A Note on Comparing Repeated Observations

C. Mitchell Dayton
University of Maryland

It is often of interest to compare means based on the same cohort of individuals at several points in time. For example, Ramsey and Schafer (1997) provide mental aptitude test scores (IQ scores) for 62 children at ages 2, 4, 8 and 13 years (Table 1). Conventional multivariate tests for equality of means of repeated observations (e.g., Wilk's lambda based on the three differences among successive means) suggest that the means are not all equal across the ages tested. The down-up-down pattern of the means suggests a complex trend and polynomial trend components up through cubic are each statistically significant at the .05 level. To further study the differences among the means, pairwise Hotelling T^2 statistics evaluated at the .05 level were conducted. The results suggest that IQ at age 13 is significantly lower than at any of the three previous ages and that mean scores for ages 2 and 4 differ. However, the remaining pairwise differences for ages 2 and 8 and ages 4 and 8 do not differ. Thus, as often occurs with pairwise comparison procedures, we are presented with an intransitive decision with respect to interpreting results for the three youngest ages.

In a previous paper (Dayton, 1998), I advocated the use of information criterion such as Akaike's (1973, 1974) AIC as a substitute for conventional pairwise-comparison procedures such as the Tukey procedure. Among the problems cited for conventional procedures were: (1) some arbitrary technique is utilized to control the family-wise type I error rate for the set of correlated pairwise tests; (2) issues of homogeneity of variance and differential sample size pose problems for many paired-comparison procedures; (3) intransitive decisions) are the rule rather than the since these procedures entail a series of discrete, pairwise significance tests; and (4) there exists a large variety of competing procedures that differ in how type I error is controlled and, consequently, in power.

The AIC statistic is of the form AIC = $-2Log_e(L) + 2q$ where L is the likelihood of the sample given a specified theoretical distribution and q is the number of independent parameters that is estimated by maximum likelihood when fitting the theoretical distribution to the data. A more conservative information criterion, since it generally incorporates a heavier penalty term for additional parameters, is the Schwarz (1978) statistic of the form BIC = $-2Log_e(L) + Log_e(N)q$ where N is the sample size. In summary, the procedure recommended by Dayton (1998) is as follows. For means based on P independent groups, there is a total of 2^P-1 patterns of ordered subsets with equal means within subsets. For example, with $P = 4$ groups with the means ranked and labeled 1, 2, 3, 4, the $2^3 = 8$ distinct ordered subsets are {1,2,3,4}, {12,3,4}, {1,23,4}, {1,2,34}, {12,34}, {1,234}, {123,4} and {1234} where a comma is used to separate subsets that are unequal in mean value. For each ordered subset of means, compute maximum likelihood estimates (MLE's) for the unique parameters and calculate the corresponding information criterion, AIC_m. Finally, select the pattern with min(AIC_m) as the preferred outcome. Simulation results (Dayton, 1998; Huang & Dayton, 1995) suggest that the null pattern (e.g., {1234}) does not tend to be detected at a satisfactory rate when, in fact, all population means are equal. Thus, a two-stage "protected" strategy is recommended in which an omnibus ANOVA, or MANOVA as appropriate, is used to screen for the null case and then, assuming rejection of the null hypothesis, a min(AIC_m) strategy is employed.

For repeated observations, the scores are correlated across, for example, P time periods with the result that the data are inherently multivariate. Assuming a P-variate multivariate normal density for the distribution of the scores, the log-likelihood for a sample of N P-tuples of observations is given by:

$$Log_e(L) = -.5PNLog_e(\pi) - .5NLog_e(|\Sigma|) - .5trace(\Sigma^{-1}\Theta) - .5N(\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu)$$

where $\bar{x}$ is the Px1 vector of sample means, μ is the Px1 vector of population means, Σ is the PxP variance-covariance matrix for the variates, Θ is the PxP matrix of sums of squares and sums of cross-products for the variates, and $\Theta = N\Sigma$. Let V be the usual (biased) MLE for Σ and let $A = NV$. Since $\bar{x}$ is the MLE for μ, the final term in $Log_e(L)$ vanishes when calculating the sample log-likelihood. The estimates in V, and consequently in A, vary as a function of the pattern of means being fit to the data with the extreme cases being all means distinct for the pattern {1,2,3,4} and all means equal for the pattern {1234}. Assume that A is computed from distinct means, $\bar{x}$, for each repeated observation and let $\bar{x}_b$ represent the Px1 vector of suitably restricted means for a specified pattern (e.g.,

$\bar{x}_b = [\frac{1}{2}(\bar{x}_1 + \bar{x}_2), \frac{1}{2}(\bar{x}_1 + \bar{x}_2), \bar{x}_3, \bar{x}_4]'$ for the pattern {12,3,4}). Then, the modified value of A as a result of the restrictions is $A + n(\bar{x} - \bar{x}_b)(\bar{x} - \bar{x}_b)'$ (Anderson, 1984, page 61). Thus, computations for AIC based on the set of competing models are relatively straightforward.

Returning to the IQ example, the ordered pattern {12,3,4} = {IQ2, IQ8, IQ4, IQ13] is associated with the minimum value of both AIC and BIC (Table 2). This pattern is consistent with a cubic trend for which the peaks at age 2 and 4 are equal.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B.N. Petrov and F. Csake (eds.), *Second International Symposium on Information Theory*. Budapest: Akademiai Kiado, 267-281.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, AC-19*, 716-723.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. New York: Wiley.

Dayton, C. M. (1998). Information criteria for the paired-comparison problem. *American Statistician, 52*, 144-151.

Huang, Chuen-Chuen & Dayton, C. M. (1995). Detecting patterns of bivariate mean vectors using model-selection criteria. *British Journal of Mathematical & Statistical Psychology, 48*, 129-147.

Ramsey, F. & Schafer, D. (1997). *The Statistical Sleuth*. Duxbury Press.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461-464.

Table 1

Ramsay & Schafer IQ Data

Age	Mean	Variance
2	115.63	169.84
4	112.13	169.26
8	114.48	186.32
13	106.44	235.12

Table 2

AIC and BIC Values for IQ Data

Pattern	AIC	BIC
{1,2,3,4}	1912.64	1942.42
{12,3,4}	1911.33 *	1938.98 *
{1,23,4}	1913.84	1941.49
{1,2,34}	1922.30	1949.95
{12,34}	1919.20	1944.73
{1,234}	1943.32	1968.85
{123,4}	1913.94	1939.47
{1234}	1939.81	1963.21

* Minimum

Appendix

```
new;
/* SUBSETR - program for computing Akaike AIC, Schwarz BIC and Rissanen
RIC (Sclove, 1987) statistics for ordered subsets of means based on
repeated observations. Mean are organized by repeated measures
time period in one flat flat text file and the variance-covariance
matrix in a second flat test file.

Reference: C. Dayton (1998), Information Criteria for the Paired-
Comparisons Problem, American Statistician, 52, 144-151. */
/* All screen output is also directed to the file subset.out in the
current directory */
print "Default file for all printed output is Subset.out in the current
directory";
output file = subset.out reset;
print "Program SUBSETR for Ordered Subsets of Means or Proportions";
print "Prepared by: C. Mitchell Dayton";
print "Department of Measurement &amp; Statistics";
print "University of Maryland";
print "E-Mail: CD4@UMAIL.UMD.EDU";
/* Determine number of patterns to display on output */
print "Number of AIC/BIC values to display (5 is recommended) ";; nsave
= con(1,1);
print "Sample size = ";n=con(1,1);
/* Read means */
print "Name of input file for means? ";; fmt = cons;
loadm  ydat[] = ^fmt;
k = rows(ydat); /* k is number of repeated measures */
print "Name of input file for variance-covariance matrix? ";; fmt1 =
cons;
loadm cov[] = ^fmt1;
cov = reshape(cov,k,k);
aa = (n-1)*cov;
a = (k*n/2)*ln(2*pi);
xx=cumsumc(ones(k,1));
base=2;
aic1=zeros(nsave,k+1); bic1=aic1; ric1=aic1;
x=1;
do until x > 2^(k-1); x1=x-1;
/* Generate v with binary equivalent of integer = x-1 */
if x1 == 0; v = zeros(1,k);
elseif x1 > 0;
nbase = maxc(log(x1)./log(base))+1;
zbase=reshape(base,nbase,rows(x1));
vv = rev(recserrc(x1,zbase))';
v = zeros(1,k-rows(vv'))~vv;
endif;
/* Compute u that contains ordered subsets based on changes in v.
For example, if v = {0010}, then u = {1123} */
u = zeros(1,k);
i=2;
u[1]=1;
do until i > k;
if v[i] == v[i-1]; u[i] = u[i-1];
elseif v[i] /= v[i-1]; u[i] = u[i-1]+1; endif;
```

```
i = i+1; endo;
/* Create k-mean vector, mms, with equality restrictions imposed based
on patterns in u vector
*/
mns = zeros(1,k); nns=mns;
xb = mns;
nn = ones(1,k);
i=1;
do until i > maxc(u');
mns[i]= sumc(submat(ydat,indexcat(u',i),1));
nns[i]= sumc(submat(nn',indexcat(u',i),1));
mns[i]=mns[i]/rows(mns[i]);
i=i+1; endo;
mms = (subscat(u',xx,mns'))';
nns = (subscat(u',xx,nns'))';
xb = mms./nns;
/* Compute log-likelihoods and AIC, BIC values*/
b = (n/2)*ln(det((aa+n*(ydat-xb')*(ydat-xb')')/n));
lik = -.5*sumc(diag(inv((aa+n*(ydat-xb')*(ydat-xb')')/n)*(aa+n*(ydat-
xb')*(ydat-xb')'))) -a - b;
mod=k!/((k-2)!*2)+k;
aic = -2*lik + 2*(mod+maxc(u'));
bic = -2*lik + ln(n)*(mod+maxc(u'));
ric = -2*lik + ln((n+2)/24)*(mod+maxc(u'));
/* print aic~u~v; used as diagnostic only */
/* Save AIC, BIC if smaller than any currently saved case */
if x < nsave+1; aic1[x,.] = aic~u; bic1[x,.] = bic~u; ric1[x,.] =
ric~u; goto next1; endif;
if maxc(aic1[.,1]) > aic; aic1[maxindc(aic1[.,1]),.] = aic~u; endif;
if maxc(bic1[.,1]) > bic; bic1[maxindc(bic1[.,1]),.] = bic~u; endif;
if maxc(ric1[.,1]) > ric; ric1[maxindc(ric1[.,1]),.] = ric~u; endif;
next1:;
x = x+1;
endo; /* End of loop for generating mean patterns */
/* Sort saved AIC, BIC values in ascending order */
aic1 = sortc(aic1,1); bic1 = sortc(bic1,1); ric1 = sortc(ric1,1);
/* Print results for case of means */
format /mat /rd 10, 3;
ii=1; if k == 3; ii=2; endif;
print " "; print "Ordered means are: "; print ydat';
print " "; print "Best models";
print " "; print "Smallest Akaike AIC values and ordered subsets: ";
print aic1[ii:nsave,.];
print " "; print "Smallest Schwarz BIC values and ordered subsets: ";
print bic1[ii:nsave,.];
print " "; print "Smallest Rissanen RIC values and ordered subsets: ";
print ric1[ii:nsave,.];
format /mat /rd 15, 4;
output off;
end
```