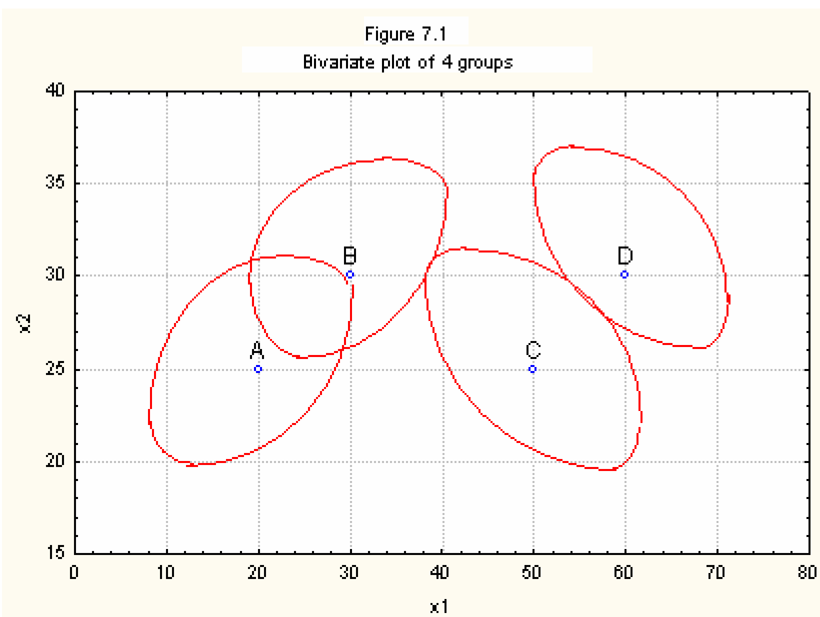
CHAPTER 7
Multivariate effect sizes indices

Seldom does one find that there is only a single dependent variable involved in a study. In Chapter 3's Example A we have the variables BDI, POMS_S and POMS_B, in Example E there are 5 cholesterol measurements as well as systolic and diastolic blood pressure measurements. Many times one will analyse and determine effect sizes indices for the variables separately. However, it is also sensible to consider the variables as a single unit where not only the different means and SDs play a role, but also the inter-correlations between the variables. Based on these analyses, called *multivariate* analyses, one can also obtain effect sizes indices.



In Figure 7.1 we reproduce Olejnik & Algina (2000)'s Figure 2 and use it to explain how a multivariate standardized difference in means can be interpreted. This representation is a scatter-plot of the two variables x_1 and x_2 in samples drawn from populations A, B, C and D. The points shown on the diagram each represent one of the group mean's $(\bar{x}_1, \bar{x}_2)$, while the ellipses are drawn in such a way that most of that group's measurements (say 90%) fall within it. In groups A and B we find a positive correlation between x_1 and x_2, while it is negative in groups C and D. Note that the groups A and B overlap, while there is almost no overlapping in groups C and D. The fact that the ellipses are almost the same size suggests that the SDs are equal for all the groups for x_1 and also for x_2. If one wants to compares the means $\bar{x}_1$ and $\bar{x}_2$ of groups A and B, then the *Mahalanobis-distance* D can be used, since it is an analogue of the univariate standardized difference. It is defined as:

$$D = \sqrt{\frac{1}{1-r_p^2}\left(\hat{\delta}_1^2 + \hat{\delta}_2^2 - 2r_p\hat{\delta}_1\hat{\delta}_2\right)} \quad (7.1)$$

where $\hat{\delta}_1$ and $\hat{\delta}_2$ are the standardized differences according to (4.3), and are defined as:

$$\hat{\delta}_1 = \frac{\bar{x}_{A1} - \bar{x}_{B1}}{s_{p1}} \quad \text{and} \quad \hat{\delta}_2 = \frac{\bar{x}_{A2} - \bar{x}_{B2}}{s_{p2}} \quad ,$$

where $\bar{x}_{A1}$ and $\bar{x}_{B1}$ are the means of x_1 for groups A and B and s_{p1} is the common SD for x_1 over groups A and B, etc.

Further, r_p is the pooled within-group correlation :

$$r_p = \frac{s_{p12}}{s_{p1}s_{p2}} \quad ,$$

where S_{p12} is the pooled covariance between x_1 and x_2.

Suppose that, in the example of the data in Figure 7.1, the pooled SD for x_1 is 8 and for x_2 it is 3. Further, let the r_p for groups A and B be 0,5 and for C and D let it be equal to -0,5. From the figure it is clear that if we want to compare groups A and B then we calculate the following: $\hat{\delta}_1 = (30 - 20)/8 = 1,25$ and

$\hat{\delta}_2 = (30 - 20)/3 = 1,67$, while

$$D^2 = \frac{1}{1 - 0,5^2}\left(1,25^2 + 1,67^2 - 2 \times 0,5 \times 1,25 \times 1,67\right) = 2,264/0,75$$
$$= 3,019,$$

Therefore, $D = 1,74$.

If we want to compare groups C and D, then $\hat{\delta}_1$ and $\hat{\delta}_2$ are the same as before but, because $r_p = -0,5$, we find that $D = 2,93$.

The following three points are important to note:

- First, the value of D is larger than the values of either of $\hat{\delta}_1$ and $\hat{\delta}_2$ because it represents the Euclidean distance between the 2-dimensional means on a scale which is determined by both the SDs of, and the correlation between, the variables x_1 and x_2. The $\hat{\delta}$ value, on the other hand, is only the difference between the one-dimensional means in the same units as the SD.
- Second, there is a larger separation between the points in groups C and D than in groups A and B, even though the mean points are equally far from one another. This separation can be ascribed to the fact that A and B have a positive correlation between x_1 and x_2, while the correlation between the variables for groups C and D is negative. The values of D reflect this occurrence, because, for A and B, the value $2\,r_p\hat{\delta}_1\hat{\delta}_2$ is subtracted from the positive value $\hat{\delta}_1^2 + \hat{\delta}_2^2$, while for C and D it is added (because $r_p < 0$). Thus while $D = 1,74$ for A and B, it is substantially higher, i.e., $2,93$, for C and D.
- Third, in addition to the sign, the magnitude of r_p also plays a role.

In the following paragraph the Mahalanobis-distance D is generalized to more than 2 variables. Later in the chapter we look at determining omnibus effects as well as contrast effects for the case with more than 2 groups.

7.1 Comparing two groups with m variables

In the introduction we illustrated the case where one compares the two-variable means of two populations.

Suppose now that two m-variable populations are completely observable and that

$$\boldsymbol{\delta} = \left(\delta_1, \delta_{2,}, ..., \delta_m\right) \tag{7.2}$$

represents the *vector* of *effect size indices* for the variables $x_1, x_2, ..., x_m$ for both populations.

Further, let

$$\boldsymbol{R} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \cdots \rho_{1m} \\ \rho_{12} & 1 \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \rho_{1m} & & \ddots & 1 \end{pmatrix} \tag{7.3}$$

denote the joint correlation matrix consisting of the inter-correlations between $x_1, x_2, ..., x_m$.

For comparing the m-variable means between two populations the *generalized Mahalanobis-distance* for populations can be used, namely

$$D = \sqrt{\boldsymbol{\delta}\ \boldsymbol{R}^{-1}\boldsymbol{\delta}'} \quad . \tag{7.4}$$

This involves matrix calculations which are difficult to do by hand. An easier formula for calculating this quantity is:

$$D^2 = \frac{\left(N_A + N_B\right)^2 \left(1 - \Lambda\right)}{N_A N_B \Lambda} \quad , \tag{7.5}$$

where N_A and N_B are the population sizes. The quantity Λ is known as *Wilk's lambda*, it calculated in a multivariate analysis of variance (MANOVA) where two mean vectors are compared and is provided in the output of any statistical computer package capable of MANOVA such as SPSS, SAS or STATISTICA.

Note that the value of $\hat{U}^{(1)} = \frac{1-\hat{\Lambda}}{\hat{\Lambda}}$ (which is called the *Hotelling-statistic*) is usually reported along with Wilk's lambda ($\hat{\Lambda}$) in most of the above mentioned computer packages.

In the case where random samples of size n_A and n_B are drawn from two populations, D can be estimated by $\tilde{D}$, defined as

$$\tilde{D}^2 = \frac{(n_A + n_B - 2)(1-\hat{\Lambda})\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}{\hat{\Lambda}} \quad , \tag{7.6}$$

or $$\tilde{D}^2 = (n_A + n_B - 2)\hat{U}(1)\left(\frac{1}{n_A}+\frac{1}{n_B}\right).$$

Here $\hat{\Lambda}$ is based on the sample data. Unfortunately, this estimator is actually biased. The estimator $\hat{D}$ is considered preferable since it is approximately unbiased. It is defined as

$$\hat{D}^2 = \left[(n_A + n_B - m - 3)\frac{1-\hat{\Lambda}}{\hat{\Lambda}} - m\right]\left(\frac{1}{n_A}+\frac{1}{n_B}\right) \quad , \tag{7.7}$$

or $$\hat{D}^2 = \left[(n_A + n_B - m - 3)\hat{U}(1) - m\right]\left(\frac{1}{n_A}+\frac{1}{n_B}\right).$$

Note also that if $\hat{D}^2$ is negative, then the value is changed to 0, seeing as $D^2 \geq 0$ per definition.

Further, the so called *Hotelling's* T^2 is obtained using the expression

$$T^2 = (n_A + n_B - 2)\frac{1-\hat{\Lambda}}{\hat{\Lambda}} = (n_A + n_B - 2)\hat{U}(1) \ ,$$

so that D , $\tilde{D}$ and $\hat{D}$ can also be calculated using $U^{(1)}$, $\hat{U}^{(1)}$ or T^2.

An approximate $(1-\alpha)100\%$ *CI* for D can also be calculated. This follows since $c.\frac{1-\hat{\Lambda}}{\hat{\Lambda}}$ has an approximate central-F distribution with degrees of freedom which are together with *c,* functions of the non-centrality parameter $ncp_D = \frac{n_A n_B}{(n_A + n_B)} D^2$, m , n_A and n_B. More details regarding this topic can be found in Appendix B.

A SAS-program named **VI_D** is available on the web page of this manual and make use of the following inputs: $\hat{\Lambda}$, n_A , n_B , m and α.

Example 7.1

Consider Example A from Chapter 3. Suppose that we want to compare the experimental and control groups with respect to the mean vector of the BDI before tests, after tests and follow-up tests. The MANOVA output obtained from STATISTICA is:

$$\hat{\Lambda} = 0{,}1513 \quad ; \quad F(3;46) = 86{,}02 \ (p < 0{,}0001).$$

With $n_A = n_B = 25$ and $m = 3$; it follows that $\hat{U}^{(1)} = \frac{1-\hat{\Lambda}}{\hat{\Lambda}} = 5{,}61$ and

$$\tilde{D} = \sqrt{48 \times 5{,}61\left(\frac{1}{25} + \frac{1}{25}\right)} = 4{,}64$$

and

$\hat{D} = \sqrt{(44 \times 5{,}61 - 3)(0{,}04 + 0{,}04)} = 4{,}42$, which is smaller than $\tilde{D}$ because it is approximately unbiased.

Applying the program VI_D produces the 95% *CI* for D: $(3{,}44;5{,}60)$.

This means that the Mahalonobis distance between the two groups' 3-dimensional mean vectors can be as small as $3,44$ and as large as $5,60$ with a 95% probability. □

7.1.1 Guideline values for *D*:

Since the value D is a 'distance' between 2 mean vectors which is weighted by the inter-correlations of $x_1,...,x_m$, it is measured on a different scale to $\delta_1,\delta_2,...,\delta_m$ and the individual effect size indices for the m variables. Only in very special circumstances are there relationships between D $\delta_1,...,\delta_m$. For simplicity, we will restrict our discussion to the case involving only 2 variables where D is given by equation (7.1).

In the extreme case, where $r_p = 0$, then D simplifies to $\sqrt{\hat{\delta}_1^2 + \hat{\delta}_2^2}$, but since there is no correlation between x_1 and x_2, each one can be considered separately and a multivariate analysis would be redundant. When r_p tends to 1, then D tends to infinity, but in these cases x_1 describes almost all of the variance of x_2 so that one only needs to analyse x_1 or x_2, and not both. In the case where $\hat{\delta}_1 = \hat{\delta}_2$ and where r_p tends to 1, then D tends to $\hat{\delta}_1 = \hat{\delta}_2$. Table 7.1 displays, for selected values of $\hat{\delta}_1 = \hat{\delta}_2$ and r_p, the values of D. It is clear that:

- $D > \hat{\delta}_1 = \hat{\delta}_2$
- D becomes larger when r becomes smaller
- The value of D is at a maximum value when r_p is close to -1. Similar to the case where r_p is close to 1, if r_p is close to -1, then only one of x_1 or x_2 should be analysed because almost all of the variance of the one variable is described by the other.

Table 7.1 provides some indications of the possible guideline values when we know the r_p values and when $\hat{\delta}_1$ and $\hat{\delta}_2$ are almost equal to one another, it is not possible in the more general case with m variables.

Table 7.1:
Values of Mahalanobis-distance if k=2, m=2,
$\delta_1 = \delta_2 = \delta$

δ \ r	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1
0.1	0.45	0.32	0.26	0.22	0.20	0.18	0.17	0.16	0.15
0.2	0.89	0.63	0.52	0.45	0.40	0.37	0.34	0.32	0.30
0.3	1.34	0.95	0.77	0.67	0.60	0.55	0.51	0.47	0.45
0.4	1.79	1.26	1.03	0.89	0.80	0.73	0.68	0.63	0.60
0.5	2.24	1.58	1.29	1.12	1.00	0.91	0.85	0.79	0.75
0.6	2.68	1.90	1.55	1.34	1.20	1.10	1.01	0.95	0.89
0.7	3.13	2.21	1.81	1.57	1.40	1.28	1.18	1.11	1.04
0.8	3.58	2.53	2.07	1.79	1.60	1.46	1.35	1.26	1.19
0.9	4.02	2.85	2.32	2.01	1.80	1.64	1.52	1.42	1.34
1	4.47	3.16	2.58	2.24	2.00	1.83	1.69	1.58	1.49

δ \ r	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.14	0.13	0.13	0.12	0.12	0.12	0.11	0.11	0.11	0.10
0.2	0.28	0.27	0.26	0.25	0.24	0.23	0.22	0.22	0.21	0.21
0.3	0.42	0.40	0.39	0.37	0.36	0.35	0.34	0.33	0.32	0.31
0.4	0.57	0.54	0.52	0.50	0.48	0.46	0.45	0.43	0.42	0.41
0.5	0.71	0.67	0.65	0.62	0.60	0.58	0.56	0.54	0.53	0.51
0.6	0.85	0.81	0.77	0.74	0.72	0.69	0.67	0.65	0.63	0.62
0.7	0.99	0.94	0.90	0.87	0.84	0.81	0.78	0.76	0.74	0.72
0.8	1.13	1.08	1.03	0.99	0.96	0.92	0.89	0.87	0.84	0.82
0.9	1.27	1.21	1.16	1.12	1.08	1.04	1.01	0.98	0.95	0.92
1	1.41	1.35	1.29	1.24	1.20	1.15	1.12	1.08	1.05	1.03

7.2 Effect sizes of contrast effects for m variables and k groups

As in Chapter 6, one can compare more than 2 groups using contrasts. A *contrast* now compares population mean vectors with one another and is defined as:

$$\boldsymbol{\psi} = c_1\boldsymbol{\mu}_1 + c_2\boldsymbol{\mu}_2 + ... + c_k\boldsymbol{\mu}_k \quad , \tag{7.8}$$

where

$$\boldsymbol{\mu}_1 = \left(\mu_{i1}, \mu_{i2}, ..., \mu_{im}\right)$$

and

$$\boldsymbol{\psi} = \left(\psi_1, \psi_2, ..., \psi_m\right) \quad ,$$

while c_1, $c_2, \ldots, c_k$ are the contrast weights of the k populations.

Example 7.2

In Example E of Chapter 3 the mean vectors of cholesterol-values (chol_0-chol_4) for the heart patients which did little or no exercise can be compared to those patients who exercised moderately to frequently with the contrast:

$$\boldsymbol{\psi} = 0,5\boldsymbol{\mu}_1 + 0,5\boldsymbol{\mu}_2 - 0,5\boldsymbol{\mu}_3 - 0,5\boldsymbol{\mu}_4 \quad ,$$

where $\boldsymbol{\mu}_i$ is the mean vector with the mean of each of the 5 cholesterol measurements as its components, while $\boldsymbol{\psi}$ is a vector with the contrast of each cholesterol measurement as its components. □

The Mahalanobis distance D_ψ can be used as an effect size index for $\boldsymbol{\psi}$, where:

$$D_\psi^2 = \frac{N\left(1-\Lambda_\psi\right)\left(\sum_{i=1}^{k}\frac{c_i^2}{N_i}\right)}{\Lambda_\psi} \quad , \tag{7.9}$$

and where, as before, N_i is the population size and $N = N_i + ... + N_k$, while Λ_ψ is Wilk's lambda for the contrast $\boldsymbol{\psi}$.

When k random samples are drawn from the populations, an estimator (Kline, 2004b : 10) $\tilde{D}_\psi$ is defined as

$$\tilde{D}_\psi^2 = \frac{(n-k)\left(1-\hat{\Lambda}_\psi\right)\left(\sum_{i=1}^{k}\frac{c_i^2}{n_i}\right)}{\hat{\Lambda}_\psi} \quad , \qquad (7.10)$$

where $n = n_1 + ... + n_k$ and $\hat{\Lambda}_\psi$ is now based on the sample data. The results of the MANOVA are used to test $\boldsymbol{\psi} = 0$.

An approximate $(1-\alpha)100\%$ *CI* for D_ψ is obtained using the same method as in paragraph 7.1. The SAS-program **VI_D-contrast** is available on the manual's web page and uses the following inputs:

$$\hat{\Lambda}_\psi, m, n_1, ..., n_k, c_1, ..., c_k \text{ and } \alpha .$$

Example 7.3:

Continuing with Example 7.2 the results of the MANOVA to test $\boldsymbol{\psi} = \boldsymbol{0}$ are:

$$\hat{\Lambda}_\psi = 0,765; \quad F(5;42) = 2,58\left(p = 0,0404\right).$$

Further, $n = 50$, $k = 4$, $n_1 = 19$, $n_2 = 20$, $n_3 = 9$, $n_4 = 2$ and

$c_1 = 0,5$, $c_2 = 0,5$, $c_3 = -0,5$ and $c_4 = -0,5$, so that

$$\tilde{D}_\psi^2 = \frac{(50-4)(1-0,765)\left(\frac{0,25}{19}+\frac{0,25}{20}+\frac{0,25}{9}+\frac{0,25}{2}\right)}{0,765}$$

$$= \frac{1,929}{0,765} = 2,521 \quad ,$$

thus $\tilde{D}_\psi = 1,588$

The 95% *CI* for D_ψ is then: $(0;2,248)$

(the lower bound could not be calculated, and was simply assumed to be equal to zero).

□

7.3 Multivariate omnibus effect

In the univariate case in Chapter 5, the effect size index η^2 is employed to measure the omnibus effect of differences in k population means. This was the proportion of the dependent variable's variance which could be attributed to population membership.

However, since Wilk's lambda Λ can be considered as the proportion of the generalized variance of a vector of m variables *not* attributed to population membership (see Kline, 2004b: 4), an obvious choice for effect size index is *Wilk's generalized correlation ratio:*

$$\eta^2_{mult} = 1 - \Lambda . \qquad (7.11)$$

Cohen(1988) proposed a generalisation of the effect size f^2 in multiple linear regression, based on Λ, namely

$$f_r^2 = \Lambda^{-1/r} - 1 \qquad (7.12)$$

where $$r = \sqrt{\frac{m^2(k-1)^2 - 4}{m^2 + (k-1)^2 - 5}} \ . \qquad (7.13)$$

Notes**:**

1. In multiple linear regression is m = 1, subsequently r = 1 and also $\Lambda = 1 - R^2$, therefore $f^2 = \frac{1}{1-R^2} - 1 = \frac{R^2}{1-R^2}$ (see Chapter 5, paragraph 5.2.3).
2. Table 7.2 can be used for reading off the value of r for selected k and m-values, while Table 7.3 then gives values of f_r^2 for selected Λ and R^2-values.
3. Contrary to its univariate counterpart f^2, the value of f_r^2 depends on k and m, making it difficult to state guidelines to whether f_r^2 is small, medium or large.

Table 7.2: Values for r for given values of k and m

		m											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**	**20**
	2	1	1	1	1	1	1	1	1	1	1	1	1
	3	1	2	2	2	2	2	2	2	2	2	2	2
	4	1	2	2.43	2.65	2.76	2.83	2.87	2.90	2.92	2.94	2.97	2.98
	5	1	2	2.65	3.06	3.32	3.49	3.61	3.69	3.75	3.79	3.90	3.94
	6	1	2	2.76	3.32	3.71	4.00	4.21	4.36	4.47	4.56	4.79	4.88
	7	1	2	2.83	3.49	4.00	4.39	4.69	4.92	5.10	5.24	5.62	5.78
	8	1	2	2.87	3.61	4.21	4.69	5.08	5.39	5.63	5.83	6.40	6.64
k	**9**	1	2	2.90	3.69	4.36	4.92	5.39	5.77	6.08	6.34	7.12	7.47
	10	1	2	2.92	3.75	4.47	5.10	5.63	6.08	6.46	6.78	7.78	8.25
	11	1	2	2.94	3.79	4.56	5.24	5.83	6.34	6.78	7.16	8.38	8.99
	12	1	2	2.95	3.83	4.63	5.35	5.99	6.56	7.05	7.48	8.93	9.68
	13	1	2	2.95	3.85	4.68	5.44	6.12	6.74	7.28	7.76	9.43	10.34
	14	1	2	2.96	3.87	4.73	5.51	6.23	6.89	7.47	8.00	9.89	10.95
	15	1	2	2.97	3.89	4.76	5.57	6.32	7.01	7.64	8.21	10.30	11.52
	20	1	2	2.98	3.94	4.87	5.76	6.61	7.42	8.18	8.90	11.82	13.82
	30	1	2	2.99	3.97	4.94	5.89	6.82	7.73	8.62	9.48	13.35	16.50

Table 7.3: Values for f_r^2 for selected values of r, Λ (or R^2)

	Wilks' Λ														
	0.99	**0.96**	**0.92**	**0.9**	**0.88**	**0.85**	**0.8**	**0.75**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
	R^2														
r	**0.01**	**0.04**	**0.08**	**0.1**	**0.12**	**0.15**	**0.2**	**0.25**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
1	0.01	0.04	0.09	0.11	0.14	0.18	0.25	0.33	0.43	0.67	1	1.5	2.33	4	9
2	0.01	0.02	0.04	0.05	0.07	0.08	0.12	0.15	0.2	0.29	0.41	0.58	0.83	1.24	2.16
2.5	0	0.02	0.03	0.04	0.05	0.07	0.09	0.12	0.15	0.23	0.32	0.44	0.62	0.9	1.51
3	0	0.01	0.03	0.04	0.04	0.06	0.08	0.1	0.13	0.19	0.26	0.36	0.49	0.71	1.15
3.5	0	0.01	0.02	0.03	0.04	0.05	0.07	0.09	0.11	0.16	0.22	0.3	0.41	0.58	0.93
4	0	0.01	0.02	0.03	0.03	0.04	0.06	0.07	0.09	0.14	0.19	0.26	0.35	0.5	0.78
4.5	0	0.01	0.02	0.02	0.03	0.04	0.05	0.07	0.08	0.12	0.17	0.23	0.31	0.43	0.67
5	0	0.01	0.02	0.02	0.03	0.03	0.05	0.06	0.07	0.11	0.15	0.2	0.27	0.38	0.58
r 5.5	0	0.01	0.02	0.02	0.02	0.03	0.04	0.05	0.07	0.1	0.13	0.18	0.24	0.34	0.52
6	0	0.01	0.01	0.02	0.02	0.03	0.04	0.05	0.06	0.09	0.12	0.16	0.22	0.31	0.47
6.5	0	0.01	0.01	0.02	0.02	0.03	0.03	0.05	0.06	0.08	0.11	0.15	0.2	0.28	0.43
7	0	0.01	0.01	0.02	0.02	0.02	0.03	0.04	0.05	0.08	0.1	0.14	0.19	0.26	0.39
7.5	0	0.01	0.01	0.01	0.02	0.02	0.03	0.04	0.05	0.07	0.1	0.13	0.17	0.24	0.36
8	0	0.01	0.01	0.01	0.02	0.02	0.03	0.04	0.05	0.07	0.09	0.12	0.16	0.22	0.33
9	0	0	0.01	0.01	0.01	0.02	0.03	0.03	0.04	0.06	0.08	0.11	0.14	0.2	0.29
10	0	0	0.01	0.01	0.01	0.02	0.02	0.03	0.04	0.05	0.07	0.1	0.13	0.17	0.26
12	0	0	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.04	0.06	0.08	0.11	0.14	0.21
15	0	0	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.05	0.06	0.08	0.11	0.17
Effect:	**Small**			**Medium**						**Large**					

In the k sample case, $\eta^2_{mult} = 1 - \Lambda$ can be estimated by:

$$\hat{\eta}^2_{mult} = 1 - \hat{\Lambda}\left(\frac{u+v}{v}\right)^r \quad , \qquad (7.14)$$

using a result of Cohen & Nee (1984). Here are

u = m (k – 1), v = wr + 1 – m (k – 1) / 2, with w = n – (m + k) / 2 –1. (7.15)

Substitution of (7.14) in (7.12) results in the estimator

$$\hat{f}_r^2 = \hat{\Lambda}^{-1/r}\left(\frac{v}{u+v}\right) - 1 \ . \qquad (7.16)$$

The following statistic is similar to the F-statistic in (5.12) in multiple linear regression:

$$F(u,v) = (\hat{\Lambda}^{-1/r} - 1)\frac{v}{u} \quad , \qquad (7.17)$$

which has an approximate F-distribution, with u and v degrees of freedom, under the null hypothesis of no difference between mean vectors (Cohen, 1988) . If the null hypothesis is not true, it also follows that *F(u,v)* is noncentral-F distributed, again with u and v degrees of freedom and noncentrality parameter

$$ncp = (\Lambda^{-1/r} - 1)(u + v + 1) \ . \qquad (7.18)$$

Similar to paragraph 5.2.5, $(1-\alpha)100\%$ CI's can be obtained for η^2_{mult} and f_r^2 as follows:

- Determine as in Appendix B a $(1-\alpha)100\%$ CI for ncp: (ncp_L, ncp_U) .
- From (7.18) follows: $\Lambda = \left(1 + \frac{ncp}{u+v+1}\right)^{-r}$ and $f_r^2 = \frac{ncp}{u+v+1}$. (7.19)
- By substitution of ncp_L and ncp_U in (7.19), the following CI's are obtained: $\left[\eta^2_{mult}(L), \eta^2_{mult}(U)\right]$ and $\left[f_r^2(L), f_r^2(U)\right]$ for η^2_{mult} and f_r^2 respectively.

These $(1-\alpha)100\%$ CI's for η^2_{mult} en f_r^2 can be calculated by means of the SAS-program **VI_eta2_meerv** obtained from the manual's website.

Other effect size indices based on Wilk's Λ $\left(\tau^2\right)$, the Hotelling-Lawley statistic $\left(\zeta^2\right)$ and on Pillai's statistic $\left(\xi^2\right)$, are given by Huberty (1994 : 194).

An approximate $(1-\alpha)100\%$ *CI* can be calculated for ζ^2, but we will first briefly discuss this index (see Steyn & Ellis, 2009).

It is defined as

$$\zeta^2 = \frac{U^{(s)}}{s+U^{(s)}} \quad , \qquad (7.20)$$

where $U^{(s)}$ is the Hotelling-Lawley quantity (see Appendix C) for m-variable populations and $s = min(m,k-1)$.

An estimator which is approximately unbiased for ζ^2 is given by (see Appendix C.1):

$$\hat{\zeta}_1^2 = \frac{(n\text{-}k\text{-}m\text{-}1)\hat{U}^{(s)} - m(k-1)}{ns + (n\text{-}k\text{-}m\text{-}1)\hat{U}^{(s)} - m(k-1)} \quad , \qquad (7.21)$$

where $\hat{U}^{(s)}$ is the Hotelling-Lawley statistic– based on k samples from the m-variable normal populations. This statistic is usually reported together with Wilk's lambda in the output of a MANOVA in computer packages such as SAS, SPSS or STATISTICA.

An approximate $(1-\alpha)100\%$ *CI* for ζ^2 can be calculated using the SAS-program **VI_zeta_kwadr1** on the manual's web page. Let $\left(\zeta_{L1}^2, \zeta_{U1}^2\right)$ denote this interval (see Appendix C.1).

Both the estimator and the *CI* are suitable for use with smaller samples.

When the sample size, *n*, is large, an asymptotically unbiased estimator (see Appendix C.2) is given by:

$$\hat{\zeta}_2^2 = \frac{(n-k)\hat{U}^{(s)} - m(k-1)}{ns + (n-k)\hat{U}^{(s)} - m(k-1)}. \qquad (7.22)$$

Both estimators for ζ^2 can be negative if $\hat{U}^{(s)}$ is very small. The quantity ζ^2 is theoretically non-negative, and so, if the estimator turns out to be negative we set the estimator to zero.

An asymptotic $(1-\alpha)100\%$ *CI* for ζ^2 can also be obtained using a SAS-program (**VI_zeta_kwadr2**) which is available on the manual's web page. Let $(\zeta_{L2}^2, \zeta_{U2}^2)$ be used to denote this interval (see Appendix C.2).

Example 7.4:

Consider Example F of Chapter 3. The results of a MANOVA on the variables S_Cho, S_Tri, HDL_C and LDL_C for comparing the 3 activity groups, was:

$n = 1362, \quad F(8;2710) = 22,28 \ (p < 0,0001)$.

Futher

$\hat{U}^{(2)}$=0,136, $F(8;2710) = 22,95 \ (p < 0,0001)$.

$r = \sqrt{\frac{4^2(3-1)^2 - 4}{4^2 + (3-1)^2 - 5}} = \sqrt{\frac{60}{15}} = 2$, u = 4 (3 – 1) = 8, w = 1362 –(4 + 3) / 2 = 1357,5,

v = 1357,5 (2) + 1 – 4 (2) / 2 = 2712.

$\hat{f}_r^2 = 0,88^{-1/2}\left(\frac{2712}{8+2712}\right) - 1$ = 1,066 (0,997) – 1 = 0,063.

95% CI for f_r^2: (0,045; 0,084) - with SAS-program **VI_eta2_meerv**.

$\hat{\eta}_{mult}^2 = 1 - 0,88\left(\frac{8+2712}{2712}\right)^2$ = 1 – 0,88 (1,006) = 0,115.

95% CI for η^2_{mult}: (0,084; 0,149) - with SAS-program **VI_eta2_meerv**.

Estimation of ζ^2: Since n is very large, we find that

$$\zeta_2^2 = \frac{1359 \times 0{,}136 - 8}{1362 \times 1359 \times 0{,}136 - 8} = \frac{176{,}8}{2724 \times 176{,}8}$$

$$= 0{,}0609 \quad ,$$

where the approximate 95% *CI* is: $(0{,}045\ ;\ 0{,}079)$ which means that ζ^2 can be as small as $0{,}045$ and as large as $0{,}079$ with 95% probability. □

7.4 Effect sizes indices for canonical correlation

In paragraph 5.2 $\rho^2_{y.A}$ and $R^2_{y.A}$ are considered to be the effect sizes for the proportion of y 's variance which is explained by the multiple linear regression

$$\hat{y} = a + b_1 x_1 + b_2 x_2 + ... + b_u x_u \,,$$

in a population and random sample respectively. The notation 'A' refers to the set of predictors $x_1, x_2, \ldots, x_u$. Let $b_1, b_2, \ldots, b_u$ be the regression coefficients, i.e., weights which are chosen so that the correlation, $\rho_{y.A}$ or $R_{y.A}$, between y and $\hat{y}$ is a maximum. The constants do not play a role in determining $\rho^2_{y.A}$ and $R^2_{y.A}$ and can be assumed to be zero, which implies that $y - \bar{y}$, $x_1 - \bar{x}_1$, …, $x_u - \bar{x}_u$ are used instead of y, $x_1, \ldots, x_u$.

Suppose that, instead of using only one criteria variable y, we use a set B which consists of y_1, $y_2, \ldots, y_v$. The question is now: How can we determine the relationship between the two sets of variables?

The first step is to determine the *canonical variables* or principal components of A and B. In other words, we must determine a linear combination of the elements of A, x_1, $x_2,\dots,x_u$, i.e., $A_1 = a_{11}x_1 + a_{12}x_2 + ... + a_{1u}x_u$, such that it explains the maximum proportion of the total variance in A. Next a second linear combination A_2 is found which attempts to explain the maximum proportion of the remaining variance in A. We proceed in this way until the linear combination A_u is obtained. These linear combinations of $x_1,\dots,x_u$ are known as the *canonical variables.* There are as many canonical variables as there are original variables. The same technique can be applied to B; the first canonical variable is $B_1 = b_{11}y_1 + b_{12}y_2 + ... + b_{1v}y_v$, and the remaining variables are defined similarly and denoted by $B_2,...,B_v$. Each of the correlations of $x_1,\dots,x_u$ with A_1, are known as the *loadings* (or structure coefficients) of $x_1,\dots,x_u$ on the first canonical variable. The mean of the sum of squares of the loadings describes the *proportion variance* of the set A which is explained by A_1 and is denoted by VA_1. Similarly for $VA_2,\dots,VA_u$ and also $VB_1,...,VB_v$ can be obtained.

The second step is to 'rotate' the first pair of canonical variables A_1 and B_1 so that the correlation between them is a maximum. Rotation means that the set of weights $a_{11},...,a_{1u}$ and $b_{11},...,b_{1v}$ are manipulated to obtain *canonical weights* which are defined in such a way that the proportions of variance VA_1 and VB_1 remain unchanged. The maximum correlation is known as the *first canonical correlation* and is denoted by λ_1. In a similar fashion we can obtain $\lambda_1,\dots,\lambda_s$, called the canonical correlations between the pairs (A_2, B_2),…,(A_s, B_s), where *s* is the minimum of *u* and *v*.

The *first canonical correlation* λ_1 is based on A_1 and B_1 (both of which describe the maximum proportions variances from their respective sets) and it is a good indication of the relationship between the sets A and B. The squared value, λ_1^2, is then the proportion variance which A_1 or B_1 explain, and is called the *eigen value*.

The linear combinations A_1 and B_1 only explain a portion of the total variance in A and B (the proportions VA_1 and VB_1); as a result, λ_1 can not actually be used to measure the relationship between the sets of variables $x_1,...,x_u$ and $y_1,\ldots y_v$. Rather, it represents a sort of partial correlation – the correlation between A_1 and B_1 without considering $A_2,\ldots,A_s$ and $B_2,\ldots,B_s$ (the so called "redundancy").

A better measure of the *redundancy of B given A* which is defined by Stewart & Lowe (1968) as:

$$\bar{R}^2_{B.A} = \sum_{i=1}^{s} \lambda_i^2 VB_i = \sum_{i=1}^{s} \lambda_i^2 \left(\sum_{j=1}^{v} \ell_{ij}^2 \right) / v \ , \qquad (7.23)$$

where ℓ_{ij} is the loading of y_j on B_i .

In this formulation, the set B is considered to be the criterion (response) values and A as the predictors so that $\bar{R}^2_{B.A}$ is the *redundancy of the criterion set given the predictors*. If A and B's roles are reversed, then

$$\bar{R}^2_{A.B} = \sum_{i=1}^{s} \lambda_i^2 VA_i = \sum_{i=1}^{s} \lambda_i^2 \left(\sum_{j=1}^{u} \ell_{ij}^2 \right) / u \ , \qquad (7.24)$$

where ℓ_{ij} is the loading of x_j on A_i.

Another method to obtain $\bar{R}^2_{B.A}$ is:

$$\bar{R}^2_{B.A} = \sum_{i=1}^{v} R_i^2 / v \qquad (7.25)$$

where R_i is the multiple correlation coefficient of $x_1,\ldots,x_u$ with y_i. It is for this reason that we use the notation $\bar{R}^2_{B.A}$. Likewise, it follows that

$$\overline{R}^2_{A.B} = \sum_{i=1}^{u} R_i^2 / u \ , \qquad (7.26)$$

where R_i is the multiple correlation coefficient of $y_i, \ldots, y_v$ with x_i .

The redundancies $\overline{R}^2_{B.A}$ and $\overline{R}^2_{A.B}$ are defined for populations as well as samples and because the only difference between them is the type of data used (full population data or sample data) we will not distinguish between the two. This notation will be maintained throughout the text.

In order to illustrate all of the concepts and their interpretations, we will discuss them by referring to the example that follows this paragraph. Canonical correlations and redundancy can be very difficult to calculate by hand and we usually make use of computer packages to perform canonical correlation analysis. As a result, the output of one of these computer packages, e.g., STATISTICA, will be used in the example.

Example 7.5:

In Example A of Chapter 3, we want to determine the relationship between the BDI's before tests, after tests and follow-up tests (BECKPRE, BECKPO1 and BECKPO2) on the one side (B) and the POMS_A (TENSPRE, TENSPOS1 and TENSPOS2) and POMS_D (DEPRE, DEPPOS1 and DEPPOS2) on the other (A). The output of a canonical analysis from STATISTICA is the following:

(a)

N=50	Canonical Analysis Summary (kdklerk_total scales) Canonical R: .81938 Chi²(18)=70.097 p=0.0000 Include condition: group=101 or group=103	
	Left Set	Right Set
No. of variables	3	6
Variance extracted	100.000%	77.4621%
Total redundancy	48.4597%	38.5355%
Variables: 1	BECKPRE	TENSPRE
2	BECKPO1	DEPPRE
3	BECKPO2	TENSPOS1
4		DEPPOS1
5		TENSPOS2
6		DEPPOS2

(b)

Root	Eigenvalues (kdklerk_total scales) Include condition: group=101 or group=103		
Root	Root 1	Root 2	Root 3
Value	0.671381	0.242453	0.183382

c)

Variable	Factor Structure, left set (kdklerk_total scales) Include condition: group=101 or group=103		
Variable	Root 1	Root 2	Root 3
BECKPRE	0.348103	0.739241	-0.576496
BECKPO1	0.825877	0.336229	0.452633
BECKPO2	0.982646	-0.158999	-0.095525

(d)

Factor	Variance Extracted (Proportions), left set (kdklerk_total Include condition: group=101 or group=103	
Factor	Variance extractd	Reddncy.
Root 1	0.589615	0.395856
Root 2	0.228269	0.055345
Root 3	0.182116	0.033397

(e)

Variable	Factor Structure, right set (kdklerk_total scales) Include condition: group=101 or group=103		
	Root 1	Root 2	Root 3
TENSPRE	0.148037	0.627317	-0.444477
DEPPRE	0.311308	0.732679	-0.542715
TENSPOS1	0.772995	-0.102487	0.129677
DEPPOS1	0.875580	0.369798	0.232106
TENSPOS2	0.776935	-0.181875	-0.112816
DEPPOS2	0.877712	-0.006656	-0.323257

(f)

Variable	Variance Extracted (Proportions), right set (kdklerk_tota Include condition: group=101 or group=103	
	Variance extractd	Reddncy.
Root 1	0.476166	0.319688
Root 2	0.185120	0.044883
Root 3	0.113335	0.020784

(g)

Variable	Canonical Weights, left set (kdklerk_total scales) Include condition: group=101 or group=103		
	Root 1	Root 2	Root 3
BECKPRE	0.059731	0.784921	-0.692047
BECKPO1	0.243237	0.799263	1.171768
BECKPO2	0.792069	-0.949809	-0.739669

(h)

Variable	Canonical Weights, right set (kdklerk_total scales) Include condition: group=101 or group=103		
	Root 1	Root 2	Root 3
TENSPRE	-0.248088	0.461440	-0.12587
DEPPRE	0.170837	0.358266	-0.62371
TENSPOS1	0.378857	-0.855018	-0.34040
DEPPOS1	0.282165	1.130572	1.42870
TENSPOS2	0.180081	0.349648	0.18841
DEPPOS2	0.346035	-0.889216	-1.04977

The table in (a) shows the first canonical correlation to be $\lambda_1 = 0{,}819$. The second and third canonical correlations are $\lambda_2 = \sqrt{0{,}2425} = 0{,}492$ and $\lambda_3 = \sqrt{0{,}1834} = 0{,}428$, which were obtained from λ_1^2, λ_2^2 and λ_3^2 in table (b). Note that the minimum of the number of variables in each group is $s = 3$, therefore there are only 3 canonical correlations. The BDI group's first canonical variable,

B_1, explains $0{,}59$ of the variance of the 3 variables (see table (d)) which is obtained from the sum of squares of the first column of table (c) divided by 3: $VB_1 = (0,348^2 + 0,826^2 + 0,983^2)/3 = 0,590$. Table (c) shows the loadings of each variable on the canonical variables. In the same way $VA_1 = 0{,}476$ from table (f), which is the proportion variance of the 6-POMS-variables which is explained by A_1. The canonical correlation $\lambda_1 = 0{,}819$ is thus the correlation between A_1 and B_1, but which each explain the proportions of $0{,}476$ and $0{,}590$ for their respective groups.

The redundancy of B given A:
$\overline{R}^2_{B.A} = 0{,}4846$ which is obtained from Table (a) and is calculated using equation (7.18):

$$\begin{aligned}\overline{R}^2_{B.A} &= 0{,}6714 \times 0{,}5896 + 0{,}2425 \times 0{,}2283 + 0{,}1834 \times 0{,}1822 \\ &= 0{,}3959 + 0{,}0554 + 0{,}0334 \\ &= 0{,}4847\end{aligned}$$

(Note that the last column of (d) provides the individual products calculated above).

If we run a multiple regression of all of the POMS variables on each of the BDI variables, then the multiple correlations are: $0{,}524$; $0{,}723$; $0{,}810$, and the redundancy is $\overline{R}^2_{B.A} = \left(0{,}524^2 + 0{,}723^2 + 0{,}810^2\right)/3 = 0{,}4844$.

If the POMS are used as the criterion, then $\overline{R}^2_{A.B} = 0{,}3854$ (see table (a)). □

7.5 Guidelines for multivariate omnibus-effects

As in the case of the univariate omnibus effect η^2 and its estimators discussed in Chapter 6, the question now is: When should these effects be considered 'large' or 'small'? Kline (2004b) attempts to solve this problem by first looking at the case involving two multivariate populations, where η^2_{multiv} is the proportion variance of the discriminant function (DF) explained by population membership. The DF is nothing other than the linear combination of the response variables $y_1,...,y_m$ which represents the maximum correlation of the dichotomous grouping variable. Therefore $\eta^2_{multiv} = \lambda_1^2$, with λ_1 the canonical correlation between DF and the grouping variable. The redundancy of the m response variables, given the dichotomous grouping variable, is then $\lambda_1^2 V_1$, where V_1 is the proportion variance of the DF.

Kline(2004b) supplies two reasons why η^2_{mult} can be substantially larger than any individual y_i's η^2:

(a) η^2_{mult} represents the proportion of generalized variance based on m variables simultaneously;

(b) Population membership explains only a portion of the variance in the DF, while the DF, in turn, only explains a portion of the variance of the set of response variables (see the previous paragraph: the DF is a canonical variable, which is rotated, with a variance smaller than the total variance of the response variables). Thus η^2_{mult} is the proportion of a portion of the variance and must thus be larger than the true proportion (e.g., if the DF's variance is 3,5 and the total variance of all the outcome variables is 5, while the variance accounted for population membership is assumed to be 2, then η^2_{mult} = 2/3,5 = 0,57, which is clearly larger than the proportion which would have been obtained from the total variance 5).

Kline's arguments can be extended to k populations. In this case there is a set of $k-1$ dichotomous grouping variables (also called dummy variables) which indicate population membership. Consequently the relationship between the m response variables and the $k-1$ grouping variables can be expressed as the canonical correlation $\lambda_1,...,\lambda_s, s = min(k-1,m)$. Wilk's lambda can be written in terms of canonical correlations as:

$$\Lambda = (1-\lambda_1^2)(1-\lambda_2^2)...(1-\lambda_s^2). \quad (7.27)$$

Since the values $\lambda_1^2,\ldots,\lambda_s^2$ are each a proportion variance of a canonical variable based on the response variables (with respect to a canonical variable which is based on grouping variables) it does not capture the entire proportion variance within the response variables. Consequently, the proportion η_{mult}^2 is larger than when it is expressed in terms of the response variable's variance.

According to the previous paragraph the redundancy of the *response variables, given the grouping variables* (set A), represents the true proportion. It can thus serve as an effect size index and is given by:

$$\bar{R}_{y.A}^2 = \sum_{i-1}^{s} \lambda_i^2 V_i , \quad (7.28)$$

where V_i is the proportion variance explained by the $i-$th canonical variable of $y_1,...,y_m$. In practice it is easier to calculate $\bar{R}_{y.A}^2$ using:

$$\bar{R}_{y.A}^2 = \sum_{i=1}^{m} \eta_i^2 / m \quad , \quad (7.29)$$

where η_i^2 is the proportion variance attributed to population membership for y_i, the usual effect size indices as defined in Chapter 6. Note that η_i is nothing other than the multiple correlation coefficient of y_i on the $k-1$ dichotomous grouping variables.

If one is working with samples, the redundancy is obtained by replacing η_i^2 with with its estimator $\hat{\eta}_i^2$.

Example 7.6

Once again, consider Example 7.4. A MANOVA is conducted on the variables S_CHOL, S_TRI, HDL_C and LDL_C to compare the three activity groups of men. With the aid of STATISTICA, the following results are obtained: $n = 1359$, $k = 3$, $m = 4$, $\hat{\Lambda} = 0{,}8804$. After that an ANOVA is conducted for each variable and the following F-values are obtained: $60{,}49$; $53{,}43$; $24{,}59$ and $40{,}70$. The estimator for η^2_{mult} was $0{,}1182$.

We use equation (6.11) in the estimation of S_CHOL's proportion variance which is explained by population membership:

$$\hat{\eta}_1^2 = \frac{(n-k-2)F/(n-k)-1}{(n-k-2)F/(n-k)+(n-k)/(k-1)} = \frac{1354\times 60{,}49/1356-1}{1354\times 60{,}49/1356+1356/2}$$
$$= 0{,}0804$$

Similarly, it follows that $\hat{\eta}_2^2 = 0{,}0716$, $\hat{\eta}_3^2 = 0{,}0335$ and $\hat{\eta}_4^2 = 0{,}0552$.

The redundancy is thus

$$\bar{R}^2_{y.A} = \frac{1}{4}(0{,}0804 + 0{,}0716 + 0{,}0335 + 0{,}0552)$$
$$= 0{,}0602.$$

While $0{,}1182$ represents the proportion variance which the two canonical variables contribute to population membership, the redundancy is the proportion variance of the four response variables which is explained. It is almost half the size of η^2_{mult}. □

An alternative index which is also based on Λ is (Huberty, 1997: 194):

$$\tau^2 = 1 - \Lambda^{1/s} , \qquad (7.30)$$

with estimator

$$\hat{\tau}^2 = 1 - \hat{\Lambda}^{1/s} . \qquad (7.31)$$

From expressions (7.21) and (7.23) it follows that $\Lambda^{1/s}$ is the geometric mean of the $\left(1-\lambda_i^2\right)$'s.

Then we have that

$$\tau^2 < \eta_{mult}^2 .$$

Because

$$\frac{1}{s}U^{(s)} = \frac{1}{s}\sum_{i=1}^{s}\frac{\lambda_i^2}{1-\lambda_i^2}$$, is the arithmetic mean of the $\frac{\lambda_i^2}{1-\lambda_i^2}$'s:

$$\zeta^2 = \frac{1}{s}\sum_{i=1}^{s}\frac{\lambda_i^2}{1-\lambda_i^2} \Big/ \left(n+\frac{1}{s}\sum_{i=1}^{s}\frac{\lambda_i^2}{1-\lambda_i^2}\right). \tag{7.32}$$

However ζ^2 also depends on $\lambda_1,...,\lambda_s$, and it is thus a more complex relationship than for η_{mult}^2 and τ^2.

The simplest function of $\lambda_1,...,\lambda_s$ is based on Pillai's statistic (Huberty, 1994: 194):

$$\xi^2 = \frac{1}{s}\sum_{i=1}^{s}\lambda_i^2 \ , \tag{7.33}$$

of which Roy's statistic is a special case:

$$\theta = \lambda_1^2 . \tag{7.34}$$

Note that while ξ^2 is the mean of the λ_i^2's, the redundancy, from (7.22), is a *weighted mean* of the λ_i^2's with weights equal to the proportion variances, V_i, explained by the canonical variables. If the variance explained by the first canonical variable, V_i, is close to 1, then θ, ξ^2 and $\bar{R}_{y.A}^2$ are almost the same, and any of these can be used as an effect size index.

Each of the indices η_{mult}^2, ζ^2 , τ^2 , ξ^2 and θ are functions of the canonical correlations in some way or another and they all lie between 0 and 1. As a result they can be considered as differenced proportions; the larger the value, the larger the effect. All of these quantities express proportions in terms of the *s* canonical variables which, after rotation, have the maximum correlations with the grouping variable. These canonical variables do not capture all of the variance in the *m* response variables, except when $m \le k-1$. In these cases $s = m$ and $\lambda_1,...,\lambda_s$ explains all the variance of $y_1,...,y_m$.

To recap: It is difficult to provide guidelines for determining when the the omnibus effect is 'small' or 'large'. Reasons for this include:

- The indices η^2_{mult}, τ^2, ζ^2, ξ^2 and θ are based on the MANOVA statistics: Wilks', Hotelling-Lawley's, Pillai's and Roy's statistics. However, while they are all functions of the canonical correlations, $\lambda_1,...,\lambda_s$, they are all different functions, and so all the values are different.
- The canonical correlations measure the relationship between the *s* discriminant functions (based on the *m* response variables) and the set of $k-1$ dichotomous grouping variables. They do not provide a direct relationship between *m* response variables and thus the proportions η^2_{mult}, τ^2, ζ^2, ξ^2 and θ are all too large.

The redundancy $\bar{R}^2_{y.A}$ presents the proportion variance, but compensates for the portions of the variances explained by the discriminant functions. Therefore we recommend that $\bar{R}^2_{y.A}$ be used as the *primary effect size index* for jointly measuring the proportion of the response variables which can be attributed to population membership. However, since it is the mean of the univariate η^2's, the same guidelines can be used as those provided in Chapter 6. Unfortunately, this mean can be greatly influenced by one very large η^2-value, and so a great amount of consideration and care should be employed before making use of these guidelines.

Example 7.7:

Consider Example 7.4. The following tables show the results obtained from STATISTICA:

Effect	Multivariate Tests of Significance (dvdwesthuizen) Sigma-restricted parameterization Effective hypothesis decomposition Include condition: geslag=1					
	Test	Value	F	Effect df	Error df	p
Intercept	Wilks	0.03053	10766.06	4	1356	0.00
	Pillai's	0.96947	10766.06	4	1356	0.00
	Hotellng	31.75828	10766.06	4	1356	0.00
	Roy's	31.75828	10766.06	4	1356	0.00
Akt_grp	Wilks	0.88044	22.28	8	2712	0.00
	Pillai's	0.11982	21.62	8	2714	0.00
	Hotellng	0.13550	22.95	8	2710	0.00
	Roy's	0.13330	45.22	4	1357	0.00

Dependnt Variable	Test of SS Whole Model vs. SS Residual (dvdwesthuizen) Include condition: geslag=1											
	Multiple R	Multiple R²	Adjusted R²	SS Model	df Model	MS Model	SS Residual	df Residual	MS Residual	F	p
S_CHO	0.285917	0.081749	0.080397	1827693	2	913846.6	20529796	1359	15106.55	60.49342	0.000000
S_TRI	0.269990	0.072895	0.071530	1461133	2	730566.6	18583351	1359	13674.28	53.42632	0.000000
HDL_C	0.186886	0.034926	0.033506	30022	2	15010.8	829543	1359	610.41	24.59143	0.000000
LDL_C	0.237713	0.056508	0.055119	952893	2	476446.7	15910235	1359	11707.31	40.69651	0.000000

Previously we obtained the following estimates:

$$\eta^2_{mult} \quad : \quad \hat{\omega}^2_{mult} = 0{,}1186$$

$$\zeta^2 \quad : \quad \hat{\zeta}^2 = 0{,}0609 \text{ with } 95\% \ CI : (0{,}045;\ 0{,}079).$$

When Wilk's lambda is equal to $0{,}8804$ it follows that the estimator of τ^2 is:

$$\hat{\tau}^2 = 1 - 0{,}8804^{1/2}$$
$$= 0{,}0617.$$

When Pillai's statistic is equal to $0{,}1198$ it follows that the estimator of ξ^2 is $\hat{\xi}^2 = \frac{1}{2} \times 1198 = 0{,}0599$, and when Roy's statistic is equal to $0{,}1333$ the estimator is $\hat{\theta} = 0{,}1333$.

From the second table the values of $\hat{\eta}_i^2$ follow directly as the modified R^2-values, i.e., $0{,}0804$; $0{,}0715$; $0{,}0335$ and $0{,}0551$, as calculated in Example 7.4. The mean of these values is then the redundancy:

$$\bar{R}_{y.A}^2 = 0{,}0602 .$$

All these values represent the proportions of variances of S_CHOL, S_TRIG, HDL_C and LDL_C which can be attributed to the activity groups. The redundancy index is $0{,}06$ which is also roughly given by the estimates $\hat{\zeta}^2$, $\hat{\tau}^2$, $\hat{\xi}$ and thus, in terms of the guidelines in Chapter 6, it indicates a medium effect.

□