Auditor's Approach to Statistical Sampling, Volume 5. (Supplementary Section) Ratio and Difference Estimation

American Institute of Certified Public Accountants. Professional Development Division. Individual Study Program

(Supplementary Section) **5**

AN AUDITOR'S APPROACH TO STATISTICAL SAMPLING

RATIO AND DIFFERENCE ESTIMATION

Individual Study Program
Professional Development Division
American Institute of Certified Public Accountants

NOTICE TO READERS

This programed learning text is a publication of the staff of the American Institute of Certified Public Accountants, and is not to be regarded as an official pronouncement of the Institute. It was prepared under the general supervision of the members of the Committee on Statistical Sampling, with primary advisory responsibilities assigned to a special subcommittee composed of James W. Kelley, CPA, John R. Coker, CPA, John R. Rogers, CPA, and Dr. Robert Taylor, ÇPA. Dr. Donald M. Roberts acted as a statistical consultant in preparing this text for the Committee. Stephen J. Gallopo, CPA, Assistant Manager, Special Projects coordinated the activities as AICPA staff aide.

RATIO AND DIFFERENCE ESTIMATION

Volume 5

SUPPLEMENTARY SECTION

TABLE OF CONTENTS

EXHIBITS

APPENDICES

Description of Problem

The PQR Emporium is a medium-sized specialty store. It has an inventory of 5,000 items ranging from thumbtacks to TV sets. The company took a physical inventory as of March 31. The auditor was present to observe and make test counts. The inventory has now been priced and totaled at $2,500,000.

The auditor wants to test the client's pricing and extensions of the inventory. Since the client's procedures are - in the auditor's opinion - reasonably adequate, the auditor believes that it should not be necessary for him to test pricing and extensions for all 5,000 items.

For this test, he decides to use statistical sampling as a basis for making his own estimate of the total inventory value. In relation to the financial statements taken as a whole, he believes that a misstatement of the inventory by more than $140,000 would be material. On this basis he chooses a desired precision of $70,000 and a reliability level of 95%.

If the inventory book value falls within his estimated value, ± $70,000, the auditor decides in advance that he will accept the client's total as fairly stated with respect to pricing and extension.

If the book value does not fall within the precision interval, the auditor will have to decide what additional evidence he requires to determine whether the client's value is fairly stated.

EXHIBIT 2
PQR EMPORIUM

DATA SHEET

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	B	Total inventory value from priced perpetual record	Exhibit 1	$2,500,000
2	$\hat{X}$		Sample Data	
3	N	Population size	Exhibit 1	
4	A	Desired precision	Exhibit 1	
5	R	Desired reliability	Exhibit 1	
6	U_R	Reliability factor	Appendix 2	

EXHIBIT 3
PQR EMPORIUM

SAMPLE DATA SHEET

1	2	3	4	5	6	7	8	9	10
Element No.	Random No.	Item Description	Recorded			Audited			Difference
			Quantity	Price	Extended Value	Quantity	Price	Extended Value	Extended Value
1	0163	Electric widgets G-73215	70	$1.00	$70	70	$1.00	$70	
2	2748	Trays T-4226	80	$1.50	$120	80	$1.00	$80	$(40)
3	0509	Pens P-4275	512	.50	$256	512	$.50	$256	
4	1472	Thumbtacks T-42170	900	.01	$9	900	$.01	$9	
5	0024	Chairs C-4228	208	$10.00	$2080	208	$8.65	$1800	$(280)

EXHIBIT 4
PQR EMPORIUM

PRELIMINARY SAMPLE DATA

	b_i BOOK VALUE	a_i AUDITED VALUE	$a_i - b_i$ DIFFERENCE
1	70	70	
2	120	80	(40)
3	256	256	
4	9	9	
5	2,080	1,800	(280)
6	564	564	
7	63	63	
8	185	185	
9	1,731	1,731	
10	163	119	(44)
11	817	817	
12	315	315	
13	22	22	
14	965	965	
15	278	278	
16	140	140	
17	2,411	2,566	155
18	390	390	
19	243	243	
20	53	53	
21	922	922	
22	289	289	
23	122	600	478
24	686	426	(260)
25	487	487	

PRELIMINARY SAMPLE DATA, contd.

	BOOK VALUE	AUDITED VALUE	DIFFERENCE
26	136	136	
27	755	755	
28	90	152	62
29	208	208	
30	1,427	1,568	141
31	525	525	
32	1,020	1,020	
33	87	87	
34	350	350	
35	875	875	
36	127	109	(18)
37	1,325	1,325	
38	80	80	
39	300	530	230
40	450	450	
41	700	700	
42	201	201	
43	495	495	
44	301	400	99
45	625	625	
46	387	387	
47	725	725	
48	637	637	
49	460	460	
50	485	485	
Total	26,152	26,675	523

EXHIBIT 5
PQR EMPORIUM

DATA SHEET

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum b_i$	Sum of sample book values	Exhibit 4 data	
2	$\sum b_i^2$	Sum of squares of sample book values	Exhibit 4 data	
3	$\sum d_i$	Sum of sample differences	Exhibit 4 data	
4	$\sum d_i^2$	Sum of squares of sample differences	Exhibit 4 data	
5	$\sum b_i \, d_i$	Sum of cross products	Exhibit 4 data	
6	$(\hat{R}-1)$	Estimated ratio -1.0	Row 3 ÷ Row 1	
7	$(\hat{R}-1)^2$		Row 6 squared	
8	$2(\hat{R}-1)$		Row 6 doubled	

EXHIBIT 6
PQR EMPORIUM

DATA SHEET (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum b_i$	Sum of sample book values	Exhibit 4 data	26,152
2	$\sum b_i^2$	Sum of squares of sample book values	Exhibit 4 data	27,012,204
3	$\sum d_i$	Sum of sample differences	Exhibit 4 data	523
4	$\sum d_i^2$	Sum of squares of sample differences	Exhibit 4 data	488,795
5	$\sum b_i\ d_i$	Sum of cross products	Exhibit 4 data	(37,411)
6	$(\hat{R}-1)$	Estimated ratio -1.0	Row 3 ÷ Row 1	.02
7	$(\hat{R}-1)^2$		Row 6 squared	.0004
8	$2(\hat{R}-1)$		Row 6 doubled	.04

EXHIBIT 7
PQR EMPORIUM

$$S^2_{R_j} = \frac{\sum d_i^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1) \sum b_i d_i}{n-1}$$

ESTIMATED STANDARD DEVIATION
OF POPULATION OF RATIOS

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum b_i^2$	Sum of squares of sample book values	Exhibit 6, Row 2	
2	$\sum d_i^2$	Sum of squares of sample differences	Exhibit 6, Row 4	
3	$(\hat{R}-1)^2$		Exhibit 6, Row 7	
4	$(\hat{R}-1)^2 \sum b_i^2$		Row 1 x Row 3	
5	$\sum b_i d_i$	Sum of cross products	Exhibit 6, Row 5	
6	$2(\hat{R}-1)$		Exhibit 6, Row 8	
7	$2(\hat{R}-1) \sum b_i d_i$		Row 5 x Row 6	
8	$\sum d_i^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1)\sum b_i d_i$		Row 2 + Row 4 - Row 7	
9	$n-1$	Sample size minus one	Exhibit 4	
10	$S^2_{R_j}$		Row 8 ÷ Row 9	
11	S_{R_j}	Estimated standard deviation of population of ratios	Square root of Row 10	

EXHIBIT 8
PQR EMPORIUM

$$s^2_{R_j} = \frac{\sum d_1^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1)\sum b_i d_i}{n-1}$$

ESTIMATED STANDARD DEVIATION OF POPULATION OF RATIOS (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum b_i^2$	Sum of squares of sample book values	Exhibit 6, Row 2	27,012,204
2	$\sum d_i^2$	Sum of squares of sample differences	Exhibit 6, Row 4	488,795
3	$(\hat{R}-1)^2$		Exhibit 6 Row 7	.0004
4	$(\hat{R}-1)^2 \sum b_1^2$		Row 1 x Row 3	10,805
5	$\sum b_i d_i$	Sum of cross products	Exhibit 6, Row 5	(37,411)
6	$2(\hat{R}-1)$		Exhibit 6, Row 8	.04
7	$2(\hat{R}-1)\sum b_i d_1$		Row 5 x Row 6	(1,496)
8	$\sum d_i^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1)\sum b_i d_i$		Row 2 + Row 4 - Row 7	501,096
9	$n-1$	Sample size minus one		49
10	$s^2_{R_j}$		Row 8 ÷ Row 9	10,226
11	s_{R_j}	Estimated standard deviation of population of ratios	Square root of Row 10	101.1

EXHIBIT 9
PQR EMPORIUM

$$n = \frac{n'}{1 + \frac{n'}{N}} \quad \text{where } n' = \left(\frac{S_{R_j} \cdot U_R \cdot N}{A} \right)^2$$

COMPUTATION OF SAMPLE SIZE

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	
2	S_{R_j}	Estimated standard deviation of population of ratios	Exhibit 7, Row 11	
3	R	Desired reliability	Exhibit 2	
4	U_R	Reliability factor	Appendix 2	
5	A	Desired precision	Exhibit 2	
6	$\frac{S_{R_j} \cdot U_R \cdot N}{A}$		(Row 1 x Row 2 x Row 4) ÷ Row 5	
7	n'	Sample size (with replacement)	Row 6 squared	
8	$\frac{n'}{N}$	Adjustment factor	Row 7 ÷ Row 1	
9	$1 + \frac{n'}{N}$		1.0 + Row 8	
10	n	Sample size (without replacement)	Row 7 ÷ Row 9	
11	$n + .1n$	Sample size with 10% safety factor	Row 10 + .1 x Row 10	

$$n = \frac{n'}{1 + \frac{n'}{N}} \quad \text{where } n' = \left(\frac{S_{R_j} \cdot U_R \cdot N}{A} \right)^2$$

COMPUTATION OF SAMPLE SI~~E (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	5,000
2	S_{R_j}	Estimated standard deviation of population of ratios	Exhibit 7, Row 11	101.1
3	R	Desired reliability	Exhibit 2	95%
4	U_R	Reliability factor	Appendix 2	1.96
5	A	Desired precision	Exhibit 2	70,000
6	$\frac{S_{D_j} \cdot U_R \cdot N}{A}$		(Row 1 x Row 2 x Row 4) ÷ Row 5	14.15
7	n'	Sample size (with replacement)	Row 6 squared	200
8	$\frac{n'}{N}$	Adjustment factor	Row 7 ÷ Row 1	.04
9	$1 + \frac{n'}{N}$		1.0 + Row 8	1.04
10	n	Sample size (without replacement)	Row 7 ÷ Row 9	192
11	$n + .1n$	Sample size with 10% safety factor	Row 10 + .1 x Row 10	211

COMBINED SAMPLE DATA SHEET

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum b_i$	Sum of preliminary sample book values	Row 1, Exhibit 6	
2	$\sum b_i$	Sum of additional sample book values		89,978
3	$\sum b_i$	Sum of combined sample book values	Row 1 plus Row 2	
4	$\sum b_i^2$	Sum of squares of preliminary sample	Row 1, Exhibit 8	
5	$\sum b_i^2$	Sum of squares of additional sample		89,707,686
6	$\sum b_i^2$	Sum of squares of combined sample	Row 4 plus Row 5	
7	$\sum d_i$	Sum of preliminary sample differences	Row 3, Exhibit 6	
8	$\sum d_i$	Sum of additional sample differences		1,724
9	$\sum d_i$	Sum of combined sample differences	Row 7 plus Row 8	
10	$\sum d_i^2$	Sum of squares of preliminary sample	Row 2, Exhibit 8	
11	$\sum d_i^2$	Sum of squares of additional sample		1,448,829
12	$\sum d_i^2$	Sum of squares of combined sample	Row 10 plus Row 11	
13	$\sum b_i d_i$	Sum of cross products of preliminary sample	Row 5, Exhibit 8	
14	$\sum b_i d_i$	Sum of cross products of additional sample		143,038
15	$\sum b_i d_i$	Sum of cross products of combined sample	Row 13 plus Row 14	

COMBINED SAMPLE DATA SHEET, contd.

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
16	$(\hat{R}-1)$	Estimated ratio -1.0	Row 9 ÷ Row 3	
17	$(\hat{R}-1)^2$		Row 16 squared	
18	$2(\hat{R}-1)$		2.0 times (Row 16)	

COMBINED SAMPLE DATA SHEET (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum b_i$	Sum of preliminary sample book values	Row 1, Exhibit 6	26,152
2	$\sum b_i$	Sum of additional sample book values		89,978
3	$\sum b_i$	Sum of combined sample book values	Row 1 plus Row 2	116,130
4	$\sum b_i^2$	Sum of squares of preliminary sample	Row 1, Exhibit 8	27,012,204
5	$\sum b_i^2$	Sum of squares of additional sample		89,707,686
6	$\sum b_i^2$	Sum of squares of combined sample	Row 4 plus Row 5	116,719,890
7	$\sum d_i$	Sum of preliminary sample differences	Row 3, Exhibit 6	523
8	$\sum d_i$	Sum of additional sample differences		1,724
9	$\sum d_i$	Sum of combined sample differences	Row 7 plus Row 8	2,247
10	$\sum d_i^2$	Sum of squares of preliminary sample	Row 2, Exhibit 8	488,795
11	$\sum d_i^2$	Sum of squares of additional sample		1,448,829
12	$\sum d_i^2$	Sum of squares of combined sample	Row 10 plus Row 11	1,937,624
13	$\sum b_i d_i$	Sum of cross products of preliminary sample	Row 5, Exhibit 8	(37,411)
14	$\sum b_i d_i$	Sum of cross products of additional sample		143,038
15	$\sum b_i d_i$	Sum of cross products of combined sample	Row 13 plus Row 14	105,627

COMBINED SAMPLE DATA SHEET (ANSWERS), contd.

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
16	$(\hat{R}-1)$	Estimated ratio -1.0	Row 9 ÷ Row 3	.02
17	$(\hat{R}-1)^2$		Row 16 squared	.0004
18	$2(\hat{R}-1)$		2.0 times (Row 16)	.04

EXHIBIT 13
PQR EMPORIUM

$$S^2_{R_j} = \frac{\sum d_i^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1) \sum b_i d_i}{n-1}$$

ESTIMATED STANDARD DEVIATION
OF POPULATION OF RATIOS

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum b_i^2$	Sum of squares of sample book values	Exhibit 12, Row 6	
2	$\sum d_i^2$	Sum of squares of sample differences	Exhibit 12, Row 12	
3	$(\hat{R}-1)^2$		Exhibit 12, Row 17	
4	$(\hat{R}-1)^2 \sum b_i^2$		Row 1 x Row 3	
5	$\sum b_i d_i$	Sum of cross products	Exhibit 12, Row 15	
6	$2(\hat{R}-1)$		Exhibit 12, Row 18	
7	$2(\hat{R}-1) \sum b_i d_i$		Row 5 x Row 6	
8	$\sum d_i^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1)\sum b_i d_i$		Row 2 + Row 4 - Row 7	
9	n-1	Sample size minus one	Exhibit 10, Row 11	
10	$S^2_{R_j}$		Row 8 ÷ Row 9	
11	S_{R_j}	Estimated standard deviation of population of ratios	Square root of Row 10	

$$S^2_{R_j} = \frac{\sum d_i^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1) \sum b_i d_i}{n-1}$$

ESTIMATED STANDARD DEVIATION
OF POPULATION OF RATIOS (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum b_i^2$	Sum of squares of sample book values	Exhibit 12, Row 6	116,719,890
2	$\sum d_i^2$	Sum of squares of sample differences	Exhibit 12, Row 12	1,937,624
3	$(\hat{R}-1)^2$		Exhibit 12, Row 17	.0004
4	$(\hat{R}-1)^2 \sum b_i^2$		Row 1 x Row 3	46,688
5	$\sum b_i d_i$	Sum of cross products	Exhibit 12, Row 15	105,627
6	$2(\hat{R}-1)$		Exhibit 12, Row 18	.04
7	$2(\hat{R}-1)\sum b_i d_i$		Row 5 x Row 6	4,225
8	$\sum d_i^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1)\sum b_i d_i$		Row 2 + Row 4 - Row 7	1,980,087
9	n-1	Sample size minus one	Exhibit 10, Row 11	210
10	$S^2_{R_j}$		Row 8 ÷ Row 9	9,429
11	S_{R_j}	Estimated standard deviation of population of ratios	Square root Row 10	97.1

$$A = \frac{S_{R_j} \cdot U_R \cdot N}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

COMPUTATION OF PRECISION

Row	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	
2	R	Reliability	Exhibit 2	
3	U_R	Reliability factor	Appendix 2	
4	n	Sample size	Exhibit 10, Row 11	
5	$\sqrt{n}$			
6	S_{R_j}	Estimated standard deviation of population of ratios	Exhibit 14, Row 11	
7	$\frac{S_{R_j} \cdot U_R \cdot N}{\sqrt{n}}$	Precision (with replacement)	(Row 1 x Row 3 x Row 6) ÷ Row 5	
8	$\frac{n}{N}$	Adjustment factor	Row 4 ÷ Row 1	
9	$1 - \frac{n}{N}$		1.0 - Row 8	
10	$\sqrt{1 - \frac{n}{N}}$			
11	A	Precision	Row 7 x Row 10	

EXHIBIT 16
PQR EMPORIUM

$$A = \frac{S_{R_j} \cdot U_R \cdot N}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

COMPUTATION OF PRECISION (ANSWERS)

Row	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	5,000
2	R	Reliability	Exhibit 2	95%
3	U_R	Reliability factor	Appendix 2	1.96
4	n	Sample size	Exhibit 10, Row 11	211
5	$\sqrt{n}$			14.5
6	S_{R_j}	Estimated standard deviation of population of ratios	Exhibit 14, Row 11	97.1
7	$\frac{S_{R_j} \cdot U_R \cdot N}{\sqrt{n}}$	Precision (with replacement)	(Row 1 x Row 3 x Row 6) ÷ Row 5	65,626
8	$\frac{n}{N}$	Adjustment factor	Row 4 ÷ Row 1	.04
9	$1 - \frac{n}{N}$		1.0 - Row 8	.96
10	$\sqrt{1 - \frac{n}{N}}$			.98
11	A	Precision	Row 7 x Row 10	64,313

FINAL ESTIMATE OF POPULATION TOTAL

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\hat{R}$	Estimated ratio	Exhibit 11, Row 16	
2	B	Total book value	Exhibit 1	
3	$\hat{X} = \hat{R}B$	Estimated total audited value	Row 1 x Row 2	
4	A	Achieved precision at desired relia-bility	Row 11, Exhibit 15	
5	$\hat{X} - A$	Lower precision limit	Row 3 - Row 4	
6	$\hat{X} + A$	Upper precision limit	Row 3 + Row 4	

FINAL ESTIMATE OF POPULATION TOTAL (ANSWERS)

Row	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\hat{R}$	Estimated ratio	Exhibit 11, Row 16	1.02
2	B	Total book value	Exhibit 1	2,500,000
3	$\hat{X} = \hat{R}B$	Estimated total audited value	Row 1 x Row 2	2,550,000
4	A	Achieved precision at desired reliability	Row 1, Exhibit 15	64,000 (rounded)
5	$\hat{X} - A$	Lower precision unit	Row 3 - Row 4	2,486,000
6	$\hat{X} + A$	Upper precision unit	Row 3 + Row 4	2,614,000

Description of Problem

The MNO Corporation has 15,150 accounts receivable with trial balance totaling $29,000,000. A total of 150 accounts are in excess of $10,000. The trial balance of these accounts totals $4,000,000.

The auditor has learned from prior examinations that many of the accounts are subject to disputes as to the quality and condition of the merchandise delivered and that disputed amounts are normally settled in favor of the customer.

In last year's audit, all accounts showing balances over $10,000 were examined 100%. The remaining accounts were stratified into 2 strata - under $2,000, and between $2,000 and $10,000. Using the techniques of Volume 3, a sample of about 800 was selected, and this produced a precision of about $900,000.

In anticipation of this year's examination, the auditors made an analysis of the 15,000 accounts below $10,000. Using the same strata boundaries as in previous years, the 13,000 items in the lower stratum showed an average balance of $1,000 with a standard deviation of $600, while the 2,000 items in the upper stratum showed an average of $6,000 with a standard deviation of $2,400. Calculations using the techniques of Volume 3 showed that a sample of 800 (500 in the lower stratum and 300 in the upper stratum) would produce a precision of about $900,000 provided the audited population showed about the same variability as the book population.

However, because of the financial position of the company, the auditors believe that the desired precision should be no greater than $600,000. This conclusion is based on the judgment that $1,200,000 would represent a material amount. Furthermore, because the number of accounts showing differences had been running about 12% in previous years, the auditors believe that better precision can be achieved if they use ratio estimation.

The total sample of 800 can be divided into 10 subsamples of 80 items each (50 from stratum 1 and 30 from stratum 2).

EXHIBIT 20
MNO CORPORATION

DATA SHEET

STRATUM 1

Subsample No.	Total Book Value	Total Audited Value	Total Difference
1	55,900	53,900	(2,000)
2	47,100	46,700	(400)
3	52,300	50,350	(1,950)
4	47,850	47,250	(600)
5	60,000	56,650	(3,350)
6	40,600	40,300	(300)
7	43,750	43,550	(200)
8	50,450	49,150	(1,300)
9	51,000	52,850	1,850
10	41,400	41,300	(100)
Total	490,350	482,000	(8,350)

STRATUM 2

Subsample No.	Total Book Value	Total Audited Value	Total Difference
1	174,720	172,380	(2,340)
2	199,740	195,150	(4,590)
3	169,020	166,800	(2,220)
4	183,330	180,030	(3,300)
5	194,310	189,960	(4,350)
6	162,540	161,730	(810)
7	185,700	182,370	(3,330)
8	166,530	165,180	(1,350)
9	204,030	197,970	(6,060)
10	174,450	171,180	(3,270)
Total	1,814,370	1,782,750	(31,620)

EXHIBIT 21
MNO CORPORATION

STRATIFIED RATIO ESTIMATE

SUB-SAMPLE NO.	ESTIMATED TOTAL AUDITED VALUE	ESTIMATED TOTAL BOOK VALUE	RATIO	RATIO X TOTAL BOOK VALUE
1	25,506,000	26,182,000	.97	
2				
3				
4				
5				
6				
7				
8				
9				
10				
Pooled				

EXHIBIT 22
MNO CORPORATION

STRATIFIED RATIO ESTIMATE (ANSWERS)

SUB-SAMPLE NO.	ESTIMATED TOTAL AUDITED VALUE	ESTIMATED TOTAL BOOK VALUE	RATIO	RATIO X TOTAL BOOK VALUE
1	25,506,000	26,182,000	.97	24,250,000
2	25,152,000	25,562,000	.98	24,500,000
3	24,211,000	24,864,000	.97	24,250,000
4	24,287,000	24,663,000	.98	24,500,000
5	27,393,000	28,554,000	.96	24,000,000
6	21,250,000	21,392,000	.99	24,750,000
7	23,481,000	23,655,000	.99	24,750,000
8	23,791,000	24,219,000	.98	24,500,000
9	26,939,000	26,862,000	1.00	25,000,000
10	22,140,000	22,394,000	.99	24,750,000
Pooled	24,417,000	24,845,000	.98	24,500,000

$$S_{D_j} = \sqrt{\frac{\sum_{j=1}^{n} d_j^2 - n\bar{d}^2}{n - 1}}$$

EXHIBIT 23
PQR EMPORIUM

ESTIMATED STANDARD DEVIATION
OF POPULATION OF DIFFERENCES

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum d_j$	Sum of differences	Exhibit 5, Row 3	
2	$\sum d_j^2$	Sum of squares of differences	Exhibit 5, Row 4	
3	n	Sample size	Exhibit 4	
4	$\bar{d}$	Average difference	Row 1 ÷ Row 3	
5	$(\bar{d})^2$	Square of average difference	Row 4 x Row 4	
6	$n\bar{d}^2$		Row 3 x Row 5	
7	$\sum d_j^2 - n\bar{d}^2$		Row 2 - Row 6	
8	$S_{D_j}^2$		Row 7 ÷ (n - 1)	
9	S_{D_j}	Estimated standard deviation of difference population	$\sqrt{\text{Row 8}}$	

$$s_{D_j} = \sqrt{\frac{\sum_{j=1}^{n} d_j^2 - n\bar{d}^2}{n - 1}}$$

EXHIBIT 24
PQR EMPORIUM

ESTIMATED STANDARD DEVIATION
OF POPULATION OF DIFFERENCES (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum d_j$	Sum of differences	Exhibit 5, Row 3	523
2	$\sum d_j^2$	Sum of squares of differences	Exhibit 5, Row 4	488,795
3	n	Sample size	Exhibit 4	50
4	$\bar{d}$	Average difference	Row 1 ÷ Row 3	10.46
5	$(\bar{d})^2$	Square of average difference	Row 4 x Row 4	109.4
6	$n\bar{d}^2$		Row 3 x Row 5	5,470
7	$\sum d_j^2 - n\bar{d}^2$		Row 2 - Row 6	483,325
8	$s_{D_j}^2$		Row 7 ÷ (n-1)	9,864
9	s_{D_j}	Estimated standard deviation of difference population	$\sqrt{\text{Row 8}}$	99.3

$$n = \frac{n'}{1 + \frac{n'}{N}} \text{ where } n' = \left(\frac{S_{D_j} \cdot U_R \cdot N}{A} \right)^2$$

COMPUTATION OF SAMPLE SIZE

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	
2	S_{D_j}	Estimated standard deviation of population of differences	Exhibit 24 Row 9	
3	R	Desired reliability	Exhibit 2	
4	U_R	Reliability factor	Appendix 2	
5	A	Desired precision	Exhibit 2	
6	$\frac{S_{D_j} \cdot U_R \cdot N}{A}$		(Row 1 x Row 2 x Row 4) / Row 5	
7	n'	Sample size (with replacement)	Row 6 squared	
8	$\frac{n'}{N}$	Adjustment factor	Row 7 ÷ Row 1	
9	$1 + \frac{n'}{N}$		1.0 + Row 8	
10	n	Sample size (without replacement)	Row 7 ÷ Row 9	
11	$n + .1n$	Sample size with 10% safety factor	Row 10 + (.1) (Row 10)	

$$n = \frac{n'}{1 + \frac{n'}{N}} \text{ where } n' = \left(\frac{S_{D_j} \cdot U_R \cdot N}{A} \right)^2$$

COMPUTATION OF SAMPLE SIZE (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	5,000
2	S_{D_j}	Estimated standard deviation of population of differences	Exhibit 24, Row 9	99.3
3	R	Desired reliability	Exhibit 2	95%
4	U_R	Reliability factor	Appendix 2	1.96
5	A	Desired precision	Exhibit 2	70,000
6	$S_{D_j} \cdot U_R \cdot N$		$\frac{\text{Row 1 x Row 2 x Row 4}}{\text{Row 5}}$	13.9
7	n'	Sample size (with replacement)	Row 6 squared	193
8	$\frac{n'}{N}$	Adjustment factor	Row 7 ÷ Row 1	.039
9	$1 + \frac{n'}{N}$		1.0 + Row 8	1.039
10	n	Sample size (without replacement)	Row 7 ÷ Row 9	186
11	n + .1n	Sample size with 10% safety factor	Row 10 + (.1) (Row 10)	205

$$s_{D_j} = \sqrt{\frac{\sum_{j=1}^{n} d_j^2 - n\bar{d}^2}{n - 1}}$$

EXHIBIT 27
PQR EMPORIUM

ESTIMATED STANDARD DEVIATION
OF POPULATION OF DIFFERENCES

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum d_j$	Sum of differences	Exhibit 12, Row 9	
2	$\sum d_j^2$	Sum of squares of differences	Exhibit 12, Row 12	
3	n	Sample size	Exhibit 10, Row 11	
4	$\bar{d}$	Average difference	Row 1 ÷ Row 3	
5	$(\bar{d})^2$	Square of average difference	Row 4 x Row 4	
6	$n\bar{d}^2$		Row 3 x Row 5	
7	$\sum d_j^2 - n\bar{d}^2$		Row 2 - Row 6	
8	$s_{D_j}^2$		Row 7 ÷ (n - 1)	
9	s_{D_j}	Estimated standard deviation of difference population	$\sqrt{\text{Row 8}}$	

$$s_{D_j} = \sqrt{\frac{\sum_{j=1}^{n} d_j^2 - n\bar{d}^2}{n - 1}}$$

EXHIBIT 28
PQR EMPORIUM

ESTIMATED STANDARD DEVIATION
OF POPULATION OF DIFFERENCES (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	$\sum d_j$	Sum of differences	Exhibit 12, Row 9	2,247
2	$\sum d_j^2$	Sum of squares of differences	Exhibit 12, Row 12	1,937,624
3	n	Sample size	Exhibit 10, Row 11	211
4	$\bar{d}$	Average difference	Row 1 ÷ Row 3	10.65
5	$(\bar{d})^2$	Square of average difference	Row 4 x Row 4	113.4
6	$n\bar{d}^2$		Row 3 x Row 5	23,927
7	$\sum d_j^2 - n\bar{d}^2$		Row 2 - Row 6	1,913,697
8	$s_{D_j}^2$		Row 7 ÷ (n - 1)	9,113
9	s_{D_j}	Estimated standard deviation of difference population	$\sqrt{\text{Row 8}}$	95.5

$$A = \frac{S_{D_j} \cdot U_R \cdot N}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

COMPUTATION OF PRECISION

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	
2	R	Reliability	Exhibit 2	
3	U_R	Reliability factor	Appendix 2	
4	n	Sample size	Exhibit 10, Row 11	
5	$\sqrt{n}$		Square root of Row 4	
6	S_{D_j}	Estimated standard deviation of population of differences	Exhibit 28, Row 9	
7	$\frac{S_{D_j} \cdot U_R \cdot N}{\sqrt{n}}$	Precision (with replacement)	(Row 1 x Row 3 x Row 6) ÷ Row 5	
8	$\frac{n}{N}$	Adjustment factor	Row 4 ÷ Row 1	
9	$1 - \frac{n}{N}$		1.0 - Row 8	
10	$\sqrt{1 - \frac{n}{N}}$		Square root of Row 7	
11	A	Precision	Row 7 x Row 10	

$$A = \frac{S_{D_j} \cdot U_R \cdot N}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

COMPUTATION OF PRECISION (ANSWERS)

Row	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	5,000
2	R	Reliability	Exhibit 2	95%
3	U_R	Reliability factor	Appendix 2	1.96
4	n	Sample size	Exhibit 10, Row 11	211
5	$\sqrt{n}$		Square root of Row 4	14.5
6	S_{D_j}	Estimated standard deviation of population of differences	Exhibit 28, Row 9	95.5
7	$\frac{S_{D_j} \cdot U_R \cdot N}{\sqrt{n}}$	Precision (with replacement	(Row 1 x Row 3 x Row 6) ÷ Row 5	64,545
8	$\frac{n}{N}$	Adjustment factor	Row 4 ÷ Row 1	.042
9	$1 - \frac{n}{N}$		1.0 - Row 8	.957
10	$\sqrt{1 - \frac{n}{N}}$		Square root of Row 9	.98
11	A	Precision	Row 7 x Row 10	63,254

FINAL ESTIMATE OF POPULATION TOTAL

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	
2	$\overline{d}$	Estimated average difference	Exhibit 28, Row 4	
3	$\hat{D}$	Estimated total difference	Row 1 x Row 2	
4	B	Total book value	Exhibit 2	
5	$\hat{X}$	Estimated total audited value	Row 4 less Row 3	
6	A	Achieved precision	Exhibit 30, Row 11	
7	$\hat{X} - A$	Lower precision limit	Row 5 minus Row 6	
8	$\hat{X} + A$	Upper precision limit	Row 5 plus Row 6	

EXHIBIT 32
PQR EMPORIUM

FINAL ESTIMATE OF POPULATION TOTAL (ANSWERS)

ROW	NOTATION	DESCRIPTION	SOURCE	RESULT
1	N	Population size	Exhibit 2	5,000
2	$\bar{d}$	Estimated average difference	Exhibit 28, Row 4	10.70
3	$\hat{D}$	Estimated total difference	Row 1 x Row 2	53,250
4	B	Total book value	Exhibit 2	2,500,000
5	$\hat{X}$	Estimated total audited value	Row 4 less Row 3	2,553,000 (rounded)
6	A	Achieved precision	Exhibit 30 Row 11	63,000 (rounded)
7	$\hat{X} - A$	Lower precision limit	Row 5 minus Row 6	2,490,000
8	$\hat{X} + A$	Upper precision limit	Row 5 plus Row 6	2,616,000

EXHIBIT 33
PQR EMPORIUM

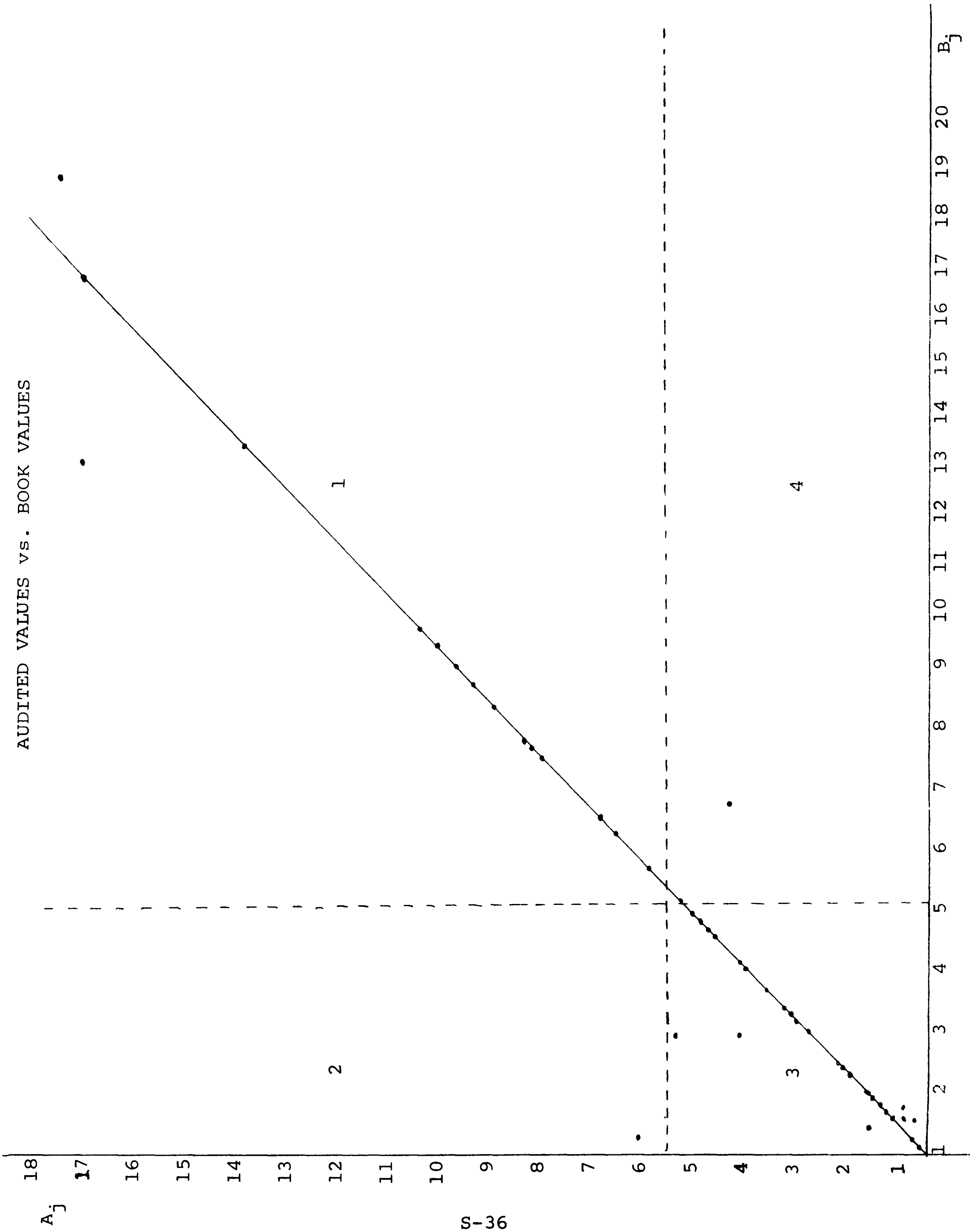

EXHIBIT 34
MNO CORPORATION

DIFFERENCE ESTIMATE BY SUBSAMPLE

SUB-SAMPLE NO.	STRATUM 1 TOTAL ESTIMATED DIFFERENCE $N_1 \bar{d}_1$	STRATUM 2 TOTAL ESTIMATED DIFFERENCE $N_2 \bar{d}_2$	TOTAL $N_1\bar{d}_1 + N_2 \bar{d}_2$
1	(520,000)	(156,008)	(686,008)
2	(104,000)	(306,015)	(410,015)
3	(507,000)	(148,007)	(655,007)
4	(156,000)	(220,011)	(376,011)
5	(871,000)	(290,015)	(1,161,015)
6	(78,000)	(54,003)	(132,003)
7	(52,000)	(222,011)	(274,011)
8	(338,000	(90,004)	(428,004)
9	481,000	(404,020)	76,980
10	(26,000)	(218,011)	(244,011)
Pooled	(217,100)	(210,810)	(427,910)

BASIC FORMULAS

Sample ratio

$$\hat{R} = \frac{\sum_{1}^{n} a_i}{\sum_{1}^{n} b_i}$$

Estimated standard deviation of ratio population

$$S_{R_j} = \sqrt{\sum_{1}^{n} \frac{\left(a_i - \hat{R}b_i\right)^2}{n - 1}}$$

computational form

$$S_{R_j} = \sqrt{\frac{\sum_{1}^{n} d_j^2 + (\hat{R}-1)^2 \sum_{1}^{n} b_i^2 - 2(\hat{R}-1)\sum_{1}^{n} b_i d_i}{n - 1}}$$

Sample size

with replacement

$$n' = \left(\frac{S_{R_j} \cdot U_R \cdot N}{A}\right)^2$$

without replacement

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

Estimate of total audited value $\hat{X} = \hat{R}\,B$

Precision of $\hat{X} = \hat{R}\ B$

$$A = \frac{S_{R_j} \cdot U_R \cdot N}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

Combined sample ratio

$$\hat{R}_c = \frac{\sum N_i\ \overline{a}_i}{\sum N_i\ \overline{b}_i}$$

(N_i represents the number of items in the i^{th} stratum, $\bar{a}_i$ represents the sample average audited value in the i^{th} stratum, $\bar{b}_i$ represents the sample average book value in the i^{th} stratum.)

Precision of $\hat{X} = \hat{R}_c \cdot B$

$$A = U_R \sqrt{\sum \frac{N_i \left(N_i - n_i\right)}{n_i} \cdot S_i^2}$$

where

$$S_i^2 = \frac{\sum_{j=1}^{n_i} \left[\left(a_{ij} - \bar{a}_i\right) - \hat{R}_c \left(b_{ij} - \bar{b}_i\right)\right]^2}{n_i - 1}$$

Separate ratio estimate of the total audited value

$$\hat{X} = \sum \hat{R}_i B_i$$

where $\hat{R}_i$ represents the sample ratio in the i^{th} stratum and B_i represents the book value in the i^{th} stratum

Precision of $\hat{X} = \sum \hat{R}_i B_i$

$$A = U_R \cdot \sqrt{\frac{\sum N_i \left(N_i - n_i\right)}{n_i} \cdot S_i^2}$$

where

$$S_i^2 = \frac{\sum_{j=1}^{n_i} \left[\left(a_{ij} - \bar{a}_i\right) - \hat{R}_i \left(b_{ij} - \bar{b}_i\right)\right]^2}{n_i - 1}$$

Computational form

$$S_i^2 = \frac{\sum_{j=1}^{n_i} d_{ij}^2 + (\hat{R}_i - 1)^2 \sum_{j=1}^{n_i} b_{ij}^2 - 2(\hat{R}-1) \sum_{j=1}^{n_i} b_{ij} d_{ij}}{n_i - 1}$$

Sample average difference

$$\bar{d} = \sum_{i=1}^{n} \frac{d_i}{n}$$

Estimated standard deviation of the difference population

$$S_{D_j} = \sqrt{\frac{\sum_{j=1}^{n} (d_j - \bar{d})^2}{n - 1}}$$

Computational form

$$S_{D_j} = \sqrt{\frac{\sum_{j=1}^{n} d_j^2 - n\bar{d}^2}{n - 1}}$$

Sample size

with replacement

$$n' = \left(\frac{S_{D_j} \cdot U_R \cdot N}{A} \right)^2$$

without replacement

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

Estimate of total difference $\hat{D} = N \cdot \bar{d}$

Estimate of total audited value $\hat{X} = B + \hat{D}$

Precision of $\hat{X} = B + \hat{D}$

$$A = \frac{S_{D_j} \cdot U_R \cdot N}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

Estimated total difference in stratified sampling

$$\hat{D} = \sum N_i \bar{d}_i$$

where N_i represents the number of items in the i^{th} stratum and $\bar{d}_i$ represents the sample average difference in the i^{th} stratum

Estimated standard deviation of differences in i^{th} stratum

$$S_i = \sqrt{\sum_{j=1}^{n_i} \frac{(d_{ij} - \bar{d}_i)^2}{n_i - 1}}$$

Computational form

$$S_i = \sqrt{\sum_{j=1}^{n_i} \frac{d_{ij}^2 - n_i \bar{d}_i}{n_i - 1}}$$

Sample size

$$n = \frac{U_R^2}{A^2} \sum \frac{N_i S_i^2 (N_i - P_i n)}{P_i}$$

where

$$P_i = \frac{N_i S_i}{\sum N_i S_i}$$

Precision of $\hat{X} = B + \hat{D}$

$$A = U_R \sqrt{\sum \frac{N_i S_i^2 (N_i - n_i)}{n_i}}$$

Conversion of Reliability Percentages to U_R Values

Desired R	U_R
65%	.94
70%	1.04
75%	1.15
80%	1.28
85%	1.44
90%	1.64
95%	1.96
98%	2.33
99%	2.58

SUMMARY OF VOLUME 5

Chapter 1

Ratio estimation is appropriate when:

1. There is a book value for each population item.
2. The total book value is known and corresponds to the addition of all the individual book values.
3. There are some observed differences between audited and book values.
4. The audited values are nearly proportional to the book values.

The estimated ratio is formed by dividing the total sample audited value by the total sample book value when using unrestricted random sampling. In symbols,

$$\hat{R} = \frac{\sum a_i}{\sum b_i}$$

To obtain the estimated total audited value, the estimated ratio is extended by the total book value. This is expressed symbolically by

$$\hat{X} = \hat{R}B$$

The ratio estimate of the total audited value will generally have better precision than the mean estimate. This occurs because the standard deviation of the ratio population is smaller than the standard deviation of the population of dollar values.

Chapter 2

To determine an appropriate sample size for a specified precision and reliability, a preliminary sample may be selected, using a random number table exactly as in previous volumes. The primary purpose of the preliminary sample is to estimate the standard deviation of the ratio population.

The preliminary sample should consist of at least 50 items. This number is generally adequate if the incidence of difference between book and audited values exceeds 20%. When the incidence of difference lies between 5% and 20%, 80-100 items should be selected in the preliminary sample. When the incidence of difference is smaller than 5%, the size of a preliminary sample may be even larger than 100 in order to observe some differences.

The calculation of the estimated standard deviation of the ratio population is simplified by using the difference for each item in the sample. This difference is defined by $d_i = a_i - b_i$: the difference in the i^{th} sample item is its audited value minus its book value. Following the steps in Exhibit 7, you then calculate the estimate by means of the formula

$$S^2_{R_j} = \frac{\sum d_i^2 + (\hat{R}-1)^2 \sum b_i^2 - 2(\hat{R}-1) \sum b_i d_i}{n-1}$$

Chapter 3

Once the estimated standard deviation of the ratio population is available, and given the desired precision and reliability, the required sample size is calculated by following the steps in Exhibit 9. The formula used in this exhibit is

$$n = \frac{n'}{1 + \frac{n'}{N}} \quad \text{where } n' = \left(\frac{S_{R_j} \cdot U_R \cdot N}{A} \right)^2$$

In this formula, n' represents the sample size required when sampling on a with replacement basis. This is the same formula used in Volume 1 with S_{R_j} taking the place of S_{X_j}. The extra calculation involved in computing n provides for the fact that sampling is done without replacement. If the calculated value of n' is smaller than 5% of the population size N, these extra calculations can be omitted. In addition, the exhibit adds an extra 10% to serve as a reminder that the results are only approximate.

Chapter 4

After an audited value has been established for each item in the sample, we estimate the total audited value of the population and compute the achieved precision of this estimate at the desired reliability level.

To do this we first combine the data from the preliminary sample with that obtained from the additional sample elements. The estimated ratio is computed by dividing the total sample audited value by the total sample book value. In symbols,

$$\hat{R} = \frac{\sum a_i}{\sum b_i}$$

This estimated ratio is then extended by the known total book value to produce the estimated total audited value. In symbols,

$$\hat{X} = \hat{R}B$$

To compute the precision of this estimate, it is necessary to obtain an estimate of the standard deviation of the ratio population based on the entire sample. This involves using the same formula as used in Chapter 2 for the preliminary sample. Thus, Exhibit 13 is the same as Exhibit 7.

The calculation of the achieved precision is accomplished by following the steps in Exhibit 15. The formula for the achieved precision is

$$A = \frac{S_{R_j} \cdot U_R \cdot N}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

Except for the factor $\sqrt{1 - \frac{n}{N}}$, this is the same formula used in Volume 1 with S_{R_j} taking the place of S_{X_j}, the estimated standard deviation of the (audited) dollar values. The factor $\sqrt{1 - \frac{n}{N}}$ occurs because the sampling is done without replacement.

The precision interval with lower bound $\hat{X}$ - A and upper bound $\hat{X}$ + A will contain the total audited value at the specified reliability level. If the book value is contained in the precision interval, we can say that the statistical evidence supports fair statement with precision A and reliability R. In other words, with reliability R, the maximum possible difference between the book value and audited value is twice the precision.

If the book value does not fall within the precision interval, the statistical evidence does not support fair statement with precision A and reliability R. This means that potentially the difference between the book and audited values could be larger than twice the precision. Consequently additional audit evidence would be required to support fair statement -- or to determine that there is indeed some material misstatement.

Chapter 5

To use stratification with ratio estimation it is necessary to

1. Choose a basis for stratification that is related to the size of the ratio. The book values will often serve this purpose.

2. Select the number of strata. Usually three to five with the top stratum sampled 100% will be reasonable.

3. Select the stratum boundaries. One reasonable procedure is to select boundaries so as to equalize the dollars within each stratum (except the top stratum).

If feasible, you may use a preliminary sample from each stratum to obtain an estimate of the stratum standard deviation of ratios. These estimates may be used in place of the S_i of Volume 3 in computing an appropriate sample size and allocation of the sample to the strata. Exhibits 8, 9 and 11 of Volume 3 may be used for this.

If it is not feasible to audit a preliminary sample, you could examine the book values of a sample from each stratum, compute the estimated standard deviation of book values, and use these estimates instead of the S_i values in Volume 3, Exhibits 8, 9 and 11. Beside providing a conservative sample size, this procedure has the advantage of assuring the representativeness of the sample as reflected by the distribution of book values.

Another alternative would be to choose a sample size without considering stratification; i.e., use the procedures outlined in Chapter 2 and then allocate the resulting sample to the strata in proportion to the dollar amount in the stratum. If the strata have nearly equal dollars, this would result in approximately equal sample size for each stratum.

After audited values have been determined for each sample item, two ratio estimates of the total audited value are possible -- a separate estimate and a combined estimate.

The separate ratio estimate is appropriate when the ratios vary greatly from stratum to stratum. As the name implies, it is formed by computing an estimated ratio for each stratum separately, extending each of the ratios by the total book value in the stratum, and adding the results over all the strata. To compute the precision of the separate ratio estimate, you need an estimate of the standard deviation of the ratio population in each stratum. As a working rule, this requires at least 30 observations in each stratum as well as some observed differences in each stratum.

The computational steps involved in obtaining the precision of the separate ratio estimate are rather formidable and hence are not presented in the chapter. The formula appears in Appendix 1.

In auditing,the combined ratio estimate is used more often than the separate ratio estimate -- both because the ratios tend to vary little over the strata, and the data requirements are less stringent. As a working rule, it is suggested that you have a total of 50 observations and some observed differences when using the combined ratio estimate.

As the name implies, the combined ratio estimate is formed by combining the data from each stratum to form a single ratio which is then extended by the total book value to obtain the combined ratio estimate of the total audited value.

The combined ratio is equal to the estimated total audited value divided by the estimated total book value. These estimates are formed in exactly the same way as the estimated total value was formed in Volume 3 (Exhibit 7) using mean estimation. In symbols, if there are 3 strata:

$$\hat{R}_c = \frac{N_1\bar{a}_1 + N_2\bar{a}_2 + N_3\bar{a}_3}{N_1\bar{b}_1 + N_2\bar{b}_2 + N_3\bar{b}_3}$$

Where N_i equals the number of items in the i^{th} stratum, $\bar{a}_i$ = sample average audited value within the i^{th} stratum, and $\bar{b}_i$ = sample average book value within the i^{th} stratum.

The formula for computing the precision of this estimate is not included in this chapter, but may be found in Appendix 1.

Since the formula for the standard deviation of either of these estimates is complicated, we present an alternative procedure known as replicated sampling. This procedure consists of dividing the total sample into several subsamples (usually 8 - 10), so that each subsample is a replication of the sampling procedure. In other words, if a stratified sampling procedure calls for 120 items in stratum 1 and 80 items in stratum 2, then each of the ten subsamples would have 12 items from stratum 1 and 8 from stratum 2.

For each subsample, you calculate an estimated ratio -- either separate or combined -- then compute the range of these estimates (largest minus smallest), and divide this range by the number of subsamples. This number when extended by the reliability factor serves as the precision of the estimate. The estimate itself is calculated on the basis of the whole sample.

Chapter 6

Difference estimation is appropriate when:

1. There is a book value for each population item.
2. There are some observed differences between audited and book values.
3. The total book value is known and corresponds to the addition of all the individual book values.
4. The audited values are not proportional to the book values, or similarly, the differences are not proportional to the book values.

In using unrestricted random sampling without replacement, the estimate of the total population difference is formed by extending the sample average difference by the number of items in the population. In symbols,

$$\hat{D} = N\bar{d}, \text{ where } \bar{d} = \frac{1}{n}\sum_{1}^{n} d_i \text{ and } d_i = a_i - b_i$$

To determine an appropriate sample size, a preliminary sample may be used if it is large enough to observe serveral differences. Preliminary samples of 50 will suffice if the incidence of difference is at least 20%.

The formula for estimating the standard deviation of the difference population is exactly the same as that used in Volume 1, with $\bar{d}_j$ taking the place of x_j and $\bar{d}$ taking the place of $\bar{x}$.

In symbols,

$$S_{D_j} = \sqrt{\frac{\sum_{j=1}^{n} (d_j - \bar{d})^2}{n - 1}} = \sqrt{\frac{\sum_{j=1}^{n} d_j^2 - n\bar{d}^2}{n - 1}}$$

The necessary sample size to achieve a precision A at reliability R is given by the formula

$$n = \frac{n'}{1 + \frac{n'}{N}} \text{ where } n' = \left(\frac{S_{D_j} \cdot U_R \cdot N}{A}\right)^2$$

The formula for n' is the same formula used in Volume 1, with S_{D_j} taking the place of S_{X_j}. The additional steps to obtain n are required to take account of sampling without replacement.

After the sample has been selected and audited values established for each of the sample items, the estimated total difference is calculated by means of the formula

$$\hat{D} = N\bar{d},$$

where N represents the number of items in the population and $\bar{d}$ represents the average sample difference. For each item in the sample, the difference equals the audited value minus the book value.

The estimated total audited value can be obtained by adding the estimated total difference to the known book value. In symbols,

$$\hat{X} = B + \hat{D}$$

The precision of this estimate is calculated by means of the formula

$$A = \frac{S_{D_j} \cdot U_R \cdot N}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

Except for the factor $\sqrt{1 - \frac{n}{N}}$, this formula is similar to that used in Volume 1 with S_{D_j} replacing S_{X_j}. The factor $\sqrt{1 - \frac{n}{N}}$ takes account of sampling without replacement. This formula is also the same as used in Chapter 4 with S_{D_j} replacing S_{R_j}.

Difference estimation may be used with stratified random sampling. The basis for stratification should be some known quantity or quality that is related to the magnitude of the differences in the population. Frequently, the size of the book amount is used as the basis.

In manual application, three to five strata are commonly employed, with stratum boundaries selected so that each stratum contains nearly an equal dollar amount -- except the top stratum if it is sampled 100%. Preliminary samples may be chosen from each stratum to provide estimates of the standard deviation of differences. These estimates take the place of the S_i in Exhibits 8, 9, and 11 of Volume 3, which may be used to choose an appropriate sample size and allocation of the sample to the strata.

When it is not feasible to audit a preliminary sample, a sample of book values may be selected from each stratum and used to estimate the standard deviation of the book value population with each stratum. These estimates take the place of the S_i in Exhibits 8, 9, and 11 of Volume 3 and can be used to select a sample size and allocation. The resulting sample size will be larger than required, but does provide the advantage of being representative of the book value population.

After the sample has been selected and audited values established, the estimated total difference is calculated in the following formula, assuming three strata:

$$\hat{D} = N_1\bar{d}_1 + N_2\bar{d}_2 + N_3\bar{d}_3,$$

where N_i represents the number of items in the i^{th} stratum and $\bar{d}_i$ represents the average sample difference in the i^{th} stratum, $i = 1, 2$ and 3.

To compute the precision of this estimate, you can follow the procedures of Volume 3 substituting the differences for the dollar values, or use the approximation involving subsamples outlined in Chapter 5.