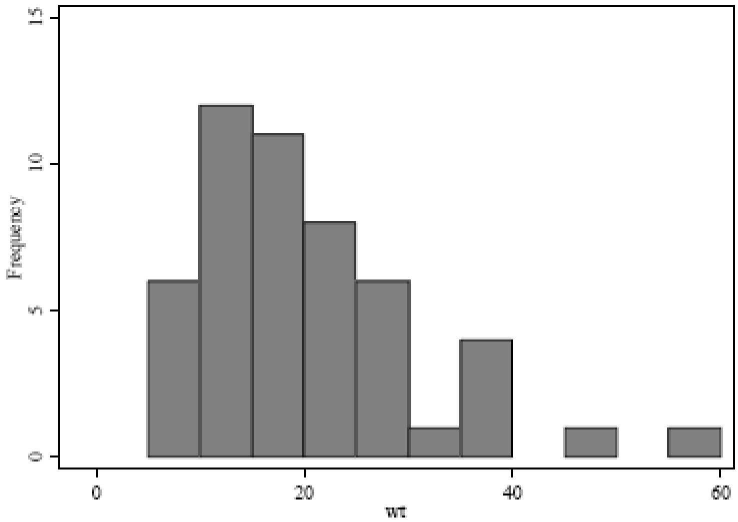
# Session 1: Practical Exercise and Commentary

1. The population:

The distribution of the weights of the 50 stones in the population is as follows:




Note that the weights do not appear to be Normally distributed. The distribution is asymmetric; it is rather skewed to the right. Some population statistics are given below.

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| wt | 50 | 19.32 | 10.34396 | 7 | 56 |

2. Desirable properties of an estimator

You have been asked to construct three samples, each of 10 stones, and hence obtain three different estimates of the mean weight in the population. The sampling processes you adopt should have the following characteristics:

a) The estimators should be unbiased. This will be reflected in sample means that are equally spread above and below the population mean. However, a sample mean well above or below the population mean may indicate bias.
b) The standard error of the mean should be small. This will be reflected in sample means that are close to the population mean.
c) Additionally, for a random sample, the estimated standard error of the mean should be close to the theoretical value [population standard deviation/($\sqrt{10}$) – for a sample size of 10].

Complete the following table to see to what extent your sampling processes satisfy these requirements.

| Sample type | Mean | | Standard deviation | | Standard error of mean | |
|---|---|---|---|---|---|---|
| | Sample | Theory | Sample | Theory | Sample | Theory |
| 1. Random | | ≈19.32 | | ≈10.34 | | ≈3.27 |
| 2. ‘Typical’ | | ≈19.32? | | ≈10.34? | | ≈3.27? |
| 3. ‘Optimal’ | | ≈19.32? | | <10.34? | | <3.27? |

**Questions**

a) Are all three of your sample means close to 19.32. Are any obviously biased?

______________________________________________________________

b) Does your ‘typical’ sample have similar characteristics to the random one?

______________________________________________________________

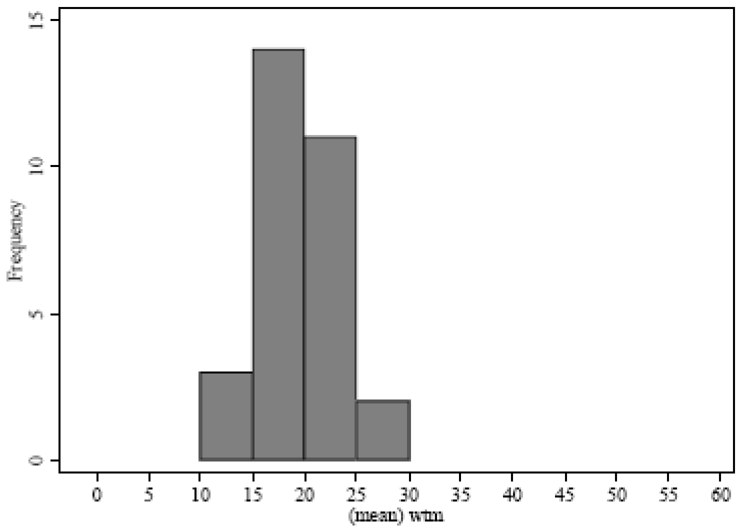
c) Has your ‘optimal’ sample resulted in a more precise estimate?

______________________________________________________________

Drawing conclusions on the basis of a single small sample is obviously unreliable. We will now consider repeated sampling – first illustrating results obtained using random and biased sampling techniques; and then considering the results obtained by you and your classmates as a whole.

3 Random sampling:

The following distribution of sample means was obtained from 30 samples each of size 10.




Summary statistics for the sample means and for the estimated standard errors of the sample means are as follows:

```
    Variable   |       Obs        Mean       Std. Dev.       Min         Max
---------------+-----------------------------------------------------------
 sample mean   |        30       19.15667    3.363463        13.3        26.4
 standard error|        30       3.241279      (omitted for clarity)
```

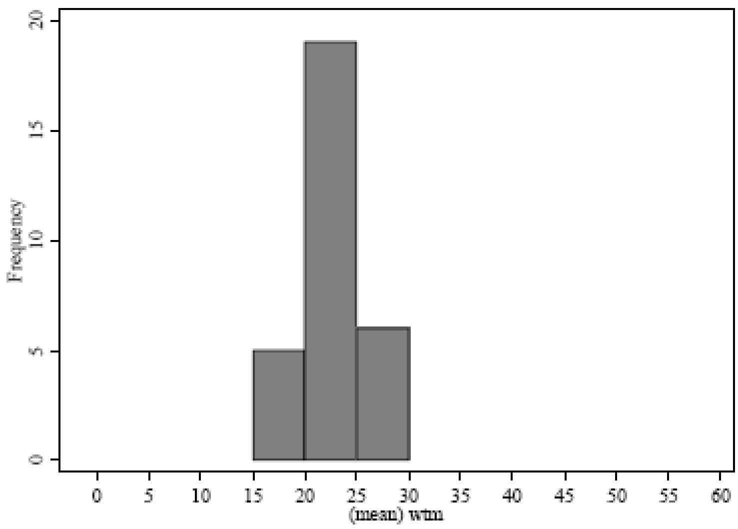
**Key results**

a) The distribution of the sample means is more Normally distributed than the individual weights. This follows from the central limit theorem.
b) The standard deviation of the sample means is less than that of the individual weights.
c) Because of random sampling the mean of the sample means (19.16) is close to the population mean (19.32).
d) The standard deviation of the sample means (3.36) is close to its theoretical value (3.27) calculated above. This demonstrates the validity of estimating the standard error as [population standard deviation/($\sqrt{10}$)].
e) The standard error of the mean can also be estimated from a single sample. The fact that this is (approximately) unbiased is illustrated by the fact that the mean of such estimates (3.24) is also close to the theoretical value (3.27) calculated above.

This final result is a very important one because, in practice, we usually only have access to a single sample. The fact that we can obtain an (approximately only – see Inference 2) unbiased estimate of the standard error of an estimator such as a mean from a *single sample* is the key to procedures such as significance tests and constructing confidence intervals etc..

4. Biased sampling:

The following distribution of sample means was obtained from 30 samples each of size 10. Samples were selected in such a way that heavier stones were more likely to be selected than lighter ones.




Summary statistics for the sample means and estimated standard errors of the sample means are as follows:

```
   Variable   |       Obs        Mean       Std. Dev.       Min         Max
--------------+--------------------------------------------------------------
 sample mean  |        30       22.69667    2.435299        17.8        27.4
 standard error|       30       3.778333         (omitted for clarity)
```

**Key results:**

a) Because of biased sampling the mean of the sample means (22.70) lies above and not very close to the population mean (19.32).

b) The mean of the estimated standard errors of the mean (3.78) is larger than before and not close to the standard deviation of the sample means (2.44). The reason for this difference is that almost all of the samples included both of the stones with outlying values. The inclusion of these two stones results in large within sample standard deviations and hence large estimates of the standard error of the mean when estimated from a single sample. In addition the fact that almost all of the samples contained these two stones has tended to make the sample means more similar one to another (hence with a smaller standard deviation) than is the case with random sampling.

These results illustrate the dangers of attempting to draw conclusions from non-random samples.

5. <u>'MSc Medical Statistics' sampling</u>

For the three types of sample drawn by the class, write down the following results which will be calculated for you: the mean and standard deviation of the sample means together with the mean of the estimated standard errors. Comment on the results.

**Random samples**

Mean of sample means ____________
SD of sample means (*observed* standard error using repeat samples) ___________
Mean of estimated standard errors of means (each calculated using a single sample) _________
Comments

**'Typical' samples**

Mean of sample means ____________
SD of sample means (*observed* standard error using repeat samples) ___________
Mean of estimated standard errors of means (each calculated using a single sample) _________
Comments

**'Optimal' samples**

Mean of sample means ____________
SD of sample means (*observed* standard error using repeat samples) ___________
Mean of estimated standard errors of means (each calculated using a single sample) _________
Comments

# A couple of issues swept under the carpet

### *1 Population variance*

On page 1 of these solutions, the summary printout describing the population of 50 stones gives their mean weight as 19.32 and their standard deviation as 10.34.

In fact, the summary was a standard Stata command, given a dataset of the stone weights, and for this summary Stata – quite naturally – uses the standard deviation formula that is appropriate for a sample (with denominator 1/(n-1)). As you will see in Inference session 2, this formula is appropriate for estimating an unknown population standard deviation – but it's not appropriate for calculating the known standard deviation of a population, where values from each member of the population are known (denominator 1/n should be used). In fact, the correct standard deviation of the population of 50 stones is 10.24.

This was not a deliberate mistake: I only realized after I'd given the practical for a couple of years. I've left it in to show how easy an error it is to make – though we very rarely have the opportunity to make it, since we almost always, as inferential (rather than descriptive) statisticians, deal with samples rather than populations.

### *2 Sampling with/without replacement [not examinable]*

In the practical we sampled without replacement. When a population is very large, it makes negligible difference if a relatively small sample is sampled with or without replacement. However, when a population is not large relative to the sample size – as is the case here – sampling without replacement induces a lack of independence between sampled items: the probability of picking a certain stone will depend on how many stones have been picked before it, and whether it has already been picked. To analyse sampling with replacement correctly, the relative sizes of sample and population need to be taken into account. This is discussed in Rice (see references at front of manual in acnowledgements) p207 (paperback) or p193 (hardback).