Multiple Linear Regression: Squared Semi-partial Correlation ΔR^2

1. Purpose of Squared Semi-partial (or Part) Correlation ΔR^2

The squared semi-partial correlation, or the squared part correlation, is mathematically equivalent to ΔR^2 — the change in model R^2 between full (all relevant predictors included) and reduced models (predictors of interest omitted). It is the increase in model R^2 from the addition of a variable or set of variables to the regression equation.

The squared semi-partial correlation

- offers a standardized measure of partial effect upon the DV for each predictor,
- it is a measure of effect size (measure for determining relative effect of a predictor), and
- serves for hypothesis testing the combined statistical effect of a set of variables or vectors in regression.

2. Calculating ΔR^2

The squared semi-partial correlation is found comparing the change in model R^2 between two regression models, the reduced and full model:

$$\Delta R^2(X) = R_f^2 - R_r^2$$

where f = full and r = reduced and X indicates the predictor or predictors for which one may calculate the squared semi-partial correlation.

Table 1: Example Calculating ΔR^2 for Variable X_1

Model	Equation	R^2 Values
Full Model	$Y = b_0 + b_1X_1 + b_2X_2 + e.$	$R_f^2 = .40$
Reduced Model (X_1 omitted)	$Y = b_0 \qquad + b_2X_2 + e.$	$R_r^2 = .30$
		$\Delta R^2(X_1) = .4 - .3$
		$= .10$

Recall the fictional mathematics scores data in Table 2 below.

SPSS Data File: http://www.bwgriffin.com/gsu/courses/edur8132/notes/fictional_math_scores.sav

Table 2: Fictional Mathematics Scores, Height, Sex, and Other Mathematic Scores

Math Scores	Height	Sex	Other Math	Math Scores	Height	Sex	Other Math
9	11	1	.	3	5	0	.
8	10	1	.	2	4	0	.
9	10	1	10	3	4	0	3
10	10	1	11	4	4	0	2
7	9	1	.	1	3	0	.
8	9	1	12	2	3	0	4
9	9	1	.	3	3	0	3
10	9	1	11	4	3	0	.
11	9	1	.	5	3	0	.
8	8	1	.	2	2	0	5
9	8	1	12	3	2	0	4
10	8	1	13	4	2	0	.
9	7	1	.	3	1	0	.

What is the partial effect of sex for the model of "Other Math" scores? The partial effect, as measured by ΔR^2(sex), can be calculated as follows. The full model contains both sex and height as predictors:

Full Model: Other $Math_i = b_0 + b_1 Height_{i1} + b_2 Sex_{i2} + e_i$

and the reduced model omits sex:

Reduced Model: Other $Math_i = b_0 + b_1 Height_{i1} + \ldots + e_i$

1. What values of R^2 are obtained in both full and reduced models?
2. What is the value obtained for ΔR^2(sex)?
3. What is the value obtained for ΔR^2(height)?

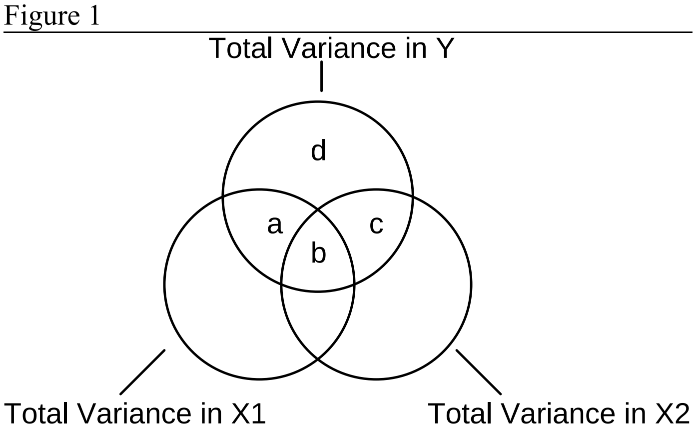
Note that ΔR^2(sex) represents the increase in model R^2 that is attributable to the variable sex over and above the contribution of height. It is the partial effect of sex on model R^2; it is the unique contribution of sex to the regression model predicting Other Math. Similarly, ΔR^2(height) is the partial effect of height on model R^2 taking into account sex—the increment in model R^2 due to adding height to a model that already contains sex.

The squared semi-partial correlation, ΔR^2, represents predicted variance in Y attributable, uniquely, to a given X or set of Xs.

3. Graphical Illustration ΔR^2

Below in Figure 1 is a Venn Diagram demonstrating semi-partial, or part, correlations.

Figure 1



Note the following:

Y = DV : The total variance of Y is the sum of a + b + c + d
X1 = IV
X2 = IV
a = The variance of Y predicted, uniquely, by X1 : $\Delta R^2(X1) = a$
c = The variance of Y predicted, uniquely, by X2 : $\Delta R^2(X2) = c$

1. What is b?
2. What is d?
3. How could d be calculated in terms of model R^2?

4. Finding ΔR^2 for House prices in Albuquerque 1993

Recall the selling prices of homes in Albuquerque for 1993:

Price	=	Prices in thousands of dollars.
Square Feet	=	Size of house in square feet living space.
Age	=	Age of house in years.
Features	=	Number out of 11 features (dishwasher, refrigerator, microwave, disposer, washer, intercom, skylight(s), compactor, dryer, handicap fit, cable TV access)
Tax	=	Annual taxes in dollars

The data can be downloaded here:

Excel: http://www.bwgriffin.com/gsu/courses/edur8131/data/house-prices.xls
SPSS: http://www.bwgriffin.com/gsu/courses/edur8131/data/house-prices.sav
Minitab: http://www.bwgriffin.com/gsu/courses/edur8131/data/house-prices.MTW

Run the following regression model:

Price = square feet + age + number of features

1. What value of of R^2 is obtained for this full model?
2. What is the value obtained for ΔR^2(square feet)?
3. What is the value obtained for ΔR^2(age)?
4. What is the value obtained for ΔR^2(features)?

5. $\Delta R^2(X_1,X_2)$: Squared Semi-partial for Sets of Predictors

It is possible to calculate ΔR^2 for sets of predictors to measure the combined contribution of several variables in terms of Y variance predicted. The value of ΔR^2 for several predictors is calculated as the change in model R^2:

$$\Delta R^2(X_1, X_2) = R_f^2 - R_r^2$$

where f = full and r = reduced and X indicates the predictor or predictors for which one may calculate the squared semi-partial correlation.

Table 1: Example Calculating ΔR^2 for Variable X_1 and X_2

Model	Equation	R^2 Values
Full Model	$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + e.$	$R_f^2 = .17$
Reduced Model (X_1, X_2 omitted)	$Y = b_0 \quad + \quad + b_3X_3 + e.$	$R_r^2 = .05$
		$\Delta R^2(X_1, X_2) = .17 - .05$
		$= .12$

1. What is the value obtained for ΔR^2(square feet, age)?
2. What is the value obtained for ΔR^2(square feet, features)?
3. What is the value obtained for ΔR^2(age, features)?

6. Hypothesis Testing ΔR^2 for One Predictor

The null hypothesis indicates that the partial effects of a predictor equals zero:

H_0: $\Delta R^2(X) = 0.00$

The alternative hypothesis:

H_1: $\Delta R^2(X) \neq 0.00$

If fail to reject H_0, then one concludes the partial effects of X controlling for other predictors does not statistically contribute to model R^2 (X does not contribute to explained variation in Y), or that X is not associated with Y once other predictors are taken into account.

Note the following equivalence when only one predictor with a single degree of freedom is considered—that is, a variable that consumes only one column of data (a single vector):

H_0: $\Delta R^2(X) = 0.00$ is the same as

H_0:$\beta_x = 0.00$.

For tests of single degree of freedom predictors (one vector, one column IVs), both t-test and partial F-test are equivalent.

Partial F-test is often used for testing H_0: $\Delta R^2(X)$

$$F = \frac{\Delta R^2(X)/(df_{2r} - df_{2f})}{(1-R_f^2)/df_{2f}}$$

where

$\Delta R^2(X)$ is the partial effect to be tested;
df_{2f} is the error degrees of freedom for the *full* model ($n - k_f - 1$);
df_{2r} is the error degrees of freedom for the *reduced* model ($n - k_r - 1$); and
R_f^2 is the *full* model R^2 value.

The partial F-ratio is test against critical F value with degrees of freedom equal to

$df_1 = df_{2r} - df_{2f}$,

and

$df_2 = df_{2f}$.

An example with the Other Math data:

Full Model: Other $Math_i = b_0 + b_1 Height_{i1} + b_2 Sex_{i2} + e_i$

$R_f^2 = .985$
$df_{2f} = 9$

Reduced Model: Other $Math_i = b_0 + b_1 Height_{i1} + \ldots + e_i$

$R_f^2 = .785$
$df_{2f} = 10$

So the following partial F-ratio is obtained:

ΔR^2(sex) = .985 - .785 = .20, so

$$F = \frac{\Delta R^2(X)/(df_{2r} - df_{2f})}{(1-R_f^2)/df_{2f}} = \frac{.20/(10-9)}{(1-.985)/9} = \frac{.20/(1)}{(.015)/9} = \frac{.20}{.001666667} = 119.9999 \approx 120$$

This calculated, or obtained, F-ratio is compared against a critical F value with the following degrees of freedom:

df_1 (df_b) = df_{2r} - df_{2f} = 10 – 9 = 1

and

df_2 (df_w)= df_{2f} = 9

Using the table of critical F-ratios linked on the course web site, the critical F value for α=.05 is

Critical ${}_{.05}F_{1,9} = 5.12$

Decision rule:

If $F \geq$ critical F reject H_0, otherwise fail to reject H_0

Since F = 120 is greater than critical F = 5.12, H_0 is rejected and we conclude that the partial effect of sex is statistically significant; sex does contribute to variance predicted in other math scores.

What is the partial F-test result for ΔR^2(height)?

When only one predictor is examined—variables that occupy only one column of data, or a vector—the following equivalence results for the F-ratio with one degree of freedom (df_1 = 1) to t-ratio:

$F_{(df1 = 1,\ df2 = \infty)} = t^2$

For sex the partial F-ratio = 120, and from the regression results the obtained t-ratio was 11.03:

$120 \approx 11.03^2$ (= 121.6609)

Note discrepancy is due to round error of reported results in SPSS for the R^2 values for the full and reduced models. SPSS provides only the thousandths place, but the actual R^2 values for the full and reduced models are:

$R_f^2 = .985222$
$R_r^2 = .785459$

Using these figures, the partial F-ratio would be:

ΔR^2(sex) = .985222 - .785459 = .199763, so

$$F = \frac{\Delta R^2(X)/(df_{2r} - df_{2f})}{(1-R_f^2)/df_{2f}} = \frac{.199763/(10-9)}{(1-.985222)/9} = \frac{.199763/(1)}{(.014778)/9} = \frac{.199763}{.001642} = 121.658$$

The above value is, again, within rounding error of the value reported in SPSS.

7. Partial F test with SPSS

SPSS can be used to calculate ΔR^2 values and partial F-tests.

a. Linear Regression →Statistics→check both "R squared change"
b. Linear Regression→Add Height to "Independents" Box→Next→Add Sex to "Independents" Box→OK

Using the Albuquerque housing data, what is the partial F-test result for ΔR^2(size)?
Using the Albuquerque housing data, what is the partial F-test result for ΔR^2(features)?

8. Hypothesis Testing ΔR^2 for a Set of Predictors

The partial F test can also be calculated for a set of predictors. Using the Albuquerque Housing Data, the following figures are obtained for the set size and features:

Full Model: Price = $b_0 + b_1$Size+ b_2Features + b_3Age + e_i

$R_f^2 = .80$
$df_{2f} = 62$

Reduced Model: Price = $b_0 + b_3$Age + e_i

$R_f^2 = .028$
$df_{2f} = 64$

So the following partial F-ratio is obtained:

ΔR^2(size, features) = .80 - .028 = .772, so

$$F = \frac{\Delta R^2(X)/(df_{2r} - df_{2f})}{(1-R_f^2)/df_{2f}} = \frac{.772/(64-62)}{(1-.8)/62} = \frac{.772/(2)}{(.20)/62} = \frac{.386}{.0032258} = 119.66$$

This calculated, or obtained, F-ratio is compared against a critical F value with the following degrees of freedom:

df_1 (df_b) = df_{2r} - df_{2f} = 64 – 62 = 2

and

df_2 (df_w)= df_{2f} = 64

Using the table of critical F-ratios linked on the course web site, the critical F value for α=.05 (using tabled df_2 = 60) is

Critical $_{.05}F_{2,64(60)} = 3.15$

Since F = 119.66 is greater than critical F = 3.15, H_0 is rejected and we conclude that the partial effect for combined size and features is statistically significant; together, these two predictors contribute to variance predicted in selling price of houses.

Housing Sales Data

1. What is the partial F-ratio for ΔR^2(square feet, age)?
2. What is the partial F-ratio for ΔR^2(square feet, features)?
3. What is the partial F-ratio for ΔR^2(age, features)?

9. Ice Cream Sales Data

Ice Cream Sales Data:
Excel: http://www.bwgriffin.com/gsu/courses/edur8131/data/ice-cream.xls
SPSS: http://www.bwgriffin.com/gsu/courses/edur8131/data/ice-cream.sav

Sales (consumption)	=	Measured in pints per capita.
Price	=	Price of ice cream in dollars.
Income	=	Weekly family income in dollars.
Temperature	=	Mean temperature in degrees Fahrenheit.

Assume the following full model:

Sales' = $b_0 + b_1$Price + b_2Income + b_3Temperature

1. What is the partial F-ratio and ΔR^2(income)?
2. What is the partial F-ratio and ΔR^2(temperature)?
3. What is the partial F-ratio and ΔR^2(price)?
4. What is the partial F-ratio and ΔR^2(temperature, income)?
5. What is the partial F-ratio and ΔR^2(temperature, price)?

Answers

	F-ratio	ΔR^2
1. What is the partial F-ratio and ΔR^2(income)?	7.97	.086
2. What is the partial F-ratio and ΔR^2(temperature)?	60.25	.651
3. What is the partial F-ratio and ΔR^2(price)?	1.57	.017
4. What is the partial F-ratio and ΔR^2(temperature, income)?	30.14	.652
5. What is the partial F-ratio and ΔR^2(temperature, price)?	33.15	.717