**High Redundancy and Unique Effects in Multiple Regression**

 We have (simulated) data on three variables (the data are available should you wish to play with them). The variables are individual persons’ wealth, IQ, and [SAp](https://www.psychologicalscience.org/pdf/ps/Frey.pdf?origin=publication_detail). SAp is a measure of scholastic aptitude. Here are the descriptive statistics.

|  |
| --- |
| **Descriptive Statistics** |
|  | N | Minimum | Maximum | Mean | Std. Deviation |
| Wealth | 100 | 17534 | 77794 | 40407.34 | 11436.455 |
| IQ | 100 | 60 | 133 | 99.46 | 15.236 |
| SAp | 100 | 419 | 1547 | 987.22 | 252.749 |
| Valid N (listwise) | 100 |  |  |  |  |

|  |
| --- |
| **Correlations** |
|  | Wealth | IQ | SAp |
| Wealth | Pearson Correlation | 1 | .316\*\* | .335\*\* |
| Sig. (2-tailed) |  | .001 | .001 |
| N | 100 | 100 | 100 |
| IQ | Pearson Correlation | .316\*\* | 1 | .834\*\* |
| Sig. (2-tailed) | .001 |  | .000 |
| N | 100 | 100 | 100 |
| SAp | Pearson Correlation | .335\*\* | .834\*\* | 1 |
| Sig. (2-tailed) | .001 | .000 |  |
| N | 100 | 100 | 100 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). |

 Both IQ and SAp have medium-sized correlations with wealth and they are strongly correlated with each other.

|  |
| --- |
| **Variables Entered/Removeda** |
| Model | Variables Entered | Variables Removed | Method |
| 1 | SAp, IQb | . | Enter |
| a. Dependent Variable: Wealth |
| b. All requested variables entered. |

|  |
| --- |
| **Model Summary** |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .342a | .117 | .098 | 10859.003 |
| a. Predictors: (Constant), SAp, IQ |

 Notice that the multiple *R* is not much higher than the zero-order correlations for predicting wealth.

|  |
| --- |
| **ANOVAa** |
| Model | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 1510415867.238 | 2 | 755207933.619 | 6.405 | .002b |
| Residual | 11438041371.202 | 97 | 117917952.280 |  |  |
| Total | 12948457238.440 | 99 |  |  |  |
| a. Dependent Variable: Wealth |
| b. Predictors: (Constant), SAp, IQ |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Standardized Coefficients | t | Sig. | Correlations | Collinearity Statistics |
| Beta | Zero-order | Partial | Part | Tolerance | VIF |
| 1 | (Constant) |  | 2.681 | .009 |  |  |  |  |  |
| IQ | .119 | .688 | .493 | .316 | .070 | .066 | .304 | 3.291 |
| SAp | .236 | 1.362 | .176 | .335 | .137 | .130 | .304 | 3.291 |

 Neither intelligence nor scholastic aptitude have significant unique effects on wealth. Would it make sense for us to conclude that neither intelligence nor scholastic aptitude are related to wealth?

 Here I conduct a sequential analysis, entering IQ first and SAp second. My theoretical reason for doing so that I believe that high intelligence causes the development of scholastic aptitude.

|  |
| --- |
| **Variables Entered/Removeda** |
| Model | Variables Entered | Variables Removed | Method |
| 1 | IQb | . | Enter |
| 2 | SApb | . | Enter |
| a. Dependent Variable: Wealth |
| b. All requested variables entered. |

|  |
| --- |
| **Model Summary** |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .316a | .100 | .091 | 10906.294 | .100 | 10.859 | 1 | 98 | .001 |
| 2 | .342b | .117 | .098 | 10859.003 | .017 | 1.855 | 1 | 97 | .176 |
| a. Predictors: (Constant), IQ |
| b. Predictors: (Constant), IQ, SAp |

 Adding SAp to the model that already has IQ does not significantly increase the *R*, so I conclude that it is really IQ, not SAp, that is the ultimate cause of variance in wealth.

|  |
| --- |
| **ANOVAa** |
| Model | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 1291626310.409 | 1 | 1291626310.409 | 10.859 | .001b |
| Residual | 11656830928.031 | 98 | 118947254.368 |  |  |
| Total | 12948457238.440 | 99 |  |  |  |
| 2 | Regression | 1510415867.238 | 2 | 755207933.619 | 6.405 | .002c |
| Residual | 11438041371.202 | 97 | 117917952.280 |  |  |
| Total | 12948457238.440 | 99 |  |  |  |
| a. Dependent Variable: Wealth |
| b. Predictors: (Constant), IQ |
| c. Predictors: (Constant), IQ, SAp |

|  |
| --- |
| **Coefficientsa** |
| Model | Unstandardized Coefficients | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 16828.901 | 7237.875 |  | 2.325 | .022 |
| IQ | 237.065 | 71.941 | .316 | 3.295 | .001 |
| 2 | (Constant) | 20983.205 | 7825.276 |  | 2.681 | .009 |
| IQ | 89.388 | 129.940 | .119 | .688 | .493 |
| SAp | 10.670 | 7.833 | .236 | 1.362 | .176 |
| a. Dependent Variable: Wealth |

 Now my friend uses the same data for a sequential analysis, but she enters SAp first. I asked her why and she told me “if you don’t know I am not going to tell you.”

|  |
| --- |
| **Variables Entered/Removeda** |
| Model | Variables Entered | Variables Removed | Method |
| 1 | SApb | . | Enter |
| 2 | IQb | . | Enter |
| a. Dependent Variable: Wealth |
| b. All requested variables entered. |

|  |
| --- |
| **Model Summary** |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .335a | .112 | .103 | 10829.780 | .112 | 12.402 | 1 | 98 | .001 |
| 2 | .342b | .117 | .098 | 10859.003 | .004 | .473 | 1 | 97 | .493 |
| a. Predictors: (Constant), SAp |
| b. Predictors: (Constant), SAp, IQ |

 Oh my, adding IQ to a model that already has SAp in it does not significantly increase *R*. Does that mean that it is actually SAp that is causing wealth, not IQ?

|  |
| --- |
| **ANOVAa** |
| Model | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 1454612943.245 | 1 | 1454612943.245 | 12.402 | .001b |
| Residual | 11493844295.195 | 98 | 117284125.461 |  |  |
| Total | 12948457238.440 | 99 |  |  |  |
| 2 | Regression | 1510415867.238 | 2 | 755207933.619 | 6.405 | .002c |
| Residual | 11438041371.202 | 97 | 117917952.280 |  |  |
| Total | 12948457238.440 | 99 |  |  |  |
| a. Dependent Variable: Wealth |
| b. Predictors: (Constant), SAp |
| c. Predictors: (Constant), SAp, IQ |

|  |
| --- |
| **Coefficientsa** |
| Model | Unstandardized Coefficients | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 25435.286 | 4387.124 |  | 5.798 | .000 |
| SAp | 15.166 | 4.306 | .335 | 3.522 | .001 |
| 2 | (Constant) | 20983.205 | 7825.276 |  | 2.681 | .009 |
| SAp | 10.670 | 7.833 | .236 | 1.362 | .176 |
| IQ | 89.388 | 129.940 | .119 | .688 | .493 |
| a. Dependent Variable: Wealth |

 It should be clear that my friend was trying to show me that testing causal models with sequential analysis is tricky business when no variables are manipulated.

 So, what should we conclude here? I opine that we should conclude that both IQ and SAp are well correlated with wealth, but if you simply want to predict wealth you do not need both of them. You can do quite well with just one of them. I’d use whichever one is less expensive (in terms of gathering the data).

Karl L. Wuensch, January, 2021.