

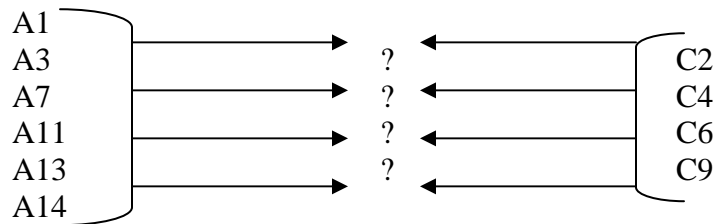
CANONICAL CORRELATION

Amanda Glatfelter
COM 531, Spring 2009

To find patterns of linear relationships between two sets of variables, in this case items from the pre-existing scales of communication apprehension and neuroticism, using the Neuendorf, Skalski, Powers Humor data set.

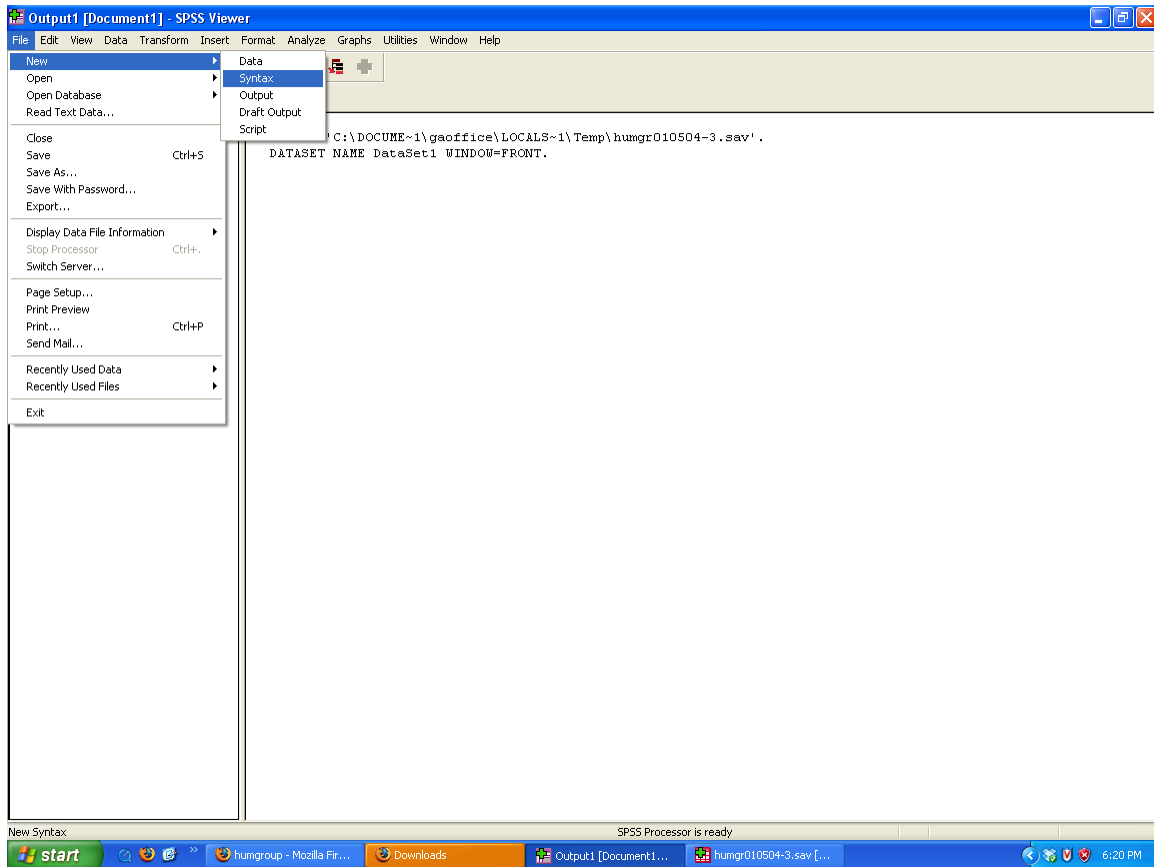
SET 1: Communication Apprehension	SET 2: Neuroticism
A1: I dislike participating in groups.	C2: Are you inclined to be moody?
A3: I am tense and nervous while participating in group discussion.	C4: Does your mind often wander when you are trying to concentrate?
A7: Generally, I am nervous when I have to participate in a meeting.	C6: Do you sometimes feel happy, sometimes depressed, without any apparent reason?
A11: Communicating at meetings usually makes me uncomfortable.	C9: Are you frequently lost in thought even when you are supposed to be taking part in a conversation?
A13: While participating in a conversation with a new acquaintance, I feel very nervous.	
A14: I have no fear of speaking up in public.	

Note: *Set 1* used a 1-5 point scale, with 1=strongly agree, 3= undecided and 5=strongly disagree. *Set 2* used a 0-10 point scale, with 0= definitely no and 10=definitely yes.



Canonical Correlation can only be completed by using **SYNTAX**.

Click **File** → **New** → **Syntax**



In the new SPSS syntax editor box type the general form:

INCLUDE 'C:\Program Files\SPSS\Canonical correlation.sps'.
CANCORR SET1=enter 1st block of variables/
SET2=enter 2nd block of variables/.

As noted in class, canonical correlation will NOT run on SPSS 16.0. Also, on non-departmental computers the macro command ('C:\.....') may be different, and therefore not work. See Dr. Neuendorf's handout for alternate syntax options if a problem occurs.

The screenshot shows the SPSS Data Editor window with a data view of 1020 variables. A Syntax Editor window is open, displaying the following syntax:

```
INCLUDE 'C:\Program Files\SPSS\Canonical correlation.sps'.
CANCORR SET1=A1,A3,A7,A11,A13,A14/
SET2=C2,C4,C6,C9/.
```

The data view shows columns for 'id', 'var00001', 'g_comp', and variables 'a1' through 'a9'. The data is organized into rows, with the first 24 rows showing a sequence of IDs and corresponding values for the variables. The Syntax Editor window is positioned over the data view, and the SPSS Processor is ready.

Click on **Run** → **All**

The screenshot shows the SPSS Data Editor window with a data table and a Syntax Editor window. The data table has columns for 'id', 'var00001', 'q_comp', and variables 'a1' through 'a9'. The Syntax Editor window contains the following code:

```
INCLUDE 'C:\Program Files\SPSS\Canonical c
CANCORR SET1=A1,A3,A7,A11,A13/
SET2=C2,C4,C6,C9/
```

A context menu is open over the 'Run' button in the Syntax Editor, with 'All' selected. The status bar at the bottom of the Syntax Editor window says 'Run All' and 'SPSS Processor is ready'. The Windows taskbar at the bottom shows the Start button and several open applications including 'humgr010504-3...', 'Syntax2 - SPSS...', and 'CANONICAL CO...'. The system clock shows 6:24 PM.

```

INCLUDE 'C:\Program Files\SPSS\Canonical correlation.sps'.
1033 * Canonical correlation.sps. This version allows long variable names and
      uses datasets.Canonical correlation.sps.
1034
1036 preserve.
1038 set printback=off.
1742 RESTORE.
1743
1745 * End of INSERT and INCLUDE nesting level 01.
CANCORR SET1=A1,A3,A7,A11,A13,A14/
        SET2=C2,C4,C6,C9/.

```

Matrix

Run MATRIX procedure:

Correlations for Set-1

	a1	a3	a7	a11	a13	a14
a1	1.0000	.2570	.2356	.2247	.1989	-.0811
a3	.2570	1.0000	.5199	.4152	.3557	-.3813
a7	.2356	.5199	1.0000	.5120	.4454	-.3630
a11	.2247	.4152	.5120	1.0000	.4245	-.2828
a13	.1989	.3557	.4454	.4245	1.0000	-.4045
a14	-.0811	-.3813	-.3630	-.2828	-.4045	1.0000

Correlations for Set-2

	c2	c4	c6	c9
c2	1.0000	.2385	.4219	.1804
c4	.2385	1.0000	.2700	.3487
c6	.4219	.2700	1.0000	.3060
c9	.1804	.3487	.3060	1.0000

Correlations Between Set-1 and Set-2

	c2	c4	c6	c9
a1	.1057	.0763	.1153	.2284
a3	.0001	.1339	.1365	.3147
a7	.0532	.1501	.0637	.2511
a11	-.0142	.1137	.1161	.2000
a13	-.0168	.1099	.1559	.2543
a14	-.0611	-.0922	-.0712	-.1693

Canonical Correlations

1	.391
2	.200
3	.096
4	.041

Test that remaining correlations are zero:

	Wilk's	Chi-SQ	DF	Sig.
1	.804	58.121	24.000	.000
2	.950	13.808	15.000	.540
3	.989	2.934	8.000	.938
4	.998	.452	3.000	.929

Standardized Canonical Coefficients for Set-1

	1	2	3	4
a1	-.297	.518	-.779	-.034

a3	-.578	-.353	.115	.853
a7	-.019	.844	.818	.032
a11	-.079	-.447	.135	-.829
a13	-.428	-.619	-.225	-.056
a14	-.033	-.477	.192	.489

Raw Canonical Coefficients for Set-1

	1	2	3	4
a1	-.130	.226	-.340	-.015
a3	-.523	-.319	.104	.770
a7	-.017	.757	.734	.028
a11	-.074	-.419	.127	-.777
a13	-.411	-.594	-.216	-.054
a14	-.029	-.422	.170	.432

Standardized Canonical Coefficients for Set-2

	1	2	3	4
c2	.243	.979	-.443	-.168
c4	-.069	.093	.807	-.731
c6	-.293	-.774	-.600	-.527
c9	-.885	.292	-.008	.580

Raw Canonical Coefficients for Set-2

	1	2	3	4
c2	.095	.383	-.174	-.066
c4	-.028	.038	.329	-.298
c6	-.102	-.268	-.208	-.182
c9	-.321	.106	-.003	.210

Canonical Loadings for Set-1

	1	2	3	4
a1	-.551	.441	-.587	-.045
a3	-.837	-.005	.243	.310
a7	-.609	.450	.593	-.160
a11	-.568	-.174	.277	-.629
a13	-.722	-.263	.005	-.295
a14	.415	-.314	-.032	.412

Cross Loadings for Set-1

	1	2	3	4
a1	-.216	.088	-.056	-.002
a3	-.328	-.001	.023	.013
a7	-.238	.090	.057	-.007
a11	-.222	-.035	.027	-.026
a13	-.282	-.053	.000	-.012
a14	.162	-.063	-.003	.017

Canonical Loadings for Set-2

	1	2	3	4
c2	-.057	.727	-.506	-.460
c4	-.399	.219	.536	-.711
c6	-.480	-.246	-.572	-.618
c9	-.955	.264	.009	.134

Cross Loadings for Set-2

	1	2	3	4
c2	-.022	.145	-.049	-.019
c4	-.156	.044	.052	-.029
c6	-.188	-.049	-.055	-.025
c9	-.374	.053	.001	.005

Redundancy Analysis:

Proportion of Variance of Set-1 Explained by Its Own Can. Var.

	Prop Var
CV1-1	.399
CV1-2	.099
CV1-3	.139
CV1-4	.129

Proportion of Variance of Set-1 Explained by Opposite Can.Var.

	Prop Var
CV2-1	.061
CV2-2	.004
CV2-3	.001
CV2-4	.000

Proportion of Variance of Set-2 Explained by Its Own Can. Var.

	Prop Var
CV2-1	.326
CV2-2	.177
CV2-3	.218
CV2-4	.279

Proportion of Variance of Set-2 Explained by Opposite Can. Var.

	Prop Var
CV1-1	.050
CV1-2	.007
CV1-3	.002
CV1-4	.000

----- END MATRIX -----

Create

Created Series

SET_NUM	Series Name	Case Number of Non-Missing Values		N of Valid Cases	Creating Function
		First	Last		
1.00 1	VARSEQ	1	6	6	CSUM(VARSEQ)
2.00 1	VARSEQ	7	10	4	CSUM(VARSEQ)

The canonical scores have been written to the active dataset.
Also, a file containing an SPSS Scoring program has been written

To use this file, GET a system file with the SAME variables that were used in the present analysis. Then use an INSERT command to run the scoring program.

For example :

```
GET FILE anotherfilename
INSERT FILE= 'canonical_correlation_scoring_syntax.sps' .
EXECUTE.
INCLUDE 'C:\Program Files\SPSS\Canonical correlation.sps'.
2057 * Canonical correlation.sps. This version allows long variable names and
      uses datasets.Canonical correlation.sps.
2058
2060 preserve.
2062 set printback=off.
2766 RESTORE.
2767
2769 * End of INSERT and INCLUDE nesting level 01.
CANCORR SET1=A1,A3,A7,A11,A13,A14/
        SET2=C2,C4,C6,C9/.
```

Matrix

Run MATRIX procedure:

Correlations for Set-1

	a1	a3	a7	a11	a13	a14
a1	1.0000	.2570	.2356	.2247	.1989	-.0811
a3	.2570	1.0000	.5199	.4152	.3557	-.3813
a7	.2356	.5199	1.0000	.5120	.4454	-.3630
a11	.2247	.4152	.5120	1.0000	.4245	-.2828
a13	.1989	.3557	.4454	.4245	1.0000	-.4045
a14	-.0811	-.3813	-.3630	-.2828	-.4045	1.0000

Correlations for Set-2

	c2	c4	c6	c9
c2	1.0000	.2385	.4219	.1804
c4	.2385	1.0000	.2700	.3487
c6	.4219	.2700	1.0000	.3060
c9	.1804	.3487	.3060	1.0000

Correlations Between Set-1 and Set-2

	c2	c4	c6	c9
a1	.1057	.0763	.1153	.2284
a3	.0001	.1339	.1365	.3147
a7	.0532	.1501	.0637	.2511
a11	-.0142	.1137	.1161	.2000
a13	-.0168	.1099	.1559	.2543
a14	-.0611	-.0922	-.0712	-.1693

Canonical Correlations

1	.391
2	.200
3	.096
4	.041

Test that remaining correlations are zero:

Wilk's	Chi-SQ	DF	Sig.
--------	--------	----	------

1	.804	58.121	24.000	.000
2	.950	13.808	15.000	.540
3	.989	2.934	8.000	.938
4	.998	.452	3.000	.929

Standardized Canonical Coefficients for Set-1

	1	2	3	4
a1	-.297	.518	-.779	-.034
a3	-.578	-.353	.115	.853
a7	-.019	.844	.818	.032
a11	-.079	-.447	.135	-.829
a13	-.428	-.619	-.225	-.056
a14	-.033	-.477	.192	.489

Raw Canonical Coefficients for Set-1

	1	2	3	4
a1	-.130	.226	-.340	-.015
a3	-.523	-.319	.104	.770
a7	-.017	.757	.734	.028
a11	-.074	-.419	.127	-.777
a13	-.411	-.594	-.216	-.054
a14	-.029	-.422	.170	.432

Standardized Canonical Coefficients for Set-2

	1	2	3	4
c2	.243	.979	-.443	-.168
c4	-.069	.093	.807	-.731
c6	-.293	-.774	-.600	-.527
c9	-.885	.292	-.008	.580

Raw Canonical Coefficients for Set-2

	1	2	3	4
c2	.095	.383	-.174	-.066
c4	-.028	.038	.329	-.298
c6	-.102	-.268	-.208	-.182
c9	-.321	.106	-.003	.210

Canonical Loadings for Set-1

	1	2	3	4
a1	-.551	.441	-.587	-.045
a3	-.837	-.005	.243	.310
a7	-.609	.450	.593	-.160
a11	-.568	-.174	.277	-.629
a13	-.722	-.263	.005	-.295
a14	.415	-.314	-.032	.412

Cross Loadings for Set-1

	1	2	3	4
a1	-.216	.088	-.056	-.002
a3	-.328	-.001	.023	.013
a7	-.238	.090	.057	-.007

a11	-.222	-.035	.027	-.026
a13	-.282	-.053	.000	-.012
a14	.162	-.063	-.003	.017

Canonical Loadings for Set-2

	1	2	3	4
c2	-.057	.727	-.506	-.460
c4	-.399	.219	.536	-.711
c6	-.480	-.246	-.572	-.618
c9	-.955	.264	.009	.134

Cross Loadings for Set-2

	1	2	3	4
c2	-.022	.145	-.049	-.019
c4	-.156	.044	.052	-.029
c6	-.188	-.049	-.055	-.025
c9	-.374	.053	.001	.005

Redundancy Analysis:

Proportion of Variance of Set-1 Explained by Its Own Can. Var.

	Prop Var
CV1-1	.399
CV1-2	.099
CV1-3	.139
CV1-4	.129

Proportion of Variance of Set-1 Explained by Opposite Can.Var.

	Prop Var
CV2-1	.061
CV2-2	.004
CV2-3	.001
CV2-4	.000

Proportion of Variance of Set-2 Explained by Its Own Can. Var.

	Prop Var
CV2-1	.326
CV2-2	.177
CV2-3	.218
CV2-4	.279

Proportion of Variance of Set-2 Explained by Opposite Can. Var.

	Prop Var
CV1-1	.050
CV1-2	.007
CV1-3	.002

CV1-4 .000

----- END MATRIX -----

Create

Created Series

SET_NUM	Series Name	Case Number of Non-Missing Values		N of Valid Cases	Creating Function
		First	Last		
1.00 1	VARSEQ	1	6	6	CSUM(VARSEQ)
2.00 1	VARSEQ	7	10	4	CSUM(VARSEQ)

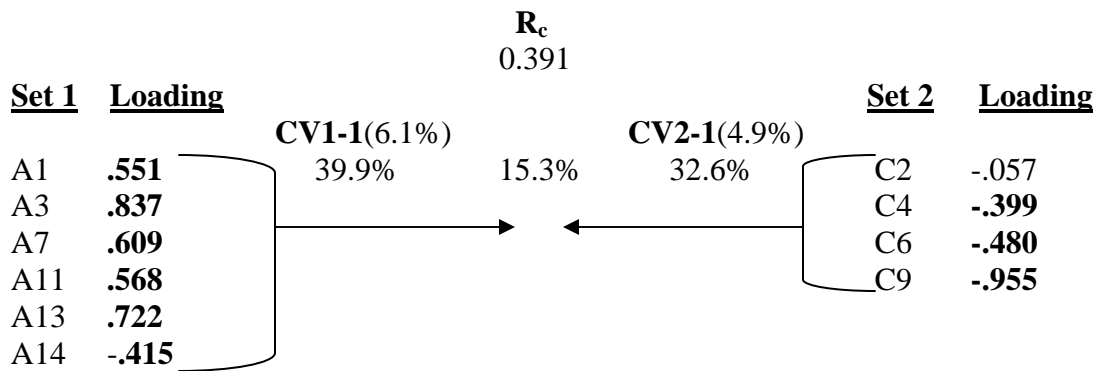
The canonical scores have been written to the active dataset. Also, a file containing an SPSS Scoring program has been written. To use this file, GET a system file with the SAME variables that were used in the present analysis. Then use an INSERT command to run the scoring program.

For example :

```
GET FILE anotherfilename
INSERT FILE= 'canonical_correlation_scoring_syntax_.sps' .
EXECUTE.
SAVE OUTFILE='C:\Documents and Settings\gaoffice\Desktop\SPSS stuff\humgr010504-
2.sav [canonical_correlation_active_data].sav'
/COMPRESSED.
```

Chart 1: Canonical Variates

Canonical Correlation 1



BOLD: Indicates a significance level of $p < .05$, according to Table 3.2 in Hair et. al for a loading of .35 for $n = 250$. Sample $n = 276$.

Wilk's Lambda = .804

Chi² = 58.121

df = 24

Sig = $p < .001$

Note: The numbers in parentheses indicate the redundancy analysis figures.

Note: the signs for Set 1 loadings have been reversed for greater interpretability (since 1=strongly agree and 5= strongly disagree in the original measure).

A canonical correlation was performed in order to determine if there is a relationship between two sets of variables, one measuring communication apprehension and the other measuring neuroticism. One significant canonical correlation function was found from the analysis. In this function, $R_c = .391$, indicating that 15.3% of the variance between the two variates is shared. The Wilk's Lambda statistic was found to be significant through a chi-square test, with $df = 24$ and $p < .001$. The other three, possible canonical correlation functions were not found significant.

Of the six variables included in Set 1, all six variables have significant loadings in CV1-1 (*A1: I dislike participating in group discussions; A3: I am tense and nervous while participating in group discussion; A7: Generally, I am nervous when I have to participate in a meeting; A11: Communicating at meetings usually makes me uncomfortable; A13: While participating in a conversation with a new acquaintance, I feel very nervous; A14: I have no fear of speaking up in conversation*). CV1-1 accounts for 39.9% of the variance in Set 1, while the other variate, CV2-1, shares 6.1 % of its variance with Set 1.

Of the four variables in Set 2, three have significant loadings (*C4: Does your mind often wander when you are trying to concentrate?; C6: Do you sometimes feel happy, sometimes depressed, without any apparent reason?; C9: Are you frequently "lost in thought" even when supposed to be taking part in a conversation?*). CV2-1 accounts for 32.6% of the variance in Set 2. The other variate, CV1-1, shares 4.9% of the variance with Set 2.

There was only one significant relationship found in the canonical correlation. CV1-1 had significant loadings for all of the variables in set 1 (six total variables), while CV2-1 had significant loadings for three out of the four total variables in set 2. Therefore, in Set 1, the canonical variate will remain labeled as “Communication Apprehension,” and Set 2 will be re-named “Neuroticism without moodiness.”

All of the loadings for the canonical variate in Set 1 are positive (see *Note*) except for question A14 (“I have no fear of speaking up in public”). All of the loadings for the canonical variate in Set 2 are negative. Therefore, the more communication apprehension a subject shows, the less neurotic (without moodiness) the respondent will be.