$Var(A\epsilon) = A\,Var(\epsilon)\,A^T$

ϵ is assumed to follow multivariate norm. dist. with mean vector 0 + covariance matrix $\sigma^2 I$

$\hat{\beta} = (X^TX)^{-1}X^Ty$

$Var(\hat{\beta}) = (X^TX)^{-1}\sigma$

$\hat{\sigma}^2 = \frac{1}{N-p-1}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$

$\vec{a}\cdot\vec{b} = |\vec{a}||\vec{b}|\cos\theta$

$Z_j = \frac{\hat{\beta}_j}{\hat{\sigma}\sqrt{v_j}}$ v_j = j^{th} diagonal of the cov-var matrix XX^T

$F = \frac{\overbrace{(RSS_0 - RSS_1)/(p_1 - p_0)}^{\text{avg reduction in RSS per addit. param}}}{RSS_1/(N-p_1-1)}$

RSS_0 = RSS of model with fewer parameters p_0+1, a subset of larger model

RSS_1 = RSS of larger model with p_1+1 params

derivation: sub $y = X\beta + \epsilon$

$\hat{\beta} = (X^TX)^{-1}X^T(X\beta + \epsilon)$

$= (X^TX)^{-1}X^TX\beta + \cancel{X^T\epsilon}\ (X^TX)^{-1}X^T\epsilon$

$= \cancel{I}\beta + (X^TX)^{-1}X^T\epsilon$

Since β is a constant vector, its variance is 0, so we only need to consider the other term

let $A = (X^TX)^{-1}X^T$ for eqn. above

$Var(\hat{\beta}) = Var((X^TX)^{-1}X^T\epsilon)$

$= (X^TX)^{-1}X^T\,Var(\epsilon)\,[(X^TX)^{-1}X^T]^T$

$= (X^TX)^{-1}X^T\,Var(\epsilon)\,X(X^TX)^{-1}$

sub $Var(\epsilon) = \sigma^2 I$

$Var(\hat{\beta}) = (X^TX)^{-1}X^T\sigma^2 I X(X^TX)^{-1}$

$= (X^TX)^{-1}\sigma^2$