Regression analysis Module 10: adjusted R^2 practice problems

(The attached PDF file has better formatting.)

** Exercise 10.1: Adjusted R^2

A statistician regresses the response variable Y on *k* explanatory variables $X_1, X_2, \ldots, X_k$ and one intercept.

The variance of the observed Y values is 10 and the estimated σ^2_ε (the error variance) is 2.

A. What is the total sum of squares (TSS)?
B. What is the residual sum of squares (RSS)?
C. What is the regression sum of squares?
D. What is the R^2 of the regression?
E. What is the adjusted R^2?

Part A: The total sum of squares TSS is the observed variance × (N-1), where N = the observations.

Part B: The residual sum of squares RSS is $\sigma^2_\varepsilon \times (N-k-1)$.

Note: These relations are generally written in the reverse form: variance = TSS / (N-1); σ^2_ε = RSS / (N-k-1).

Part C: The regression sum of squares RegSS is TSS – RSS.

Part D: The R^2 is ResSS / TSS.

Part E: The adjusted R^2 = [RegSS / (N-k-1)] / [TSS / (N-1)] = 1 – σ^2_ε / the variance of the response variable:

$$\text{adjusted } R^2 = 1 - 2 / 10 = 80\%.$$

The general formula is that the adjusted $R^2 = (\text{var}(Y) - \sigma^2_\varepsilon) / \text{var}(Y)$.

- If we are given RSS and TSS (or RSS and RegSS, or RegSS and TSS), but not N (the number of observations) or k (the number of parameters), we can derive R^2 but not the adjusted R^2.
- If we are given the variance of the Y values (the response variable) and the σ^2_ε, but not N (the number of observations) or k (the number of parameters), we can derive the adjusted R^2 but not the simple R^2.