

## Programs for Wei-Lachin Multivariate Rank Analyses

John M. Lachin  
jml@biostat.bsc.gwu.edu

Copyright (c) 2006 by John M. Lachin

The following files are included in this package:

**WeiLachinComputation.pdf** provides a technical description of the methods and the applications.

**weilach.sas** is the SAS macro for the Wei and Lachin (1984) analysis based on the fortran program published by Makuch, Merrill and Escobar (1991). The macro provides additional tests and computations described by Thall and Lachin (1988) and Lachin (1992).

Before calling the macro, a global macro variable must be defined using the statement

```
%let nr = (the number of repeated measures);
```

The Macro is called with the following arguments.

```
%macro weilach(method, option, dset, grpvar, grp1c,  
                grp2c, pref_d, pref_c);
```

The arguments are

method:

- 1 requests wilcoxon tests only
- 2 requests log-rank tests only

option:

- 0 requests no additional analyses
- 1 requests thall-lachin analysis  
(option 1 only valid if method = 1)

dset: name of the dataset

grpvar: name of the group (treatment) variable

grp1c: single digit numeric values used to identify the groups

grp2c:

pref\_d: prefix for variables containing the data values, (e.g., x1, x2, x3 ...)

pref\_c: prefix for indicator variables with censoring values. (e.g., i1, i2, i3 ...)

Depending on the specifications, the macro generates up to four data sets: wlstats, wlsigma; tlstats; and thetsig.

**wltest.sas** is a test program that reproduces results presented in Lachin (1992) with the analysis of the repeated measures cholesterol data from the National Cooperative Gallstone Study (NCGS). The NCGS data is contained in the "allchol" data set. In order to run the wltest program, the libname and %include statements in the program must be changed to provide the path to the data set and SAS macro, respectively.

The macro is invoked with the following statements.

```
%let nr = 4;
%WEILACH(1, 1, w1, tretgrp, 3, 5, y, i);
```

In the NCGS data, the two treatment groups being compared are the high dose (3) and placebo (5).

**allchol.sas7bdat** is a PCSAS, version 7 copy of the data set. **allchol.xpr** is a SAS export data set that can be “imported” to use on other platforms, i.e. SAS systems on other types of computing environments that do not support PCSAS.

**Weilach.lis.pdf** presents a listing file with the output from the wlttest program for the analysis of the cholesterol data. The following is a description of the output. To save space, each “page” of output is delimited by a row of dashes (----- etc.).

Page 1 presents the original Wei-Lachin (1984) computations. Observations 1-4 give the sample sizes, the Wei-Lachin Wilcoxon test, variance “SIG”, the chi-square statistic and p-value for the 4 repeated measures. Line 5 presents the 1 df test of stochastic ordering and line 6 presents the K=4 df omnibus test based on the repeated Wilcoxon rank statistics. These tests use the rank test values.

Page 2 presents the variance/covariance matrix of the test statistics for the 4 repeated times.

Based on Thall and Lachin (1988), Lachin (1992) describes additional statistics based on the Mann-Whitney proversion parameter estimates computed as  $P(X1 > X2) - P(X2 > X1)$ , or  $PR1 - PR2$ , where  $X1$  and  $X2$  are two random observations from groups 1 and 2, respectively. Most of the tests computed are described in Lachin (1992). These computations are presented on page 3. This data set contains 10 observations. First the variables “NCTR” through “LABEL” are printed for all 10 observations, followed by additional variables for the 10 observations. The proversion probabilities are the variables  $PR1$  and  $PR2$ , and the Mann-Whitney values are the variable  $THETA$ , followed by the variance (SIG) and SE of the theta.

Lines 1-4 repeat the Wei-Lachin analysis of the rank scores given on page 1, but with the Mann-Whitney parameters added. Lines 5-6 present the same test statistics as shown on page 1. Lines 7-10 present additional overall estimates of the Mann-Whitney location shift and the corresponding test values.

Line 7 presents an N-weighted tests of stochastic ordering using a weighted combination of the individual Mann-Whitney parameters, weighted by the sample size at each time. This test has not been published or explored in the literature. But it is a special case of the test based on (13) in Lachin (1992). We have used this test for analyses of repeated measures when the sample sizes decline markedly over time due to staggered entry.

Line 8 presents an unweighted test of stochastic ordering that is based on the unweighted mean of the Mann-Whitney shift parameters as shown in eqn (8) of Lachin (1992) whereas the original test of stochastic ordering was based on the unweighted mean of the

Wilcoxon rank statistics as in eqn (7). The two in general differ slightly. Frick (1995) shows that this is an optimal robust test (see also Lachin, 2000).

Line 9 presents the test of association based on an optimal estimate of an assumed common theta, as presented in equation (16) of Lachin (1992). This test is equivalent to an overall test of the "group" effect in a repeated measures analysis as in PROC GLM, MIXED or GLM. All are based on the average group difference over time as an estimate of an assumed common difference over time.

All of these tests control the type I error probability at the desired level under the null hypothesis. The difference is that each is more efficient (powerful) under specific alternatives. In addition to Lachin (1992), also see the discussion of the various tests in Sections 4.5 - 4.9 of Lachin (2000).

Finally, line 10 presents the test of homogeneity of the thetas over the repeated times, equation (9) of Lachin (1992). This is equivalent to the test of no group by time interaction in the theta values over time.

## References:

Frick, H. (1995). Comparing trials with multiple outcomes: The multivariate one-sided hypothesis with unknown covariances. *Biom. J.*, 8, 909-917.

Lachin JM. Some large sample distribution-free estimators and tests for multivariate partially incomplete observations from two populations. *Statistics in Medicine*, 11, 1151-1170, 1992.

Lachin, J.M. *Biostatistical Methods: The Assessment of Relative Risks*. John Wiley and Sons, 2000. (ISBN: 0-471-36996-9; 541 pages)

Makuch, R. W., Escobar, M. and Merrill, S. 'A two-sample test for incomplete multivariate data', *Applied Statistics*, 40, 202-212 (1991).

Thall PF and Lachin JM. Analysis of recurrent events: Nonparametric methods for random interval count data. *Journal of the American Statistical Association*, 1988, 83, 339-347.

Wei LJ and Lachin JM. Two-sample asymptotically distribution-free tests for incomplete multivariate observations. *Journal of the American Statistical Association*, 1984, 79, 653-661.