAMP) — FLOAT(I)) / FLOAT(ISAMP)
            IF ( PR .LE. 0.01) THEN
      CRIT(1) = STAT(I)
      ELSEIF (PR .LE. 0.05) THEN
      CRIT(2) = STAT(I)
      ELSEIF (PR .LE. 0.10) THEN
      CRIT(3) = STAT(I)
      ELSEIF (PR .GT. 0.10) THEN
      GOTO 50
      END IF
 40 CONTINUE
 50 CONTINUE
      RETURN
      END
```