Partial and Semipartial Correlation Example

This SPSS output can be obtained by adding ZPP to the STATISTICS subcommand (or with menus by checking the “Part and Partial Correlations” box on the regression *Statistics* option) for the simultaneous regression of SALARY regressed on TIME and PUBS. There also is an SPSS procedure called PARTIAL CORR which will also produce partial correlations, but I find it less convenient because you can only get partial correlations for one pair of variables per command (specifying which variables to control for with the BY keyword).

```
regression vars=salary time pubs
    /descriptives=mean stdev
    /statistics=anova coeff ses r ci zpp
    /dependent=salary
          /method=enter pubs time.
```

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients				95.0% Confidence Interval for B		Correlations		
		B	Std. Error	Beta	Std. Error	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
1	(Constant)	43082.394	3099.493			13.900	.000	36329.178	49835.610			
	time years since PhD	982.867	452.057	.570	.262	2.174	.050	-2.081	1967.815	.710	.532	.430
	pubs number of publications	121.801	149.699	.213	.262	.814	.432	-204.364	447.966	.588	.229	.161

a. Dependent Variable: salary annual salary in dollars

“Zero-order” is simply the usual Pearson correlation coefficient (nothing is partialed out) and is the same as what was obtained earlier with the `correlations` procedure.

“Part” refers to the semi-partial correlation coefficient (sr = .161). The squared semi-partial coefficient for PUBS (sr^2) equals the R-square change value from the hierarchical regression when PUBS is added to the model already including TIME (or the R-square change when a single variable is added):

$$sr^2 = R^2_{change} = (.161)^2 = .026$$

Note that the partial correlation coefficient (pr) has no direct relationship to the R-square change value ($pr^2 \neq R^2_{change}$).

The other semi-partial coefficient, .430 for TIME, bears the same relationship to R-square change. Looking at the hierarchical regression in which TIME is added second to the model already including PUBS, you see that:

$$sr^2 = R^2_{change} = (.430)^2 = .185 .$$

R code

I used the ppcor package to obtain the semi-partial and partial correlation coefficients. Each function outputs the correlations in the first block, the *p*-values in the second bloc, and the *t*-value in the third block. For the spcor and the pcor functions specified below, d is the data frame name.

```
#The ppcor package can be used to obtain the partial and semi-partial correlation coefficients

library(ppcor)
#semi-partial (same as R-square change)
spcor(d[,c("SALARY","PUBS","TIME")])


$estimate
          SALARY      PUBS      TIME
SALARY 1.0000000 0.1609384 0.4300597
PUBS   0.1724497 1.0000000 0.3396850
TIME   0.4009327 0.2955400 1.0000000

$p.value
         SALARY     PUBS      TIME
SALARY 0.000000 0.582568 0.1248191
PUBS   0.555503 0.000000 0.2347422
TIME   0.155396 0.304957 0.0000000

$statistic
          SALARY      PUBS     TIME
SALARY 0.0000000 0.5648703 1.650166
PUBS   0.6064693 0.0000000 1.251094
TIME   1.5160576 1.0716508 0.000000


#partial correlation
pcor[d,c("SALARY","PUBS","TIME")])

$estimate
          SALARY      PUBS      TIME
SALARY 1.0000000 0.2286551 0.5316060
PUBS   0.2286551 1.0000000 0.4198919
TIME   0.5316060 0.4198919 1.0000000

$p.value
           SALARY      PUBS       TIME
SALARY 0.00000000 0.4317001 0.05041374
PUBS   0.43170006 0.0000000 0.13498636
TIME   0.05041374 0.1349864 0.00000000

$statistic
          SALARY      PUBS     TIME
SALARY 0.0000000 0.8136399 2.174209
PUBS   0.8136399 0.0000000 1.602677
TIME   2.1742093 1.6026773 0.000000
```