STAT/MATH 3380 B

1. **Lectures** MWF 14:35 – 15:25

2. **Instructor** Dr. R. P. Gupta
   Room 214 Chase Buildg.
   Phone 494-2701
   email: gupta@mathstat.dal.ca

3. **Office Hours**: to be announced

4. **Text Book**: Sampling Techniques by Cochran, Following Topics will be covered:
   1. Introduction
   2. Simple Random Sampling
   3. Proportion and Percentages
   4. Sample Size
   5. Stratified Random Sampling
   6. Ratio and Regression Estimate
   7. Systematic Sampling
   8. Cluster Sampling Single Stage and Two Stage

   . Hand outs will be given in the class for many of the topics. This course needs lots of calculations so minitab or S-plus knoweldge will be helpful.You must have a good working calculator.

5. **Grading Scheme**:

Assignments	25%
Test I	15%
Test II	20%
Final Exam	40%

   This course needs lots of calculations so minitab or S-plus knoweldge will be helpful.You must have a good working calculator.

6. **Grading Scheme**:

Assignments	25%
Test I	15%
Test II	20%
Final Exam	40%

STAT/MATH 3380

Final Examination December 7, 2005

3:00 – 6:00 p.m.

1. (a) A village consists of 150 children. Dr. Jones selects a simple random sample of 15 children and count the number of cavaties in each child's mouth yielding the following results:

No. of cavities	0	1	2	3	4	5
No. of children	6	3	2	2	1	1

Dr. Smith examins all 150 children and records that 50 children have no cavities in their mouths. Estimate the total number of cavities in the village children (a) using Dr. Jones results only: (b) Using both Dr. Jones' and Dr.Smith' results; © which estimate will have a lower variance.

(b) The population of a small town consists of N=1000 people. We wish to estimate the proportion of the population having blood type AB. How large a sample is needed if we wish to estimate the true proportion with 30% relative error and with confidence 95% . Assume that we have a preliminary estimate of P to be 0.10.

2. (a) Suppose the distribution of scores of students in a large statistics class is given as below:

Score	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90	90 – 100
Frequency	4	4	9	25	16	9	9

Construct L=3 strata.

(b) If the cost function is of the form $C = c_0 + \sum t_h \sqrt{n_h}$, where c_0 and t_h are are known numbers , show that in order to minimize $V(\bar{y}_{st})$ for fixed total cost n must be proportional to

$$\left(\frac{W_h^2 S_h^2}{t_h}\right)^{2/3}$$

3. (a) Suppose we wish to estimate the weight loss of clients at a popular weight-loss clinic. Let N=300, the number of clients who are taking a diet pill daily. We take a random sample of n=8. Let x denote the no. of hours per week spent exercising and y denote their weight loss in one month. Assume $\bar{x} = 6$

Client #	1	2	3	4	5	6	7	8
x	4	5	3	6	8	7	2	5
y	20	22	17	25	30	23	15	20

Obtain the estimate of the average weight loss of a client and find the 95% confidence interval for the same. $\sum x_i y_i = 922 \quad \sum y_i^2 = 3852 \quad \sum x_i^2 = 228$

(b) Show that the variance of ratio estimate is more precise than that of simple random sample estimate iff

$$\rho > cv(x) / 2\, cv(y)$$

4. (a) To emphasize safety, a taxicab company wants to estimate the proportion of unsafe tires on their 175 cabs. Selecting a simple random sample of tires is impractical, so cluster sampling is used, with each cab as a cluster. A simple random sample of 25 cabs gives the following number of unsafe tires per cab:

2	4	0	1	2	0	4	1	3	1	2	0
1	1	2	2	4	1	0	0	3	1	2	2	1

Estimate the proportion of unsafe tires being used on the company's cabs and construct a 95% confidence interval for the same.

(b) The following table shows the number of divorses (in hundreds) in Canada for a systematic sample of years between 1950 and 1990. Estimate the total number of divorsesfor this period and find the approximate variance for your estimate.

Year	1950	1955	1960	1965	1970	1975	1980	1985	1990
Divorces	385	377	393	479	708	1036	1189	1190	1175

5. (a)Accountants frequently require that their business clients to provide cost inventries. Since a complete cost inventory is costly, a quarterly cost inventory can conveniently be accomplished by sampling. Suppose a plumbing company desires a cost inventory for many small items in stock. To obtain a simple random sample is difficult. However, the items are arranged on shelves, and selecting a simple random sample of shelves is relatively easy, treating each shelf as a cluster of items. Sampling 10 of the 48 shelves gave the following results. Estimate the total dollar amount of the items on the shelves and find a 95% confidence interval for the same.

Cluster	1	2	3	4	5	6	7	8	9	10
No. of items M_i	42	27	38	63	72	12	24	14	32	41
Total \$ amount Y_i	83	62	45	112	96	58	75	58	67	80

(b) A population consists of 3000 elements and is divided into 20 clusters, each containing 150 sub units A simple random sample of n=3 clusters is drawn. The analysis of variance table of the sample dates is as follows:

Source	df	ms
Between clusters	-	2.5
Error (within subunits within Clusters)	-	3.0
Total (within subunits in The sample)	449	3.1

Find an unbiased estimate of Sw^2, S_b^2 and S^2. Find Intra class correlation coeff.