Ⅰ. Short Review

1. Two Sample Proportion Test
   - If K=2 (# of group), and N_1 and $N_2 \geq 100$, use Z test

     $Z_{obs} = \frac{(P_1 - P_2) - (\mu_{\pi 1 - \pi 2})}{\sigma_{\pi 1 - \pi 2}}$

     $\sigma_{\pi 1 - \pi 2} = \sqrt{P'(1 - P')\frac{N_1 + N_2}{N_1 N_2}}$ $P' = \frac{P_1 N_1 + P_2 N_2}{N_1 + N_2}$

2. Dependent Samples
   - Type : Matched pair samples and Repeated measures -> use t test, but N ≥ 30, Z test
   - $\overline{D} = \frac{\sum X_1 - X_2}{N} = \overline{X}_1 - \overline{X}_2$, $\mu_D = \mu_1 - \mu_2$, $\sigma_{\overline{D}} = \sqrt{\frac{\sum D^2 - (\sum D)^2/N}{N-1}} \times \frac{1}{\sqrt{N}}$
   - $t_{obs} = \frac{\overline{D} - \mu_D}{\sigma_{\overline{D}}}$, DoF : N - 1

3. ANOVA
   - Use F test (we need α , Two DoF)
   - $F_{obs} = \frac{MSb}{MSw}$; $MSb = \frac{SSb}{DFb}$, $MSw = \frac{SSw}{DFw}$

     $SSb = \sum N_i(\overline{X}_i - \overline{X})^2$, $SSw = \sum(N_i - 1)S_i^2$

     DFb = K -1 , DFw = N – K

4. Two way ANOVA
   - Main effects : If we suspect that the general relationship between our response variable and explanatory variable holds across categories of our control variable
   - Interaction effects : If we suspect that the control variable actually alters the nature of the relationship between response and explanatory variables

5. Bonferroni Multiple Comparisons and Tukey's Test
   - Bonferroni : comparing specific pair of means. use t test (N is enough large, use Z)

     $t_{obs} = \frac{\overline{X}_1 - \overline{X}_2}{\hat{\sigma}_{x1-x2}}$, $\hat{\sigma}_{x1-x2} = \hat{\sigma}\sqrt{\frac{N_1 + N_2}{N_1 N_2}}$, $\hat{\sigma}$ = MSw

     For t_{crit} , we need another α $\Rightarrow$ $\alpha = \frac{\alpha}{K(K-1)/2}$
   - Tukey : individual differences are significant (equal sample size)

     1st step : make a table to compare mean differences

     2nd step : $HSD = q\sqrt{\frac{MSw}{N}}$, 'q' needs K, DFw and α

     The difference $\geq$ HSD : the difference is significant.

6. STATA
   - Interpretation for the outcomes will be covered in the session

Ⅱ Problems

1. Test the difference of Pre and Post test result

Pre test	Post test
225	214
108	103
165	158
130	125

2. Fill the blank and Test it

Anova educ race

obs = 999 R-squared = .0078
Root MSE = 2.96709 Adj R-squared= .0058

Source	Partial SS	df	MS	F	Prob > F
Model	68.9865278	2	(?)	(?)	
race			(?)		
Residual	8768.38284	996	8.80359723		
Total	8837.36937	998	8.85507953		