Objective:
To compare the accuracies and efficiencies of different sampling methods for inventory of residual woody debris on a typical forest harvest area in North Carolina.

Method:
50 plots were established on a recently harvested site in Johnston County, North Carolina. 20 of these plots were randomly chosen from two harvest areas (13 hardwood and 7 pine). On these plots line intersect, prism sweep and 100% tally sampling designs were applied to estimate the volume of residual woody biomass remaining after typical forest harvesting. The null hypothesis is that sampling methods are not significantly different from each other ($H_0: X_1=X_2=X_3$) where $X_1$, $X_2$, and $X_3$ are the means of 100% tally sampling, line intersect sampling and prism sweep sampling respectively. The main contrast would be comparing line intersect and prism sweep sampling designs to 100% tally sampling. The response variable is the residual woody biomass left on the site.

COMMENTS:
1. Check data for possible typos, outliers and summarize data. Producing charts and plots will help to visualize the data. For example, producing bar charts (one bar for each sampling method) and scatter plots (%Tally x LI, %Tally x VPR) will show their volume means (with standard deviations) the relationships between methods.

2. The null hypothesis can be tested using one-way ANOVA approach. During ANOVA, specific contrast can be used to compare two sampling methods against 100%Tally.

3. Depending on the trends in the data, a type of regression can be fit to data (e.g., 100%Tally as independent variable, LI as dependent variable). Assuming the relationship is linear (which should be obvious in the plot) a simple linear regression can be fit. Such analysis tests whether the relationship between two sampling methods is significant (the slope parameter $b$ is significantly different from zero or not). It will also show how much bias there is between methods. A similar regression analysis can be done between VPR and 100%Tally.

4. Since this is an undergraduate student project, I suggest practicing with SAS and/or with R statistics package (freely available from the internet) would be extremely helpful to learn basic statistics with software.

I simulated some data to show how F-test can be constructed for the effects as follows:
data a;
  length Treat plot 4. ;
  retain Seed_1 0 residue;
  do treat=1 to 3;
    do plot=1 to 20;
      call ranuni (Seed_1,residue);
      output;
    end;
  end;
  drop seed_1 ;
run;

data a;
  length sampling $6. plot 4. residue 6.4 ;
  set a;
  if treat=1 then sampling='Total '\
    if treat=2 then sampling='Prism';
  if treat=3 then sampling='Line';
    if treat=1 then residue=residue+0.8 ;
    if treat=2 then residue=residue+0.84;
  drop treat ;
run;

proc print data=a ; run;

  The SAS System          16:11 Friday, June 19, 2009  34

<table>
<thead>
<tr>
<th>Obs</th>
<th>sampling</th>
<th>plot</th>
<th>residue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Total</td>
<td>1</td>
<td>1.63323</td>
</tr>
<tr>
<td>2</td>
<td>Total</td>
<td>2</td>
<td>1.29887</td>
</tr>
<tr>
<td>3</td>
<td>Total</td>
<td>3</td>
<td>1.60301</td>
</tr>
<tr>
<td>57</td>
<td>Line</td>
<td>17</td>
<td>0.95512</td>
</tr>
<tr>
<td>58</td>
<td>Line</td>
<td>18</td>
<td>0.91506</td>
</tr>
<tr>
<td>59</td>
<td>Line</td>
<td>19</td>
<td>0.30112</td>
</tr>
<tr>
<td>60</td>
<td>Line</td>
<td>20</td>
<td>0.41355</td>
</tr>
</tbody>
</table>

proc glm data=a ;
  class sampling ;
  model residue = sampling ;
  means sampling /waller ;
  *contrast 'Total verus Prism' sampling 0 -1 1;
  *contrast 'Total verus Line' sampling -1 0 1;
run; quit ;

  The GLM Procedure
Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>sampling</td>
<td>3</td>
<td>Line, Prism, Total</td>
</tr>
</tbody>
</table>

Number of Observations Used 60

Dependent Variable: residue

Source                      DF     | SS       | Mean Square | F Value | Pr > F
---                          ---- | ---------- | ----------- | ------- |-------
Model                       2     | 9.94601496 | 4.97300748  | 50.71   | <.0001
Error                       57    | 5.59026029 | 0.09807474  |         |       
Corrected Total             59    | 15.53627526|             |         |       

R-Square     Coeff Var      Root MSE    residue Mean
0.640180      29.19179      0.313169      1.072798

Source                      DF     | Type III SS | Mean Square | F Value | Pr > F
---                          ---- | ------------- | ----------- | ------- |-------
sampling                    2     | 9.94601496   | 4.97300748  | 50.71   | <.0001

The ANOVA results shows that at least one sampling method is significantly different at <0.0001 level. This is the probability of F Value.

Waller-Duncan K-ratio t Test for residue

NOTE: This test minimizes the Bayes risk under additive loss and certain other assumptions.

Kratio                              100
Error Degrees of Freedom             57
Error Mean Square                    0.098075
F Value                              50.71
Critical Value of t                  1.78647
Minimum Significant Difference       0.1769

Means with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>Waller Grouping</th>
<th>Mean</th>
<th>N</th>
<th>sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.41725</td>
<td>20</td>
<td>Prism</td>
</tr>
<tr>
<td>A</td>
<td>1.30016</td>
<td>20</td>
<td>Total</td>
</tr>
<tr>
<td>B</td>
<td>0.50099</td>
<td>20</td>
<td>Line</td>
</tr>
</tbody>
</table>

The multiple range test (Waller) shows that Total tally and the Prism methods are not different but Line method is significantly different from other two. The residue mean for Line was 0.5099 whereas the mean for Total was 1.3.

residue Mean
Line          Sampling

Prism         Total