Last name, first name: ____________________________. Student #: ________________

STA 304H1 F SUMMER 2011, Second Test, June 16 (20%)
Duration: 60 min. Allowed: nonprogrammable hand-calculator, aid-sheet, one side, with theoretical formulas only; the test contains 3 pages, + an empty page; please check

[45] 1) A city is divided into 600 blocks. A preliminary SRS of size 10 was selected from the city blocks and the following data was obtained (x - number of houses with finished basement, y - number of houses with rented basement, z – total number of houses in the block; it is assumed that an unfinished basement cannot be rented):

var	1	2	3	4	5	6	7	8	9	10	Σ var	Σ var^2
x	8	9	4	2	0	5	3	0	4	3	38	224
y	2	3	2	0	0	2	0	0	1	1	11	23
z	12	15	10	8	6	20	5	12	8	9	105	1283

$\Sigma xy = 68$

(a) [5] Estimate the total number of houses in the city,
(b) [5] Estimate the total number of houses with finished basement in the city,
(c) [5] Estimate the total number of rented basements in the city,
(d) [5] Estimate the proportion of houses with finished basement (out of all houses).
(continued)

Solutions:
From the table $\bar{x} = 3.8$, $\bar{y} = 1.1$, $\bar{z} = 10.5$.

(a) $\hat{\tau}_z = N\bar{z} = 600 \text{ x } 10.5 = 6300$, **[5]**

(b) $\hat{\tau}_x = N\ \bar{x} = 600 \text{ x } 3.8 = 2280$, **[5]**

(c) $\hat{\tau}_y = N\ \bar{y} = 600 \text{ x } 1.1 = 660$, **[5]**

(d) $\hat{R}_{x/z} = 38/105 = 0.362 = 36.2\%$, **[5]**

Oxdia @ http://www.oxdia.com
This item is shared by the uploader to help you in your studies.
It is copyrighted by the creator (copyright owner) of the content.
Distribution is prohibited without permission from the copyright owner.
Solution (if any) is NOT audited, so use at your discretion.

(e) [15] Estimate the proportion of finished basements rented and the standard deviation of that estimator.

(f) [10] It was found that the total number of houses in the city is 6200. Use this information to again estimate the total number of finished basements in the city. Do you expect this estimator be better than one in (b)? How can you check it? Just explain, don't do any calculation.

Solutions:

(e) $\hat{R}_{y/x}$ = 11/38 = 0.289 = 28.9%, **[5]**

$S_r^2 = \sum (y_i - rx_i)^2 / (n-1)$ = (23 – 2 x 0.289 x 68 + 0.289^2 x 224)/9 = 0.267, **[5]**

$\hat{V}ar(\hat{R}) = \frac{N-n}{N} S_r^2 / (n\bar{x}^2)$ = (600-10)/600 x 0.267/(10 x 3.8^2) = 1.82 x 10^{-3},

$\hat{S}d(\hat{R}) = \sqrt{1.82 \times 10^{-3}}$ = 0.043. **[5]**

(f) You may use the ratio estimator $\hat{\tau}_x = \hat{R}_{x/z}\tau_z$ = (38/105) x 6200 = 2243.8 = 2244. **[5]**
Yes, we expect this ratio estimator be better than one in (b), due to correlation between the number of basements and the block size (number of houses) **[3].** To check it, we just may calculate their variances. **[2]**

Test ID: 4075 Downloader ID: 8669

Oxdia @ http://www.oxdia.com

This item is shared by the uploader to help you in your studies.

It is copyrighted by the creator (copyright owner) of the content.

Distribution is prohibited without permission from the copyright owner.

Solution (if any) is NOT audited, so use at your discretion.

[55] 2) A students' community consists of 24 households. The households are divided into two strata, by household size. Stratum I consists of households with size 1-5 students (17 households), and stratum II of households of size 6 and more students (7 households). A stratified random sample of size 5 was selected from the community, and the data on daily income and daily food cost were recorded, as follows:

Household	Stratum	Size, z	Daily income, y ($\times$\$10)	Daily food cost, x ($\times$\$10)
1	1	3	33	21
2	1	4	39	24
3	1	3	35	22
4	2	6	62	31
5	2	6	61	29

(a) [18] Estimate the
 (i) total size of the community (number of students),
 (ii) total daily income in the community, and
 (iii) total daily food cost.

(b) [9] Place a bound on the error of estimation of the total daily food cost.

(continued)

Solutions:

(a) N = 24, N_1 = 17, N_2 = 7, n = 5, n_1 = 3, n_2 = 2.

(i) Total size: $\hat{\tau}_z = N_1\bar{z}_1 + N_2\bar{z}_2 = 17\times 3.33 + 7\times 6 = 98.61$, **[6]**

$\bar{z}_1 = \frac{1}{3}(3+4+3) = 3.33$, $\bar{z}_2 = \frac{1}{2}(6+6) = 6$

(ii) Total daily income: $\hat{\tau}_y = N_1\bar{y}_1 + N_2\bar{y}_2 = 17\times 35.67 + 7\times 61.5 = 1036.84$, **[6]**

$\bar{y}_1 = \frac{1}{3}(33+39+35) = 35.67$, $\bar{y}_2 = \frac{1}{2}(62+61) = 61.5$

(iii) Total daily food cost: $\hat{\tau}_x = N_1\bar{x}_1 + N_2\bar{x}_2 = 17\times 22.33 + 7\times 30 = 589.61$, **[6]**

$\bar{x}_1 = \frac{1}{3}(21+24+22) = 22.33$, $\bar{x}_2 = \frac{1}{2}(31+29) = 30$

(b) $\hat{V}ar(\hat{\tau}_x) = \Sigma N_i^2 \frac{N_i - n_i}{N_i}\frac{S_i^2}{n_i} = 17(17-3)\frac{2.33}{3} + 7(7-2)\frac{2}{2} = 219.85$. **[7]**

$S_1^2 = 2.33, \quad S_2^2 = 2$

$B_\tau = 2\sqrt{219.85} = 29.65$. **[2]**

Oxdia @ http://www.oxdia.com
This item is shared by the uploader to help you in your studies.
It is copyrighted by the creator (copyright owner) of the content.
Distribution is prohibited without permission from the copyright owner.
Solution (if any) is NOT audited, so use at your discretion.

(c) [12] (i) Estimate the average daily income per student.
(ii) Estimate the percentage of income spent on food in the community.
(iii) Are these estimators in (i) and (ii) biased, or unbiased? Explain.

(d) [6] Would a stratified sample with proportional allocation be better in this survey than an SRS (assume that the population is big)? Explain.

(e) [10] (i) Is this type of stratification good, or not good for this survey on expenditures? Explain.
(ii) Would an optimal allocation of the sample produce better results than a proportional in this survey (assume that the population is big)? Explain.

Solutions:

(c) (i) Average daily income: $\hat{r}_{y/z} = \frac{\hat{\tau}_y}{\hat{\tau}_z} = \frac{1036.84}{98.61} = 10.51$. **[5]**

(ii) Percentage of food coast: $\hat{r}_{x/y} = \frac{\hat{\tau}_x}{\hat{\tau}_y} = \frac{589.67}{1036.84} = 0.5687 = 56.87\%$. **[5]**

(iii) Both estimators are ratio estimators, and then biased. **[2]**

(d) Stratified sample with proportional allocation should be better than an SRS, because of different household sizes in two strata (up to 5 students, and over 5 students), which makes <u>different mean values</u> for observed variables. **[6]**

(e) (i) The stratification should be good, because it is created by household size, which is highly correlated with variables of interest. **[5]**
(ii) It is likely that the optimal allocation will produce better results than the proportional, because the variances of variables such as income and expenditures on food will be greater in greater households (i.e., different strata). **[5]**

Test ID: 4075
Downloader ID: 8669
Oxdia @ http://www.oxdia.com
This item is shared by the uploader to help you in your studies.
It is copyrighted by the creator (copyright owner) of the content.
Distribution is prohibited without permission from the copyright owner.
Solution (if any) is NOT audited, so use at your discretion.