Psychology 405: Psychometric Theory
More on Correlations

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

April, 2014
Outline

1. Applied Problems
   - Partial Correlation

2. Multiple Correlation
   - Unit Weighted correlations

3. Other correlations
Predicting scores: Question 1

1. If a person has a GRE verbal score of 680, then what would you expect his/her GRE quantitative score to be?
2. Mean GRE V = 600 SD = 80 r = .72
3. Mean GRE Q = 650 SD = 100
4. Observe GRE V = 680
5. Predicted GRE Q = ?
Predicting scores: Answer 1

1. If a person has a GRE verbal score of 680, then what would you expect his/her GRE quantitative score to be?
2. Mean GRE V = 600 SD = 80 r = .72
3. Mean GRE Q = 650 SD = 100
4. Observe GRE V = 680
5. z GRE V = (680- 600)/80 = 1.0
6. predicted z GRE Q = rxy zx = .72 * (1) = .72
7. predicted GRE Q = .72 * 100 + 650 = 722
Predicting scores: Question 2

1. If a person has a GRE Quant score of 722, then what would you expect his/her GRE Verbal score to be?

2. Mean GRE V = 600 SD = 80 r = .72

3. Mean GRE Q = 650 SD = 100

4. Observe GRE Q = 722

5. Predicted GRE V = ?
Predicting scores: Answer 2

1. If a person has a GRE Quant score of 722, then what would you expect his/her GRE Verbal score to be?

2. Mean GRE V = 600 SD = 80 r = .72

3. Mean GRE Q = 650 SD = 100

4. Observe GRE Q = 722

5. Predicted GRE V = ?

6. z GRE Q = (722- 650)/100 = .72

7. predicted z GRE V = rxy zx = .72 * (.72) = .52

8. predicted GRE Q = .52 * 100 + 600 = 652

1. For a person with an anxiety score of 16, what is the expected GPA?

2. Anxiety Mean = 12 sd = 4 r = -.39

3. GPA Mean = 3.0 sd = .5
Predicting Scores: Answer 3

1. For a person with an anxiety score of 16, what is the expected GPA?
2. Anxiety Mean = 12 sd = 4 r = -.39
3. GPA Mean = 3.0 sd = .5
4. \( z_{anx} = \frac{(16 - 12)}{4} = 1.0 \)
5. predicted \( z_{gpa} = r_{xy} z_x = -.39 \times 1 = -.39 \)
6. predicted \( gpa = -.39 \times .5 + 3 = 2.805 \)
Partial Correlations: Question

1. Assuming the classical model of partial correlations, what is the correlation between GRE Quantitative and GPA with GRE Verbal held constant?

2. \( r_{GREQ,GPA} = .34 \)

3. \( r_{GREQ,V} = .72 \) \( r_{GREV,GPA} = .38 \)

4. We want to find the covariance of Q and GPA without V.

5. All correlations are \( \frac{C_{xy}}{\sqrt{V_x V_y}} \). So we just need to find the Covariances and Variances.
Partial Correlations: Answer

1. Assuming the classical model of partial correlations, what is the correlation between GRE Quantitative and GPA with GRE Verbal held constant?

2. $r_{GREQ,GPA} = .34$

3. $r_{GREQ,V} = .72$ $r_{GREV,GPA} = .38$

4. All correlations are $\frac{c_{xy}}{\sqrt{V_x V_y}}$. So we just need to find the Covariances and Variances.

5. Partial $r_{xy.z} = \frac{r_{xy} - r_{xz} * r_{yx}}{\sqrt{(1-r_{xz}^2)*(1-r_{yx}^2)}}$

6. $r_{qgpa.v} = \frac{(.34-.72*.38)}{\sqrt{(.482*.856)}} = .103$ partial

7. Part $r = \frac{r_{xy} - r_{xz} * r_{yx}}{\sqrt{1-r_{xz}^2}} = .096$ part
Multiple Correlation: Question

1. What is the multiple correlation of GRE V and GRE Q with rated quality of the MA?
2. $r_{GRE\ V,\ MA} = .32$  $r_{GRE\ Q,\ MA} = .29$  $r_{GRE\ V,Q} = .72$
3. $beta_{y,x} = \frac{r_{xy} - r_{xz} \times r_{yz}}{1 - r^2_{xz}}$
4. $R^2 = \sum \beta_i r_{x_i,y}$
What is the multiple correlation of GRE V and GRE Q with rated quality of the MA?

1. \[ r \text{ GRE V, MA} = .32 \]
2. \[ r \text{ GRE Q, MA} = .29 \]
3. \[ r \text{ GRE V,Q} = .72 \]

\[ \beta_{y,x} = \frac{r_{xy} - r_{xz} \cdot r_{yz}}{1 - r_{xz}^2} \]

4. \[ \beta \text{ GRE V, MA} = (0.32 - 0.72 \cdot 0.29)/(1 - 0.72) = 0.231 \]
5. \[ \beta \text{ GRE Q, MA} = (0.29 - 0.72 \cdot 0.32)/(1 - 0.72) = 0.124 \]

\[ R^2 = \beta_{y,x} \cdot r_{xy} + \beta_{y,z} \cdot r_{yz} \ldots \]

6. \[ R^2 = \beta \text{ GRE Q, MA} \cdot r \text{ GRE Q, MA} + \beta \text{ GRE V, MA} \cdot r \text{ GRE V, MA} = \]
7. \[ R^2 = 0.124 \cdot 0.29 + 0.231 \cdot 0.32 = 0.108 \]
8. \[ R = 0.329 \]
What is the unit weighted correlation of GREV and GRE Q with MA?

<table>
<thead>
<tr>
<th>Variable</th>
<th>GREV</th>
<th>GREQ</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GREV</td>
<td>1.00</td>
<td>0.72</td>
<td>0.32</td>
</tr>
<tr>
<td>GREQ</td>
<td>0.72</td>
<td>1.00</td>
<td>0.29</td>
</tr>
<tr>
<td>MA</td>
<td>0.32</td>
<td>0.29</td>
<td>1.00</td>
</tr>
</tbody>
</table>

All correlations are $C_{xy} \frac{1}{\sqrt{V_x V_y}}$. So we just need to find the Covariances and Variances.
Unit Weighted Multiple R

1. Weight the two predictors equally

<table>
<thead>
<tr>
<th>Variable</th>
<th>GREV</th>
<th>GREQ</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GREV</td>
<td>1.00</td>
<td>0.72</td>
<td>0.32</td>
</tr>
<tr>
<td>GREQ</td>
<td>0.72</td>
<td>1.00</td>
<td>0.29</td>
</tr>
<tr>
<td>MA</td>
<td>0.32</td>
<td>0.29</td>
<td>1.00</td>
</tr>
</tbody>
</table>

2. All correlations are \( \frac{C_{xy}}{\sqrt{V_x V_y}} \). So we just need to find the Covariances and Variances.

3. \( C_{v+q, MA} = .32 + .29 = .61 \)

4. \( V_{v+q} = 1.00 + .72 + .72 + 1.00 = 3.44 \)

5. \( r_{v+q, MA} = \frac{.61}{\sqrt{3.44*1}} = .329 \)
Correlating two composites – unit weights

Table: Ability and Performance

<table>
<thead>
<tr>
<th>Variable</th>
<th>GREV</th>
<th>GREQ</th>
<th>GREA</th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GREV</td>
<td>1.00</td>
<td>0.72</td>
<td>0.54</td>
<td>0.38</td>
<td>0.32</td>
<td>0.27</td>
</tr>
<tr>
<td>GREQ</td>
<td>0.72</td>
<td>1.00</td>
<td>0.48</td>
<td>0.34</td>
<td>0.29</td>
<td>0.24</td>
</tr>
<tr>
<td>GREA</td>
<td>0.54</td>
<td>0.48</td>
<td>1.00</td>
<td>0.55</td>
<td>0.47</td>
<td>0.39</td>
</tr>
<tr>
<td>GPA</td>
<td>0.38</td>
<td>0.34</td>
<td>0.55</td>
<td>1.00</td>
<td>0.42</td>
<td>0.35</td>
</tr>
<tr>
<td>Pre</td>
<td>0.32</td>
<td>0.29</td>
<td>0.47</td>
<td>0.42</td>
<td>1.00</td>
<td>0.30</td>
</tr>
<tr>
<td>MA</td>
<td>0.27</td>
<td>0.24</td>
<td>0.39</td>
<td>0.35</td>
<td>0.30</td>
<td>1.00</td>
</tr>
</tbody>
</table>

1. What is the unit weighted correlation between the ability measures and the performance measures?

2. All correlations are $\frac{C_{xy}}{\sqrt{V_x V_y}}$. So we just need to find the Covariances and Variances.
Correlating two composites – unit weights

**Table**: Ability and Performance

<table>
<thead>
<tr>
<th>Hypothetical relationships</th>
<th>Variable</th>
<th>GREV</th>
<th>GREQ</th>
<th>GREA</th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GREV</td>
<td>1.00</td>
<td>0.72</td>
<td>0.54</td>
<td></td>
<td>0.38</td>
<td>0.32</td>
<td>0.27</td>
</tr>
<tr>
<td>GREQ</td>
<td>0.72</td>
<td>1.00</td>
<td>0.48</td>
<td></td>
<td>0.34</td>
<td>0.29</td>
<td>0.24</td>
</tr>
<tr>
<td>GREA</td>
<td>0.54</td>
<td>0.48</td>
<td>1.00</td>
<td></td>
<td>0.55</td>
<td>0.47</td>
<td>0.39</td>
</tr>
<tr>
<td>GPA</td>
<td>0.38</td>
<td>0.34</td>
<td>0.55</td>
<td></td>
<td>1.00</td>
<td>0.42</td>
<td>0.35</td>
</tr>
<tr>
<td>Pre</td>
<td>0.32</td>
<td>0.29</td>
<td>0.47</td>
<td></td>
<td>0.42</td>
<td>1.00</td>
<td>0.30</td>
</tr>
<tr>
<td>MA</td>
<td>0.27</td>
<td>0.24</td>
<td>0.39</td>
<td></td>
<td>0.35</td>
<td>0.30</td>
<td>1.00</td>
</tr>
</tbody>
</table>

1. \( V_{ability} = 1.0 + .72 + .54 + .72 + 1.0 + .48 + .54 + .48 + 1 = 3 \times 1 + 2 \times (0.72 + 0.54 + 0.48) = 6.48 \)
2. \( V_{performance} = 3 \times 1 + 2 \times (0.42 + 0.35 + 0.30) = 5.14 \)
3. \( C_{ability,performance} = \frac{3.25}{\sqrt{6.48 \times 5.14}} = 0.563 \)
Cohen’s Set correlation versus unit weighted correlations

1. What is the relationship between two sets of variables? Two alternative answers.

2. Set correlation: $1 - \frac{|R_{xy}|}{|R_x||R_y|}$

3. Unit weighted correlation

   $R_{uw} = \frac{1R_{yx}1'}{(1R_{yy}1').5(1R_{xx}1').5}$

4. Both are found using the set.cor function
> m1
> set.cor(4:6, 1:3, m1)

GREV GREQ GREA GPA Pre MA
GREV 1.00 0.72 0.54 0.38 0.32 0.27
GREQ 0.72 1.00 0.48 0.34 0.29 0.24
GREA 0.54 0.48 1.00 0.55 0.47 0.39
GPA 0.38 0.34 0.55 1.00 0.42 0.35
Pre 0.32 0.29 0.47 0.42 1.00 0.30
MA 0.27 0.24 0.39 0.35 0.30 1.00
Call: set.cor(y = 4:6, x = 1:3, data = m1)

Multiple Regression from matrix input

Beta weights

<table>
<thead>
<tr>
<th></th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GREV</td>
<td>0.09</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>GREQ</td>
<td>0.05</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>GREA</td>
<td>0.48</td>
<td>0.41</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Multiple R

<table>
<thead>
<tr>
<th></th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>0.56</td>
<td>0.48</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Multiple R2

<table>
<thead>
<tr>
<th></th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>0.31</td>
<td>0.23</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Unweighted multiple R

<table>
<thead>
<tr>
<th></th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>0.50</td>
<td>0.42</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Unweighted multiple R2

<table>
<thead>
<tr>
<th></th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>0.25</td>
<td>0.18</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Various estimates of between set correlations

Squared Canonical Correlations

[1] 4.1e-01 6.0e-05 1.6e-06

Average squared canonical correlation = 0.14
Cohen's Set Correlation R2 = 0.41
Unweighted correlation between the two sets = 0.56
Just use Verbal and Quant

> m1
> set.cor(4:6,1:2,m1)

<table>
<thead>
<tr>
<th></th>
<th>GREV</th>
<th>GREQ</th>
<th>GREA</th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GREV</td>
<td>1.00</td>
<td>0.72</td>
<td>0.54</td>
<td>0.38</td>
<td>0.32</td>
<td>0.27</td>
</tr>
<tr>
<td>GREQ</td>
<td>0.72</td>
<td>1.00</td>
<td>0.48</td>
<td>0.34</td>
<td>0.29</td>
<td>0.24</td>
</tr>
<tr>
<td>GREA</td>
<td>0.54</td>
<td>0.48</td>
<td>1.00</td>
<td>0.55</td>
<td>0.47</td>
<td>0.39</td>
</tr>
<tr>
<td>GPA</td>
<td>0.38</td>
<td>0.34</td>
<td>0.55</td>
<td>1.00</td>
<td>0.42</td>
<td>0.35</td>
</tr>
<tr>
<td>Pre</td>
<td>0.32</td>
<td>0.29</td>
<td>0.47</td>
<td>0.42</td>
<td>1.00</td>
<td>0.30</td>
</tr>
<tr>
<td>MA</td>
<td>0.27</td>
<td>0.24</td>
<td>0.39</td>
<td>0.35</td>
<td>0.30</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Multiple Regression from matrix input

Beta weights

<table>
<thead>
<tr>
<th></th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GREV</td>
<td>0.28</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>GREQ</td>
<td>0.14</td>
<td>0.12</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Multiple R

<table>
<thead>
<tr>
<th></th>
<th>GPA</th>
<th>Pre</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>0.39</td>
<td>0.33</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Various estimates of between set correlations

Squared Canonical Correlations

[1] 2.0e-01 4.2e-05

Average squared canonical correlation = 0.1

Cohen's Set Correlation R2 = 0.2

Unweighted correlation between the two sets = 0.44
Comorbidity

1. Symptoms are said to be comorbid if one has both symptoms.
   - This is really just one cell in a 2 x 2 table
   - We need base rates as well

2. Consider Anxiety and Depression
   - 50% of anxiety patients are also depressed
   - 67% of depressed patients are also anxious
   - base rates are 20% for anxiety, 15% for depression

\[
\text{comorbidity}(0.2, 0.15, 0.1, \text{c("Anxiety","Depression")})
\]

Call: comorbidity(d1 = 0.2, d2 = 0.15, com = 0.1, labels = c("Anxiety", "Depression"))

Comorbidity table

<table>
<thead>
<tr>
<th></th>
<th>Anxiety</th>
<th>-Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>0.1</td>
<td>0.05</td>
</tr>
<tr>
<td>-Depression</td>
<td>0.1</td>
<td>0.75</td>
</tr>
</tbody>
</table>

implies phi = 0.49 with Yule = 0.87 and tetrachoric correlation of 0.75
and normal thresholds of 1.04 0.84