List of Symbols and Concepts

Regression model equations

Y	Outcome variable (AKA dependent variable)
Y_i	Score on outcome variable for a specific subject. Subscript indicates subject number. If left as i, is a general symbol for scores on the outcome variable.
$\hat{Y}_i$	Predicted value for outcome variable
$\bar{Y}$	Mean of Y (can substitute the symbol for any variable (i.e., X1)
X_k	A regressor in a linear model. Number denotes specific regressor (i.e., X1, X2)
e	Residual or error in regression model (i.e. $Y - \hat{Y}_i$). May have a subscript to indicate the error for a specific observation.
b_0	Sample parameter estimate of β_0. Y intercept from general linear model estimated in a sample. Describes the predicted value for Y when all Xs = 0.
b_i	Sample parameter estimate of β_i. Often referred to as a regression coefficient. Subscript indicates associated predictor variable. (i.e., b_1 is regression coefficient for X1). Describes the change in predicted value for Y associated with a one unit increase in Xi, controlling for (holding constant) all other Xs.
β_0	Population parameter estimated by b_0
β_i	Population parameter estimated by b_i. Subscript indicates associated predictor variable.
SE	Standard error. Can have standard error for any parameter estimate (e.g., b_0, b_i). The specific parameter will be indicated in the subscript (i.e., SE_{b1} is the standard error for the raw score regression coefficient for X1)

Other regression components

R^2	Coefficient of Determination. Proportion of variance in outcome variable explained by predictor variables. By default the outcome variable is Y and all predictor variables are included. Subscripts will be used to indicate other combinations.
adj R^2	Adjusted R^2 (AKA shrunken R^2) A correction for the positive bias in R^2
PRE	Variance based effect size estimate for a predictor. Describes the proportional reduction in SSE associated with that specific X

Partial Eta^2 More common name for PRE.

Delta R^2 Variance based effect size estimate for a predictor. The increase in model R^2 associated with X (comparing compact model without X to augmented model that includes X)

SEE Standard error of estimate. This is the standard deviation of the residuals. Other frequently observed symbols include: $s_{y.x}$ or σ

M_i Maha distance (a measure of leverage) for a specific observation (i)

D_i Cook's distance (a measure of influence) for a specific observation (i)

CI Confidence for a parameter estimate. Can be formed for any parameter estimates. We have discussed CIs for b's and β's. Again can be formed for various confidence levels (95% most common).

Sums of Squares

SSE Sums of squares error. In ANOVA terms also Sums of squares within. This is the sum of the squared model residuals (errors). This is the variability in Y that can NOT be explained by the model

SSR Sums of squares regression (also SSR). (In ANOVA sums of squares between). This is the variability in Y that can be explained by the model or a specific X in the model.

Test statistics/distributions

t	t-statistic. Used by t-test and some tests of differences in correlations (and many other places as well). Used in the test of hypotheses about paramters in the GLM
F	F-statistic. Used in test of R^2. Can also be used in the test of hypotheses about parameters in the GLM. For single df F tests, $F = t^2$
z	z-statistic. Used in z-test for some differences in correlations.
χ^2	Chi-square statistic. Used in various chi-square tests. We used in test for all r's <>0
df	Degrees of freedom for a specific test statistic. F statistic has both a numerator and denominator degree of freedom (ndf, ddf, respectively)

Other symbols (that didn't fit elsewhere :-)

N or n	Sample size. Little n is often used for cell size (in ANOVA terms) and big N for total sample. However, this is far from consistent in the literature)
N*	Required sample size determined from a power analysis
k	Number of regressors in a model
$\sum$	Sum of
α	Alpha. Probability of a type I (false alarm) error.
β	Beta. Probability of type II error (miss). Don't confuse with standardized regression coefficient. Same symbol.
ln	Natural log (base e)