University of Waterloo

STAT 332 – W. H. Cherry

Tutorial 6

T6 – 1. The circulation manager of a newspaper wishes to estimate the average number of newspapers purchased per household in a particular community. Travel costs from household to household are substantial so that the 4,000 households in the community are listed in 400 geographical clusters of 10 households each, and four clusters are obtained by equiprobable selecting. Interviews are conducted for each selected household and they yeild the following data for the number of newspapers purchased by each household, together with the sums and sums of squares (SS) for the 10 observations in each cluster:

CLUSTER	NUMBER OF NEWSPAPERS	Sum	SS
1	1 2 1 3 3 2 1 4 1 1	19	47
2	1 3 2 2 3 1 4 1 1 2	20	50
3	2 1 1 1 1 3 2 1 3 1	16	32
4	1 1 3 2 1 5 1 2 3 1	20	56.

(a) From these data and assuming complete response, find an approximate 90% confidence interval for the average number of newspapers purchased per household:

   (i) on the basis of the clustered sampling protocol actually used;

   (ii) ignoring the clustering and assuming that the data had been obtained by equiprobable selecting of 40 households.

(b) Comment briefly on how the intervals in (i) and (ii) compare.

T6 – 2. An approximate expression for the *relative* estimating bias of our estimator (R) of the respondent population attribute $\mathbf{R}$, the ratio of the averages of two response variates, is given at the right.

$$r.e.b.(R) \equiv \frac{e.b.(R)}{\mathbf{R}} \simeq \frac{1}{\mathbf{N}-1}\sum_{i=1}^{\mathbf{N}}\frac{\mathbf{X}_i}{\overline{\mathbf{X}}}\left(\frac{\mathbf{X}_i}{\overline{\mathbf{X}}} - \frac{\mathbf{Y}_i}{\overline{\mathbf{Y}}}\right)\left(\frac{1}{n} - \frac{1}{\mathbf{N}}\right)$$

(a) Define concisely each of the symbols in this expression.

(b) Explain briefly what is meant by *estimating bias* in this context.

(c) Briefly discuss, *in point form*, the factors which affect the magnitude of the relative estimating bias; indicate in your discussion the degree to which each factor is under the control of the investigator(s).

(d) Explain briefly why it is more useful in (c) to consider *relative* estimating bias rather than just estimating bias.

University of Waterloo   STAT 332 – W. H. Cherry

Blank page

Blank page