Neuendorf

Controlling for a Third Variable: Correlation, Semi-partial Correlation, and Partial Correlation

SECTION 1: The role of residuals

---

r_{y1}

Zero-order correlation

The Pearson zero-order correlation is calculated between the two sets of raw scores (e.g., X1 & Y)

Community QOL X1	Neighborhood QOL Y
9	9
9	8
9	8
5	7
7	8
9	9
10	5
9	7
9	8
8	10
7	6
8	10
7	7
8	7
9	9
etc.	etc.

$r_{y(1\cdot2)}$ or **sr**

Semi-partial correlation
(or, Part correlation)

The semi-partial correlation is calculated between the raw scores on Y (Neighborhood QOL) and the residuals (what is "left over") for X1 (Community QOL) when the variable X2 (Value of the Family) is regressed on X1:

$$X1' = a_1 + b_1X2$$

Residual Community QOL X1 – X1'	Neighborhood QOL Y'
1.17911	9
1.46289	9
3.16557	8
-2.82089	7
-.82089	8
1.17911	9
2.17911	5
2.88179	7
1.17911	8
.17911	10
-.82089	6
.17911	10
-.82089	7
.17911	7
1.17911	9
etc.	etc.

$r_{y1\cdot2}$ or **pr**

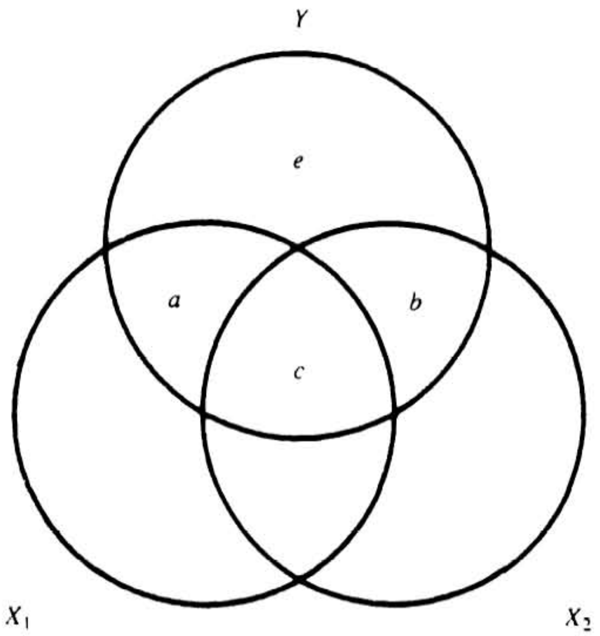
Partial correlation

The partial correlation is calculated between the residuals for X1 (Community QOL) and the residuals for Y (Neighborhood QOL) when the variable X2 (Value of the Family) is regressed on X1 and on Y:

$$X1' = a_1 + b_1X2$$
$$Y' = a_2 + b_2X2$$

Residual Community QOL X1– X1'	Residual Neighborhood QOL Y– Y'
1.17911	1.03026
1.46289	.32073
3.16557	2.06357
-2.82089	-.96974
-.82089	.03026
1.17911	1.03026
2.17911	-2.96974
2.88179	.77310
1.17911	.03026
.17911	2.03026
-.82089	-1.96974
.17911	2.03026
-.82089	-.96974
.17911	-.96974
1.17911	1.03026
etc.	etc.

SECTION 2: Measures of Association with Two Independent Variables—Expressed via Ballantines



$r^2_{Y1} = a + c$	zero-order bivariate correlation^2
$r^2_{Y2} = b + c$	zero-order bivariate correlation^2
$R^2_{Y\bullet 12} = a + b + c$	multiple correlation^2
$sr^2_1 = r^2_{Y(1\bullet 2)} = R^2_{Y\bullet 12} - r^2_{Y2} = a/1$	first order semi-partial correlation^2
$sr^2_2 = r^2_{Y(2\bullet 1)} = R^2_{Y\bullet 12} - r^2_{Y1} = b/1$	first order semi-partial correlation^2
$pr^2_1 = r^2_{Y1\bullet 2} = \dfrac{R^2_{Y\bullet 12} - r^2_{Y2}}{1 - r^2_{Y2}} = \dfrac{a}{a + e}$	first order partial correlation^2
$pr^2_2 = r^2_{Y2\bullet 1} = \dfrac{R^2_{Y\bullet 12} - r^2_{Y1}}{1 - r^2_{Y1}} = \dfrac{b}{b + e}$	first order partial correlation^2

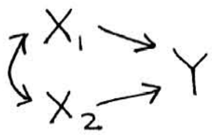
SECTION 3: Patterns of Association when Controlling for a Third Variable: Comparing Correlations (r's) and Partial Correlations (pr's)

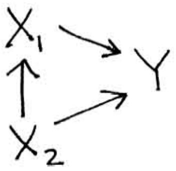
1. Partial Redundancy

   $r_{Y1} > pr_{Y1\cdot 2}$

   (but $pr_{Y1\cdot 2} \neq \emptyset$)

   Model e.g.'s:





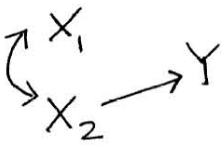
2. Full Redundancy

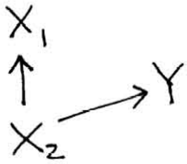
   $pr_{Y1\cdot 2} = \emptyset$

   Could be associated with either of two types of relationships:

   A. Spurious relationship

      Model e.g.'s:





B. Indirect effect

Model e.g.:

$$X_1 \longrightarrow X_2 \longrightarrow Y$$

3. Suppression

$$r_{Y1} < pr_{Y1\cdot 2}$$

Model e.g.:

$$\begin{tikzcd}
X_1 \arrow[dr] \arrow[d, dash, dotted, bend right, "\text{small or } \phi"'] & \\
X_2 \arrow[r] & Y
\end{tikzcd}$$

4. No Effect (control has no impact)

$$r_{Y1} = pr_{Y1\cdot 2}$$

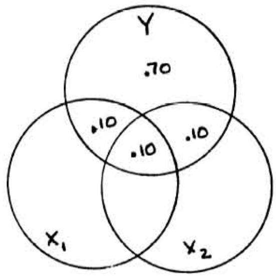
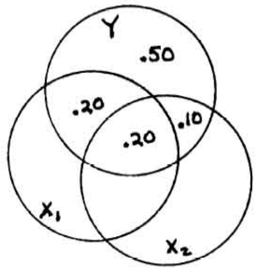
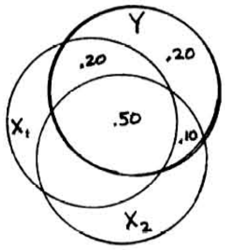
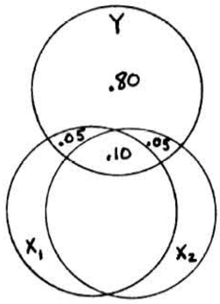
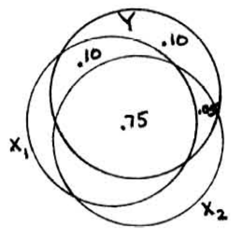
Model e.g.'s:

$$\begin{tikzcd}
X_1 \arrow[dr] \arrow[d, leftrightarrow, bend right] & \\
X_2 & Y
\end{tikzcd}$$

$$\begin{tikzcd}
X_1 \arrow[dr] \arrow[d] & \\
X_2 & Y
\end{tikzcd}$$

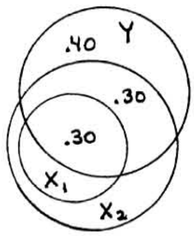
The following pages revisit these four types, with Ballantine examples.

1. Partial Redundancy

		r^2_{Y1}	$pr^2_{Y1\cdot 2}$
e.g. #1		.20	.125
e.g. #2		.40	.29
e.g. #3		.70	.50
e.g. #4		.15	.06
e.g. #5		.85	.50

2. Full Redundancy

e.g. #6



	r^2_{Y1}	$pr^2_{Y1\bullet 2}$
e.g. #6	.30	.00

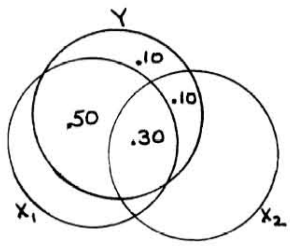
Notice that $pr^2_{Y2\bullet 1}$ is *not* 0.

X1 has no unique contribution to Y.

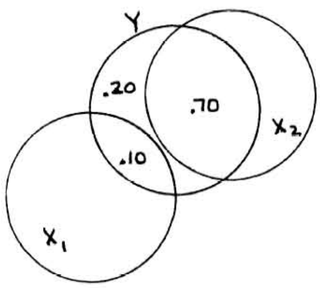
Any of the three models shown under Full Redundancy could produce these numbers—we don't know which is true from correlational findings alone.

3. Suppression

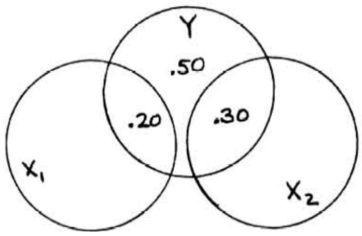
e.g. #7



e.g. #8



e.g. #9

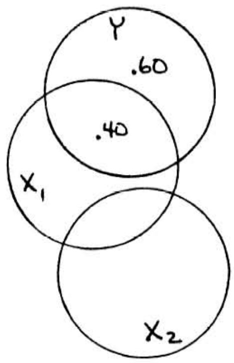


	r^2_{Y1}	$pr^2_{Y1\bullet 2}$
e.g. #7	.80	.83
e.g. #8	.10	.33
e.g. #9	.20	.29

Notice that as the overlap between X1 and X2 in the Y area is larger, redundancy grows. As *unique* contributions of X1 and X2 grow larger (i.e., little or no overlap of X1 and X2 in the Y area), suppression is greater.

4. No Effect

e.g. #10



r^2_{Y1}	$pr^2_{Y1.2}$
.40	.40