UNIVERSITY OF ESWATINI

DEPARTMENT OF STATISTICS AND DEMOGRAPHY

FINAL EXAMINATION PAPER 2018

TITLE OF PAPER : SAMPLING THEORY

COURSE CODE : STA 305/ST 306

TIME ALLOWED : 2 HOURS

REQUIREMENTS : STATISTICAL TABLES AND CALCULATOR

INSTRUCTIONS

1. Answer any three (3) questions
2. Show clearly all your working

THIS PAPER IS NOT TO BE OPENED UNTIL PERMISSION HAS BEEN GRANTED BY THE INVIGILATOR

Question 1

A local authority is investigating various aspects of the usage of its public libraries.

(a) A survey of a simple random sample of students in secondary education was undertaken, to determine their use of library services. There are 12 000 such students altogether in this local authority area. One of the survey questions was 'How many times have you visited a public library in the past four weeks?' The results were as follows.

No. of visits in past 4 weeks	Number of students
0	68
1	99
2	50
3	45
4	130
>4	58
Total	450

Obtain a point estimate and an approximate 95% confidence interval for;

i. the proportion of students in this local authority area who visited a public library in the past 4 weeks, [*4 marks*]

ii. the total number of visits made to a public library in the past four weeks by students in this local authority area. You should use a value of 5 for calculating the mean and standard deviation where the number of visits in the past four weeks is 5 or more. [*8 marks*]

(b) Define the following [*8 marks*]

i. target population

ii. sampling unit

iii. sampling frame

iv. nonsampling errors.

Question 2

Wildlife managers want to estimate the total number of caribou in the Nelchina herd located in south central Alaska. The density of caribou differs dramatically in different types of habitat. A preliminary aerial investigation has identified the area used by the herd, and divided it into six strata based on habitat type. For the main survey, the organiser decides to divide the area into sub-areas called quadrats, each of size 4 km^2. The survey is conducted by selecting a simple random sample of quadrats from each stratum, and for each quadrat the area is searched by aircraft to locate and then photograph the animals; the number of caribou, y, in each quadrat is counted in the photographs.

The sample means and standard deviations of the measurements, y, in each stratum based on a sample of 211 quadrats are as follows.

Stratum (h)	Map quadrats (N_h)	Sample quadrats (n_h)	Sample mean	Sample standard deviation
1	400	98	24.1	74.7
2	40	10	25.6	63.7
3	100	37	267.6	589.5
4	40	6	179.0	151.0
5	70	39	293.7	351.5
6	120	21	33.2	99.0
Total	770	211		

(a) The sampling frame for this survey is a land map. Discuss briefly what problems are likely to be associated with this type of sample. [*4 marks*]

(b) The formula for the variance of the estimator of a population total based on a stratified (random) sample is

$$V = \sum_{h=1}^{L} N_h(N_h - n_h)\frac{S_h^2}{n_h}$$

Define the terms N_h , S_h and n_h in the formula above. [*2 marks*]

(c) Using the data above, estimate the total number of caribou in the herd and obtain an approximate 95% confidence interval for this total. [*7 marks*]

(d) For this survey discuss briefly the merits of using stratified sampling rather than simple random sampling. [*4 marks*]

(e) Given that stratified sampling is used for this survey, discuss briefly the merits of using optimal rather than proportional allocation. [*3 marks*]

Question 3

Let $\mathcal{S} = \{s_1, s_2, ..., s_M\}$ consists of M different samples that can be possibly be drawn from the population U. Consider a population with 4 elements,

$$U = \{u_1, u_2, u_3, u_4\}.$$

A sample of size $n = 2$ is to be drawn from the population.

(a) Consider for the moment the case of simple random sampling

i. Obtain $M = |\mathcal{S}|$. [*2 marks*]

ii. Obtain π_k. [*2 marks*]

iii. Obtain the sum of all inclusion probabilities of elements $k \in U$. [*2 marks*]

(b) Consider the following sampling design: $\mathcal{S}_n = \{s_1, s_2, s_3\}$ where

$s_1 = \{u_1, u_3\}$, $s_2 = \{u_1, u_4\}$, $s_3 = \{u_2, u_4\}$.

Assume the following probabilities for the samples

$p(s_1) = 0.1$, $p(s_2) = 0.6$, $p(s_3) = 0.3$

i. Obtain all inclusion probabilities π_k. [*6 marks*]

ii. Obtain numerically the sum of inclusion probabilities. [*2 marks*]

iii. Obtain all inclusion probabilities $\pi_{k,l}$. [*6 marks*]

Question 4

A small survey was carried out in a Sub-Saharan country to estimate the total number of bunches of bananas produced in a district during a given growing period. The district was divided into 289 primary units such that each unit had about 500 to 1000 banana pits. Each pit may produce 0, 1 or more bunches of bananas. The total number of banana pits for the whole district was known to be 181 336. A simple random sample of 20 primary units was selected from the 289 units, and for each unit the number of banana pits (x) and the total number of banana bunches (y) were obtained. The results for $n = 20$ are summarised below.

	Mean	SD
Number of banana pits per unit (x)	644.35	115.9025
Total number of banana bunches per unit (y)	901.70	221.8112

i. Using the mean of a simple random sample, estimate the total number of banana bunches for the district, and estimate the standard error of your estimator. [*5 marks*]

ii. The researcher seeks your advice on how he might use the supplementary data on the numbers of banana pits per unit to estimate the total number of banana bunches in the district. He has estimated the correlation between the number of banana pits per unit and the total number of banana bunches per unit (i.e. between x and y above) and thinks a ratio estimator might be suitable.

   (a) Explain what is meant by *correlation*. Given that the estimated correlation is 0.7737, how would you respond? Briefly discuss the properties of the ratio estimator and the estimator based on the sample mean in part (i). [*5 marks*]

   (b) Give the ratio estimate of the total number of banana bunches in the district and its estimated standard error. Hence, giving a reason, say whether this estimate is better than that calculated in part (i), and use it to construct an approximate 95% confidence interval for the true total number of banana bunches in the district. Explain what this confidence interval shows. [*10 marks*]

Question 5

A simple random sample of 10 hospitals was selected from a population of 33 hospitals that had received state funding to upgrade their emergency medical services. Within each of the selected hospitals, the records were examined for all patients hospitalised in the past 12 months for trauma, which is defined as body wounds or shock produced by sudden physical injury due to accident or violence. The numbers of patients hospitalised for trauma, and the numbers of patients with trauma who were discharged dead, for the selected hospitals are given below.

Hospital	Number of patients hospitalized for trauma	Number with trauma discharged death
1	560	4
2	190	4
3	260	2
4	370	4
5	190	4
6	130	0
7	170	9
8	170	2
9	60	0
10	110	1

(a) Explain why this design may be considered as a cluster sample. What are the first-stage and second-stage units? [*2 marks*]

(b) Obtain a point estimate and an approximate 95% confidence interval for the total number of persons hospitalised for trauma conditions for the 33 hospitals. State the properties of your estimator. [*6 marks*]

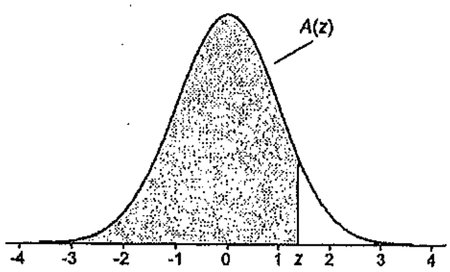
(c) Obtain a point estimate of the proportion of persons discharged dead among those hospitalised for trauma conditions for the 33 hospitals, using the cluster totals. Hence calculate an approximate 95% confidence interval for this proportion, and comment on the validity of the assumptions necessary for this calculation. [*6 marks*]

(d) Give reasons why, for this survey, cluster sampling might be preferred to stratified random sampling.

What might be the drawbacks of cluster sampling? Discuss, with reasons, any improvements you might make if another survey was being planned on the same topic. [*6 marks*]

TABLE A.1

Cumulative Standardized Normal Distribution



$A(z)$ is the integral of the standardized normal distribution from $-\infty$ to z (in other words, the area under the curve to the left of z). It gives the probability of a normal random variable not being more than z standard deviations above its mean. Values of z of particular importance:

z	$A(z)$	
1.645	0.9500	Lower limit of right 5% tail
1.960	0.9750	Lower limit of right 2.5% tail
2.326	0.9900	Lower limit of right 1% tail
2.576	0.9950	Lower limit of right 0.5% tail
3.090	0.9990	Lower limit of right 0.1% tail
3.291	0.9995	Lower limit of right 0.05% tail

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999							

TABLE A.2

t Distribution: Critical Values of t

Degrees of freedom	Two-tailed test: / One-tailed test:	10% / 5%	5% / 2.5%	2% / 1%	1% / 0.5%	0.2% / 0.1%	0.1% / 0.05%
		Significance level					
1		6.314	12.706	31.821	63.657	318.309	636.619
2		2.920	4.303	6.965	9.925	22.327	31.599
3		2.353	3.182	4.541	5.841	10.215	12.924
4		2.132	2.776	3.747	4.604	7.173	8.610
5		2.015	2.571	3.365	4.032	5.893	6.869
6		1.943	2.447	3.143	3.707	5.208	5.959
7		1.894	2.365	2.998	3.499	4.785	5.408
8		1.860	2.306	2.896	3.355	4.501	5.041
9		1.833	2.262	2.821	3.250	4.297	4.781
10		1.812	2.228	2.764	3.169	4.144	4.587
11		1.796	2.201	2.718	3.106	4.025	4.437
12		1.782	2.179	2.681	3.055	3.930	4.318
13		1.771	2.160	2.650	3.012	3.852	4.221
14		1.761	2.145	2.624	2.977	3.787	4.140
15		1.753	2.131	2.602	2.947	3.733	4.073
16		1.746	2.120	2.583	2.921	3.686	4.015
17		1.740	2.110	2.567	2.898	3.646	3.965
18		1.734	2.101	2.552	2.878	3.610	3.922
19		1.729	2.093	2.539	2.861	3.579	3.883
20		1.725	2.086	2.528	2.845	3.552	3.850
21		1.721	2.080	2.518	2.831	3.527	3.819
22		1.717	2.074	2.508	2.819	3.505	3.792
23		1.714	2.069	2.500	2.807	3.485	3.768
24		1.711	2.064	2.492	2.797	3.467	3.745
25		1.708	2.060	2.485	2.787	3.450	3.725
26		1.706	2.056	2.479	2.779	3.435	3.707
27		1.703	2.052	2.473	2.771	3.421	3.690
28		1.701	2.048	2.467	2.763	3.408	3.674
29		1.699	2.045	2.462	2.756	3.396	3.659
30		1.697	2.042	2.457	2.750	3.385	3.646
32		1.694	2.037	2.449	2.738	3.365	3.622
34		1.691	2.032	2.441	2.728	3.348	3.601
36		1.688	2.028	2.434	2.719	3.333	3.582
38		1.686	2.024	2.429	2.712	3.319	3.566
40		1.684	2.021	2.423	2.704	3.307	3.551
42		1.682	2.018	2.418	2.698	3.296	3.538
44		1.680	2.015	2.414	2.692	3.286	3.526
46		1.679	2.013	2.410	2.687	3.277	3.515
48		1.677	2.011	2.407	2.682	3.269	3.505
50		1.676	2.009	2.403	2.678	3.261	3.496
60		1.671	2.000	2.390	2.660	3.232	3.460
70		1.667	1.994	2.381	2.648	3.211	3.435
80		1.664	1.990	2.374	2.639	3.195	3.416
90		1.662	1.987	2.368	2.632	3.183	3.402
100		1.660	1.984	2.364	2.626	3.174	3.390
120		1.658	1.980	2.358	2.617	3.160	3.373
150		1.655	1.976	2.351	2.609	3.145	3.357
200		1.653	1.972	2.345	2.601	3.131	3.340
300		1.650	1.968	2.339	2.592	3.118	3.323
400		1.649	1.966	2.336	2.588	3.111	3.315
500		1.648	1.965	2.334	2.586	3.107	3.310
600		1.647	1.964	2.333	2.584	3.104	3.307
∞		1.645	1.960	2.326	2.576	3.090	3.291

Useful formulas

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} \qquad \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{\sum_{i=1}^n y_i}{n}$$

$$\hat{\mu}_{srs} = \bar{y} \qquad \hat{\mathbb{V}}(\hat{\mu}_{srs}) = \left(\frac{N-n}{N}\right)\frac{s^2}{n}$$

$$\hat{\tau}_{srs} = N\hat{\mu}_{srs} \qquad \hat{\mathbb{V}}(\hat{\tau}_{srs}) = N^2\hat{\mathbb{V}}(\hat{\mu}_{srs})$$

$$\hat{p}_{srs} = \sum_{i=1}^n \frac{y_i}{n} \qquad \left(\frac{N-n}{N}\right)\frac{\hat{p}(1-\hat{p})}{n-1}\left(\frac{N-n}{N}\right)$$

$$\hat{\tau}_{hh} = \frac{1}{n}\sum_{i=1}^n \frac{y_i}{p_i} \qquad \hat{\mathbb{V}}(\hat{\mu}_{hh}) = \frac{1}{n(n-1)}\sum_{i=1}^n \left(\frac{y_i}{p_i} - \hat{\tau}_{hh}\right)^2$$

$$\hat{\mu}_{hh} = \frac{\hat{\tau}_{hh}}{N} \qquad \hat{\mathbb{V}}(\hat{\mu}_{hh}) = \frac{1}{N^2}\hat{\mathbb{V}}(\hat{\tau}_{hh})$$

$$\hat{\tau}_{ht} = \sum_{i=1}^{\nu} \frac{y_i}{\pi_i} \qquad \hat{\mathbb{V}}(\hat{\tau}_{ht}) = \sum_{i=1}^{\nu}\left(\frac{1}{\pi_i^2} - \frac{1}{\pi_i}\right)y_i^2 + 2\sum_{i=1}^{\nu}\sum_{j>i}^{\nu}\left(\frac{1}{\pi_i\pi_j} - \frac{1}{\pi_{ij}}\right)y_iy_j$$

$$\hat{\mu}_{ht} = \frac{\hat{\tau}_{ht}}{N} \qquad \hat{\mathbb{V}}(\hat{\mu}_{ht}) = \frac{1}{N^2}\hat{\mathbb{V}}(\hat{\tau}_{ht})$$

$$\hat{r} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} \qquad \hat{\mathbb{V}}(\hat{r}) = \left(\frac{N-n}{Nn\mu_x^2}\right)\frac{\sum_{i=1}^n (y_i - rx_i)^2}{n-1}$$

$$\hat{\mu}_r = r\mu_x \qquad \hat{\mathbb{V}}(\hat{\mu}_r) = \left(\frac{N-n}{Nn}\right)\frac{\sum_{i=1}^n (y_i - rx_i)^2}{n-1}$$

$$\hat{\tau}_r = Nr\mu_x = r\tau_x \qquad \hat{\mathbb{V}}(\hat{\tau}_r) = \frac{N(N-n)}{n}\frac{\sum_{i=1}^n (y_i - rx_i)^2}{n-1}$$

$$\hat{\mu}_L = a + b\mu_x \qquad \hat{\mathbb{V}}(\hat{\mu}_L) = \frac{N-n}{Nn(n-2)}\sum_{i=1}^n (y_i - a - bx_i)^2$$

$$\hat{\tau}_L = N\mu_L \qquad \hat{\mathbb{V}}(\hat{\tau}_L) = \frac{N(N-n)}{n(n-2)}\sum_{i=1}^n (y_i - a - bx_i)^2$$

$$\hat{\mu}_{str} = \sum_{h=1}^L \frac{N_h}{N}\bar{y}_h \qquad \hat{\mathbb{V}}(\hat{\mu}_{str}) = \frac{1}{N^2}\sum_{h=1}^L N_h^2\left(\frac{N_h - n_h}{N_h}\right)\frac{s_h^2}{n_h}$$

$$\hat{\tau}_{str} = N\hat{\mu}_{str} \qquad \hat{\mathbb{V}}(\hat{\tau}_{str}) = N^2\hat{\mathbb{V}}(\hat{\mu}_{str})$$

$$\hat{p}_{str} = \sum_{h=1}^L \frac{N_h}{N}\hat{p}_h \qquad \hat{\mathbb{V}}(\hat{p}_{str}) = \frac{1}{N^2}\sum_{h=1}^L N_h^2\left(\frac{N_h - n_h}{N_h}\right)\left(\frac{\hat{p}_h(1-\hat{p}_h)}{n_h - 1}\right)$$

$$\hat{\mu}_{pstr} = \sum_{h=1}^L w_h\bar{y}_h \qquad \hat{\mathbb{V}}(\hat{\mu}_{pstr}) = \frac{1}{n}\left(\frac{N-n}{N}\right)\sum_{h=1}^L w_hs_h^2 + \frac{1}{n^2}\sum_{h=1}^L (1-w_h)s_h^2$$

$$\hat{\tau}_{cl} = \frac{M}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L} y_{ij} = \frac{N}{n}\sum_{i=1}^{n}\sum_{j=1}^{L} y_{ij} = \frac{N}{n}\sum_{i=1}^{n} y_i = N\bar{y}$$

$$\hat{\mu}_{cl} = \frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L} y_{ij} = \frac{1}{nL}\sum_{i=1}^{n} y_i = \frac{\bar{y}}{L} = \frac{\hat{\tau}_{cl}}{M}$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{\hat{\tau}_{cl}}{N}$

$$\hat{V}(\hat{\tau}_{cl}) = N(N-n)\frac{s_u^2}{n} \qquad \hat{V}(\hat{\mu}_{cl}) = \frac{N(N-n)}{M^2}\frac{s_u^2}{n}$$

where $s_u^2 = \frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}$.

$$\hat{\mu}_1 = \bar{y} = \frac{\hat{\tau}_{cl}}{N} \qquad \hat{V}(\hat{\mu}_1 = \frac{N-n}{N}\frac{s_u^2}{n}$$

The formulas for systematic sampling are the same as those used for one-stage cluster sampling. Change the subscript *cl* to *sys* to denote the fact that data were collected under systematic sampling.

$$\hat{\mu}_{c(a)} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} M_i} = \frac{\sum_{i=1}^{n} y_i}{m} \qquad \hat{V}(\hat{\mu}_{c(a)}) = \frac{(N-n)N}{n(n-1)M^2}\sum_{i=1}^{n} M_i^2(\bar{y} - \hat{\mu}_{c(a)})^2$$

$$\hat{\mu}_{c(b)} = \frac{N}{M}\frac{\sum_{i=1}^{n} y_i}{n} = \frac{N}{nM}\sum_{i=1}^{n} y_i \qquad \hat{V}(\hat{\mu}_{c(b)}) = \frac{(N-n)N}{n(n-1)M^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{(N-n)N}{nM^2}s_u^2$$

$$\hat{p}_c = \frac{\sum_{i=1}^{n} p_i}{n} \qquad \hat{V}(\hat{p}_c) = \left(\frac{N-Nn}{nN}\right)\sum_{i=1}^{n}\frac{(p_i - \hat{p}_c)^2}{n-1} = \left(\frac{1-f}{n}\right)\sum_{i=1}^{n}\frac{(p_i - \hat{p}_c)^2}{n-1}$$

$$\hat{p}_c = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} M_i} \qquad \hat{V}(\hat{p}_c) = \left(\frac{1-f}{n\bar{m}^2}\right)\frac{\sum_{i=1}^{n}(y_i - \hat{p}_c M_i)^2}{n-1}$$

To estimate τ, multiply $\hat{\mu}_{c(\cdot)}$ by M. To get the estimated variances, multiply $\hat{V}(\hat{\mu}_{c(\cdot)})$ by M^2. If M is not known, substitute M with $N\bar{m} = Nm/n$. $\bar{m} = \sum_{i=1}^{n} M_i/n$.

n for μ SRS $\quad n = \dfrac{N\sigma^2}{(N-1)(d^2/z^2) + \sigma^2}$

n for τ SRS $\quad n = \dfrac{N\sigma^2}{(N-1)(d^2/z^2N^2) + \sigma^2}$

n for p SRS $\quad n = \dfrac{Np(1-p)}{(N-1)(d^2/z^2) + p(1-p)}$

n for μ SYS $\quad n = \dfrac{N\sigma^2}{(N-1)(d^2/z^2) + \sigma^2}$

n for τ SYS $\quad n = \dfrac{N\sigma^2}{(N-1)(d^2/z^2N^2) + \sigma^2}$

n for μ STR $\quad n = \dfrac{\sum_{h=1}^{L} N_h^2(\sigma_h^2/w_h)}{N^2(d^2/z^2) + \sum_{h=1}^{L} N_h\sigma_h^2}$

n for τ STR $\quad n = \dfrac{\sum_{h=1}^{L} N_h^2(\sigma_h^2/w_h)}{N^2(d^2/z^2N^2) + \sum_{h=1}^{L} N_h\sigma_h^2}$

where $w_h = \frac{n_h}{n}$.

Allocations for STR μ:

$$n_h = (c - c_0)\left(\frac{N_h\sigma_h/\sqrt{c_h}}{\sum_{k=1}^{L} N_k\sigma_k\sqrt{c_k}}\right) \qquad n = \frac{\left(\sum_{k=1}^{L} N_k\sigma_k\sqrt{c_k}\right)\left(\sum_{k=1}^{L} N_k\sigma_k/\sqrt{c_k}\right)}{N^2(d^2/z^2) + \sum_{k=1}^{L} N_k\sigma_k^2}$$

$$n_h = n\left(\frac{N_h}{N}\right) \qquad n = \frac{\sum_{k=1}^{L} N_k\sigma_k}{N^2(d^2/z^2) + \frac{1}{N}\sum_{k=1}^{L} N_k\sigma_k^2}$$

$$n_h = n\left(\frac{N_h\sigma_h}{\sum_{k=1}^{L} N_k\sigma_k}\right) \qquad n = \frac{\left(\sum_{k=1}^{L} N_k\sigma_k\right)^2}{N^2(d^2/z^2) + \sum_{k=1}^{L} N_k\sigma_k^2}$$

Allocations for STR τ:

change $N^2(d^2/z^2)$ to $N^2(d^2/z^2N^2)$

Allocations for STR p:

$$n_h = n\left(\frac{N_i\sqrt{p_h(1-p_h)/c_h}}{\sum_{k=1}^{L} N_k\sqrt{p_k(1-p_k)/c_k}}\right) \qquad n = \frac{\sum_{k=1}^{L} N_k p_k(1-p_k)/w_k}{N^2(d^2/z^2) + \sum_{k=1}^{L} N_k p_k(1-p_k)}$$