Formulas for Exam 1

- **Two independent samples**
  
  Test Statistic for testing $H_0 : \mu_1 = \mu_2$ or $H_0 : \mu_1 - \mu_2 = 0$
  
  $t = \frac{(\bar{y}_1 - \bar{y}_2) - 0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$
  
  $s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$

  Confidence interval for $\mu_1 - \mu_2$:
  
  $(\bar{y}_1 - \bar{y}_2) - t_{\alpha/2,k(n-1)}s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ to $(\bar{y}_1 - \bar{y}_2) + t_{\alpha/2,k(n-1)}s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

  For most situations, use a value of 2 for $t_{\alpha/2,k(n-1)}$.

- **Analysis of Variance: k groups, n observations per group**
  
  $SS_{Within} = SS_{repError} = \sum (n-1)s_i^2$
  
  $df = \sum (n-1) = k(n-1)$
  
  $MS_{Within} = MS_{repError} = s_p^2 = \frac{SS_{Within}}{k(n-1)}$
  
  $LSD = t_{\alpha/2,k(n-1)}s_p \sqrt{\frac{2}{n}}$

  Replace $t_{\alpha/2,k(n-1)}$ by the value 2 for most situations. For large values of k replace $t_{\alpha/2,k(n-1)}$ by the value 3.

  A difference in sample means is statistically significant if the absolute value of the difference is greater than the value of $LSD$. 

Least Squares Regression

Linear slope estimate: \( \hat{m}_1 = \frac{\sum C_{1i}Y_i}{\sum C_{1i}^2} \)

Intercept estimate: \( \hat{b} = \bar{Y} \)

Prediction equation: \( \hat{Y} = \hat{b} + \hat{m}_1 C_1 \)

\[ SS_{Linear} = (\hat{m}_1)^2 n \left( \sum C_{1i}^2 \right) \]
\[ df_{Linear} = 1 \]
\[ MS_{Linear} = \frac{SS_{Linear}}{df_{Linear}} \]
\[ t_{Linear} = \sqrt{F} = \frac{MS_{Linear}}{MS_{repError}} \]

The linear term is statistically significant if \( t_{Linear} \) is greater than 2 or 3.

Quadratic slope estimate: \( \hat{m}_2 = \frac{\sum C_{2i}Y_i}{\sum C_{2i}^2} \)

Prediction equation: \( \hat{Y} = \hat{b} + \hat{m}_1 C_1 + \hat{m}_2 C_2 \)

\[ SS_{Quadratic} = (\hat{m}_2)^2 n \left( \sum C_{2i}^2 \right) \]
\[ df_{Quadratic} = 1 \]
\[ MS_{Quadratic} = \frac{SS_{Quadratic}}{df_{Quadratic}} \]
\[ t_{Quadratic} = \sqrt{F} = \frac{MS_{Quadratic}}{MS_{repError}} \]

The quadratic term is statistically significant if \( t_{Quadratic} \) is greater than 2 or 3.

Coefficient of Determination: \( R^2 = \frac{SS_{Model}}{SS_{Total}} \)

Simple linear model: \( SS_{Model} = SS_{Linear} \)

Model with linear and quadratic terms: \( SS_{Model} = SS_{Linear} + SS_{Quadratic} \)