

Eco220 Exercise set 4.

1. The Government, interested in the per capita gross domestic product in Canada, takes a randomly selected individual from each province, records the individual's gross product and takes an average of those personal gdp's (what each individual produces) as an estimate of the per capita gdp in Canada. What sort of sample is this? Alternatively the government could have randomly selected a province, perform a census of personal gdp's in the province and take the average. What sort of sample is this? What would it have to do to form an independent random sample?

2. A programming error results in \bar{X}_e the mean of a random sample $X_i, i = 1, 2, \dots, n$ taken from a population with a true mean μ and true variance σ^2 being calculated as:

$$\bar{X}_e = \frac{\sum_{i=1}^n X_i}{n} + \frac{1}{n}$$

Is the estimator unbiased? Is it consistent? Explain.

In correcting the above error the programmer uses a new formula where only the first $n_1 < n$ observations are employed regardless of how large n is, i.e. his formula now is:

$$\bar{X}_e = \frac{\sum_{i=1}^{n_1} X_i}{n_1}, \quad (n_1 < n)$$

Is the estimator unbiased? Is it Consistent? Explain.

3. A random sample $X_i, i = 1, \dots, n$ is drawn from a population in which $E(X) = \mu$ and $V(X) = \sigma^2$. Two estimators μ^A and μ^B are proposed which are defined by the following formulae:

$$\mu^A = \frac{\sum_{i=1}^n X_i}{n}; \quad \mu^B = \frac{4X_n + \sum_{i=1}^n X_i}{n+4}$$

show that they are both unbiased and derive their variances. Which one would you prefer and why?

4. When the population has a distribution which is symmetric both the sample mean and the sample median are unbiased estimators, their respective variances are σ^2/n and $\pi\sigma^2/2n$ discuss their relative merits as estimators of the population mean.

5. A coin has a probability of .6 of turning up heads but its owner does not know this. To estimate

the chance of it turning up heads he proposes to toss the coin n times and count the proportion p_n of times it comes up heads.

- i) Define the sample space when $n=3$ (i.e. list the set of basic outcomes) together with the associated values of p_n .
- ii) Calculate $E(p_n)$ and $V(p_n)$.
- iii) Repeat i) and ii) for $n=2$ and comment on the relationship between the sample size n and the expected value and variance of his estimator.

6. To save computing time an investigator randomly selected a proportion $n_1 / n < 1$ of the sample $X_i, i=1, \dots, n$ that she took and calculated an estimate of the true mean as:

$$\bar{X}_{n_1} = \frac{\sum_{i \in \text{selected portion}}^n X_i}{n_1}$$

Is the estimator unbiased? Is it sufficient? Why?

7. A country is split into 4 regions, the North, the South, the East, and the West the respective populations (millions) in the regions are 1, 2, 3 and 4. A person is randomly selected from each region and their waist size measured for an obesity study. Construct an estimator that would ensure an unbiased estimator of average waist size in the country.

8. The following table reports the per capita gross domestic product for the provinces of Canada together with their respective populations for the year 1991.

	Gdp per capita	Population (1000's)
Newfoundland	15878.396	578.20700
Prince Edward Island	16078.723	130.48300
Nova Scotia	19083.999	912.33500
New Brunswick	18161.636	743.21500
Quebec	22119.575	7033.3630
Ontario	26880.663	10359.231
Manitoba	20800.434	1106.2750
Saskatchewan	20637.584	1002.3460
Alberta	27726.149	2571.7960
British Columbia	24398.375	3338.4600
Yukon and NWT	32662.259	87.869000

Recalling the details of question 1 what would be the expected value of a per capita gdp estimator:

- i) Based upon a stratified sample using the provinces as the strata?
- ii) Based upon a cluster sample using the Saskatchewan as the cluster?
- iii) Based upon a simple random sample of the country?