Journal of Modern Applied Statistical Methods

Volume 1 | Issue 2

Article 34

11-1-2002

Extensions Of The Concept Of Exchangeability And Their Applications

Phillip I. Good
Information Research, Huntington Beach, CA

Journal of Modern Applied Statistical Methods
Fall 2002, Vol. 1, No 2, 243-247

1538 – 9472/02/$30.00

Extensions Of The Concept Of Exchangeability And Their Applications

Phillip Good
Information Research
Huntington Beach, California

Permutation tests provide exact p-values in a wide variety of practical testing situations. But permutation tests rely on the assumption of *exchangeability*, that is, under the hypothesis, the joint distribution of the observations is invariant under permutations of the subscripts. Observations are *exchangeable* if they are independent, identically distributed (i.i.d.), or if they are jointly normal with identical covariances. The range of applications of these exact, powerful, distribution-free tests can be enlarged through exchangeability-preserving transforms, asymptotic exchangeability, partial exchangeability, and weak exchangeability. Original exact tests for comparing the slopes of two regression lines and for the analysis of two-factor experimental designs are presented.

Key words: Permutation test, exchangeable, weak exchangeability, exact test, groups.

Introduction

Because the permutation tests can provide exact significance levels and are powerful and distribution free, they have an enormous number of applications. See, for example, Manly(1997). The observations on which these tests are based may be drawn from finite populations or represent a particular realization of a set of random variables. Rank tests are permutation tests based on the ranks of the observations rather than their original values.

Permutation tests rely on the assumption of *exchangeability*, that is, under the hypothesis, the joint distribution of the observations is invariant under permutations of the subscripts. Observations are *exchangeable* if they are independent, identically distributed (i.i.d.), or if they are jointly normal with identical covariances. For additional examples, see Galambos (1986) or Draper et al. (1993).

A caveat is that a set of units may be exchangeable for some purposes and not for others, depending on what is measured and the questions of interest. A simple example suggested by Draper et al (1993) is a circadian series in which observations within days are not exchangeable because of serial correlation, while observations between days (at the same point in time) are exchangeable as are the residuals from a model incorporating serial correlation.

The range of applications of these exact, powerful, distribution-free tests are enlarged below through exchangeability - preserving transforms, asymptotic exchangeability, partial exchangeability, and weak exchangeability. Original exact tests for comparing the slopes of two regression lines and for the analysis of two-factor experimental designs are presented.

Exchangeable Variables

Let $G\{x;\ y_1, y_2,\ \ldots y_{n-1}\}$ be a distribution function in x and symmetric in its remaining

Phillip I. Good is the author of five textbooks in statistics including *Permutation Tests, Resampling Methods, Applying Statistics in the Courtroom,* Common Errors in Statistics, *and Managers Guide to Design and Conduct of Clinical Trials.* He has published a number of short stories. See links at: http://users.oco.net/authors.htm including http://www.beachesbeaches.com/pinkie.html.

arguments—that is, permuting the remaining arguments would not affect the value of G. Let the conditional distribution function of x_i given $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ be G for all i. Then the $\{x_i\}$ are exchangeable.

It is easy to see that a set of i.i.d. variables is exchangeable. Or that the joint distribution of a set of normally distributed random variables whose covariance matrix is such that all diagonal elements have the same value σ^2 and all the off-diagonal elements have the same value ρ^2 is invariant under permutations of the variable subscripts.

Polya's urn or contagion model variables are also exchangeable. An urn contains **b** black balls, **r** red balls, **y** yellow balls, … and so forth. A series of balls is extracted from the urn. After the ith extraction, the color of the ball X_i is noted and *k* balls of the same color are added to the urn., where k can be any integer, positive, negative, or zero. The set of random events $\{X_i\}$ form an exchangeable sequence. See, also, Dubins and Freedman (1979).

Transformably Exchangeable

Suggesting the concept of transformably exchangeable is the procedure for testing a non-null two-sample hypothesis H: F[x] =G[x−d]; for if there are two sets of independent observations $\{Z_i\}$ and $\{Y_i\}$ with Z_i distributed as F and Y_i as G, an exact test of H can be obtained by first transforming the variables by subtracting 0 from each of the Z_i's and d from each of the Y_i's.

A set of observations (random variables) **X** will be said to be *transformably exchangeable* if there exists a transformation (measureable transformation) T, such that T**X** is exchangeable (Commenges, 2001).

If there are a set of observations {X[t], t= 1, 2,…n} where X[t] = a + bX[t−1] + z_t and the $\{z_t\}$ are i.i.d., then the variables {Y[t], t= 2,…n} where Y[t] = X[t] − bX[t−1] are exchangeable.

Dependent non-collinear normally distributed variables with the same mean are transformably exchangeable for as the covariance matrix is non-singular, use the inverse of this matrix may be used to transform the original variables to independent (and hence exchangeable) normal ones. By applying two successive transformations, an exact permutation test can be obtained of the non-null two-sample univariate hypothesis for dependent normally distributed variables providing the covariance matrix is known. Unfortunately, as Commenges (2001) showed, the decision to accept or reject in a specific case may depend on the transformation that was chosen.

Michael Chernick notes the preceding result applies even if the variables are collinear. Let R denote the rank of the covariance matrix in the singular case. Then, there exists a projection onto an R-dimensional subspace where R normal random variables are independent. So if there is an N dimensional (N > R) correlated and singular multivariate normal distribution, there exists a set of R linear combinations of the original N variables so that the R linear combinations are each univariate normal and independent of one other.

Exchangeability-Preserving Transforms

Suppose it is desired to test whether two regression curves are parallel, even though the value of the intercepts are not known. Given that

$$y_{ik} = a_i + b_i x_{ik} + \varepsilon_{ik} \textit{ for } i = 1,2;\ \mathrm{k} = 1,\ldots,n_i$$

where the errors $\{\varepsilon_{ij}\}$ are exchangeable. To obtain an exact permutation test for H: $b_1 = b_2$, the $\{a_i\}$ are needed to be eliminated, while preserving the exchangeability of the residuals. It is known that under the null hypothesis

$$\bar{y}_{i.} = a_i + b\bar{x}_{i.} + \bar{\varepsilon}_{i.}$$

$$y' = \frac{1}{2}(\bar{y}_1 - \bar{y}_2);\ x' = \frac{1}{2}(\bar{x}_1 - \bar{x}_2);\ \varepsilon' = \frac{1}{2}(\bar{\varepsilon}_1 - \bar{\varepsilon}_2);\ a' = \frac{1}{2}(a_1 + a_2).$$

Define
$$y'_{1k} = y_{1k} - y' \ \text{ for } \mathrm{k} = 1 \textit{ to } n_1, \text{and}$$
$$y'_{2k} = y_{2k} + y' \ \text{ for } \mathrm{k} = 1 \textit{ to } n_2.$$

Define

$$x'_{1k} = x_{1k} - x' \text{ for k} = 1 \textit{ to } n_1 \text{ and } \mathrm{x}'_{2k} = x_{2k} + x' \text{ for k} = 1 \textit{ to } n_2.$$

Then

$$y'_{ik} = a' + bx'_{ik} + \varepsilon'_{ik} \textit{ for } i = 1,2;\ \mathrm{k} = 1,\ldots,n_i$$

Two cases arise. If the original predictors were the same for both sets of observations, that is, if $x_{1k}= x_{2k}$ for all k, then the errors $\{\varepsilon'_{i\ k}\}$ are exchangeable and the method of matched pairs can be applied; see, for example, Good (2000, p51). Otherwise, proceed as follows: First, estimate the two parameters a' and b by least-squares means. Use these estimates to derive the transformed observations $\{y'_{ik}\}$. Then test the hypothesis that $b_1=b_2$ using a two-sample comparison. If the original errors were exchangeable, then the errors $\{\varepsilon'_{ik}\}$ though not independent are exchangeable also and this test is exact.

Now suppose

$$y_{ik} = A_i Z_k + b_i x_{ik} + \varepsilon_{ik} \ for\ i = 1,2;\ \mathrm{k} = 1,\ldots,n_i$$

where Z_k is a column vector of covariates with A_i a row vector of the corresponding coefficients. Defining A'_i as the mean of A_1 and A_2, then

$$y'_{ik} = A' Z_k + bx'_{ik} + \varepsilon'_{ik} \ for\ i = 1,2;\ \mathrm{k} = 1,\ldots,n_i$$

which are analogous results for the general case.

Dean and Verducci (1990) characterized the linear transformations that preserve exchangeability. Commenges (2001) characterized the linear transformations that also preserve the permutation distribution. Clearly any transformation which preserves the ordering of the order statistics preserves exchangeability.

Asymptotic Exchangeability

Illustrating the concept of asymptotic exchangeability are the residuals in a two-way complete balanced experimental design. Our model is that

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

where

$$\sum \alpha_i = \sum \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$$

and the $\{\varepsilon_{ijk}\}$ are exchangeable. Eliminating the main effects in the traditional manner, that is, setting

$$X'_{ijk} = X_{ijk} - \overline{X}_{i..} - \overline{X}_{.j.} + \overline{X}_{...},$$

the test statistic obtained is

$$\mathrm{I}=\sum_i \sum_j \left(\sum_k X'_{ijk}\right)^2,$$

which was first derived by Still and White (1981). A permutation test based on this statistic will not be exact for finite samples as the residuals

$$\varepsilon'_{ijk} = \varepsilon_{ijk} - \overline{\varepsilon}_{i..} - \overline{\varepsilon}_{.j.} + \overline{\varepsilon}_{...}$$

are weakly correlated, the correlation depending on the subscripts. It is easy to show the Studentized correlations converge to a common value as the sample size increases, thus the residuals are asymptotically exchangeable, and the permutation test of the hypothesis $\gamma_{ij} = 0$ for all i and j based on I is asymptotically exact.

Romano (1990) proved asymptotic exchangeability for the two-sample comparison of independent observations with not necessarily identical distributions providing the underlying variables have the same mean and variance under the hypothesis. Baker (1995) used simulations to demonstrate the asymptotic exchangeability of the deviates about the sample median that are used in Good's test for equal variances.

Exchangeability and Invariance

The requirement for exchangeability in testing arises in either of two ways:

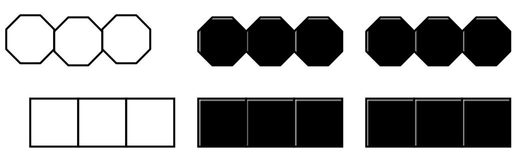
- Sufficiency—the order statistics are sufficient for a wide variety of problems.
- Invariance—the joint distribution of the observations is invariant under permutation of the subscripts.

For many testing problems, the underlying model must remain invariant under permutations of the subscripts. This can only be accomplished in many cases if the set of permutations are restricted. Recall that in the classic definition (de Finetti, 1930; Galambos, 1986) a set of n random variables is said to be *exchangeable* if the joint distribution of the variables is invariant with respect to the group S_n of all possible permutations of the subscripts.

Define the *weak exchangeability* of a set of random variables as the invariance of their joint distribution with respect to a subset of permutations. Clearly, a set of variables that is exchangeable is also weakly exchangeable.

Exchangeability is a necessary and sufficient condition for exactness in the classic testing problems to which permutation methods have been applied such as the 2- and k-sample tests. But in the two-factor experimental design considered in the previous section, only the error terms $\{\varepsilon_{ijk}\}$ are exchangeable; the $\{X_{ijk}\}$ are not. Nonetheless, because the $\{X_{ijk}\}$ are weakly exchangeable under any of the three null hypotheses (H_1: $\alpha_i = 0$ for all i, H_2: $\beta_j = 0$ for all j, and H_3: $\gamma_{ij} = 0$ for all i and j), Pesarin (2001) and Salmaso (2001) were able to derive independent exact tests for each of the main effects and the interactions.

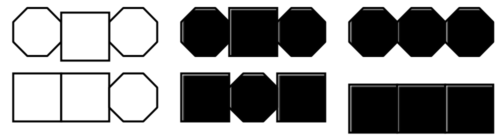
To see this, consider that the set of observations $\{X_{ijk}\}$ may be thought of in terms of a rectangular lattice L with K colored, shaped balls at each vertex. All the balls in the same column have the same color initially, a color which is distinct from the color of the balls in any other column. All the balls in the same row have the same pattern initially, a shape which is distinct from the shape of the balls in any other row.



A 2x3 design with three observations per cell.

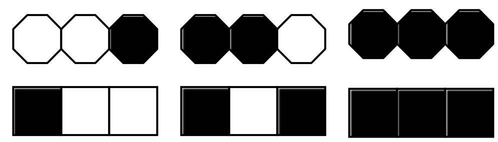
Let P denote the set of transformations that preserve the number of balls at each row and column of the lattice. P is a group**.**

Let P_R denote the set of exchanges of balls among rows which a) preserve the number of balls at each row and column of the lattice, and b) result in the numbers of each shape within each row being the same in each column. P_R is the basis of a subgroup of P.



A 2x3 design with three observations per cell after $\pi \varepsilon P_R$.

Let P_C denote the set of exchanges of balls among columns which a) preserve the number of balls at each row and column of the lattice, and b) result in the numbers of each color within each column being the same in each row. P_C is the basis of a subgroup of P.



A 2x3 design with three observations per cell after $\pi \varepsilon P_C$.

Let P_{RC} denote the set of exchanges of balls which preserve the number of balls at each row and column of the lattice, and result in a) an exchange of balls between both rows and columns (or no exchange at all), b) the numbers of each color within each column being the same in each row, c) the numbers of each shape within each row being the same in each column. P_{RC} is the basis of a subgroup of P. Moreover, $P_{RC} \cap P_R = P_{RC} \cap P_C = P_R \cap P_C = I$ and P is the group generated by the union of P_R , P_C and P_{RC}.

Define $p[\Delta; X] = \Pi_\iota \Pi_j \Pi_\kappa f[x - \Delta_{ij}]$ where

$$\Delta_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij},$$
$$\sum \alpha_i = \sum \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$$

and f is a density function that is continuous a.e.

Without loss of generality, it may be assumed μ=0, or, equivalently, the set of observations $\{X'_{ijk}\}$ obtained by subtracting μ from each element of $\{X_{ijk}\}$ may be used. Suppose, now, the hypothesis H_1: $\alpha_i = 0$ for all i holds. Then the joint distribution of the vector

($x_{i1k'}$, $x_{i2k''}$,…, x_{ijk*}) obtained by taking an arbitrary element from each column of the ith row is identical with the joint distribution of

$$(z-\beta_1-\gamma_{i1}, z-\beta_2-\gamma_{i2}, \ldots, z-\beta_J-\gamma_{iJ})$$

where f is the probability density of z. The probability density of the sum of these latter elements is identical with the probability density of nz $-\sum_{j=1}^{J}\beta_j - \sum_{j=1}^{J}\gamma_{ij}$ = nz; that is, f(z/n).

Under H_1

- f is the probability density of the mean of each of the rows of X.
- Applying any of the elements of $\mathbf{P_R}$ leaves this density unchanged.
- Applying any of the elements of $\mathbf{P_R}$ leaves the density of the test statistic $F_2 = \Sigma_i(\Sigma_j\Sigma_k x_{ijk})^2$ unchanged.

Similarly, to test H_2, the permutation distribution over $\mathbf{P_C}$ of any of the statistics $F_2 = \Sigma_j(\Sigma_i\Sigma_k x_{ijk})^2$, $F_1 = \Sigma_j |\Sigma_i\Sigma_k x_{ijk}|$, or $R_2 = \Sigma_j g[j]\Sigma_i\Sigma_k x_{ijk}$, where g[j] is a monotone function of j may be used.

If q ε $\mathbf{P_R}$ and s ε $\mathbf{P_C}$, then under H_3, the density of $S_{ij} = \Sigma_K x_{ijk}$ is invariant with respect to p = qt ε $\mathbf{P_{RC}}$, and, by induction, applying any of the elements of $\mathbf{P_{RC}}$ leaves the density of the test statistic $S = \Sigma_i\Sigma_j(S_{ij})^2$ unchanged. As only the identity I is common to the corresponding permutation groups, the permutation tests of the three hypotheses are independent of one another.

Partial and Weak Exchangeability

Consider a sequence of discrete random variables that represent the outcomes of a finite Markov Chain whose transition matrix is such that $p_{ij} = p_{ji}$ for all i and j. Such a sequence is said to be *partially exchangeable* (see, for example, Zaman, 1984). If the transition matrix is connected then the sequence is also weakly exchangeable.

References

Baker, R. D. (1995). Two permutation tests of equality of variance. *Statistics of Computation*, *5*, 289-296.

Commenges, D. (2001, in press). Transformations which preserve exchangeability and applications to permutation tests. *Nonparametric statistics*.

Dean, A. M. & Verducci, J. S. (1990). Linear transformations that preserve majorization, Schur concavity, and exchangeability. *Linear algebra and its applications*, *127*, 121-138.

Dean, A. M. & Wolfe, D. A. (1990). A note on exchangeability of random variables. *Statistica Neerlandica*, *44*, 23-27.

Draper, D., Hodges, J. S., Mallows, C.L., & Pregibon, D. (1993). Exchangeability and data analysis (with discussion). *Journal of the Royal Statistical Society*, A. *156*, 9-28.

Dubins, L. & Freedman, D. (1979). Exchangeable processes need not be mixtures of independent identically distributed random variables. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebeit*, *48*, 115-132.

Galambos, J. (1986). Exchangeability. In S. Kotz and N. L. Johnson (Eds). *Encyclopedia of statistical sciences*, *7*, 573-577. NY: Wiley.

Good, P. I. (2000) *Permutation tests*. (2nd ed.). NY: Springer.

Lehmann, E. L. (1986). *Testing statistical hypotheses*. NY: John Wiley & Sons.

Manly, B. F. J. (1997). *Randomization, bootstrap and Monte Carlo methods in biology*. (2nd ed.). London: Chapman and Hall.

Pesarin, F. (2001). *Multivariate permutation tests*. NY: Wiley.

Romano, J. P. (1990). On the behavior of randomization tests without a group invariance assumption. *Journal of the American Statistical Association*, *85*, 686-692.

Salmaso, L. (2003, in press). Synchronized permutation tests in 2k factorial designs. *Communications in Statistics - Theory and Methods.*

Still, A. W. & White, A. P. (1981). The approximate randomization test as an alternative to the F-test in the analysis of variance. *British. Journal of Mathematical and Statistical Psychology*, *3*, 243--252.

Zaman, A. (1984). Urn models for Markov exchangeability. *Annals of Probability*, *12*, 223-229.