Appendix 2: Standard deviation and confidence limits

The frequencies of measurements from a large sample from a population can be graphed as the common bell-shaped curve, or a normal curve, shown in Fig. A2.1. Plotted heights from a large sample of adult human males, for example (shown at the left of Fig. A2.1), suggest that most individuals are around the mean height of 1.7 m and that a few very short and very tall individuals are at the lower and upper tails of the graph respectively. From this graph, the probability of a single individual (chosen at random) having a height between 1.8 and 1.9 m is indicated by the shaded area under the curve.

The spread (or dispersal) of values around the mean (the 'width' of the normal curve) can be measured by the standard deviation. In the normal curve on the right-hand side of Fig. A2.1, the shaded areas under the curve show that 68.26% of all values fall between one standard deviation (σ) of the mean (μ). Similarly, 95% of all values lie between the limits of $\mu \pm 1.96\ \sigma$.

The standard deviation of the mean of n measurements from a population with a standard deviation of σ is $\sigma/\sqrt{n}$. In the same way that we used the sample mean as an estimate of the population mean, we can use the sample standard deviation (s) to estimate the population standard deviation (σ). Conventionally, the standard deviation of a mean is referred to as its standard error (SE) and is computed as $s/\sqrt{n}$. It is possible to attach confidence limits at a chosen level of probability to the estimate of the true mean. At the 0.95 level of probability (95% confidence limits) we would compute $\bar{x} \pm 1.96 \times$ SE and be able to say that the true mean has a probability of 0.95 of falling within these limits and conversely a probability of 0.05 of falling outside. That is, we are 95% confident that the true mean lies between these limits.

However, the use of the value 1.96 for 95% confidence limits applies to large samples (>30). For smaller samples, it is necessary to use tables to find a value to substitute for 1.96. The confidence limits are then calculated as $\bar{x} \pm t \times$ SE where the value of t is taken from the table given in Appendix 7. The values in the table include an

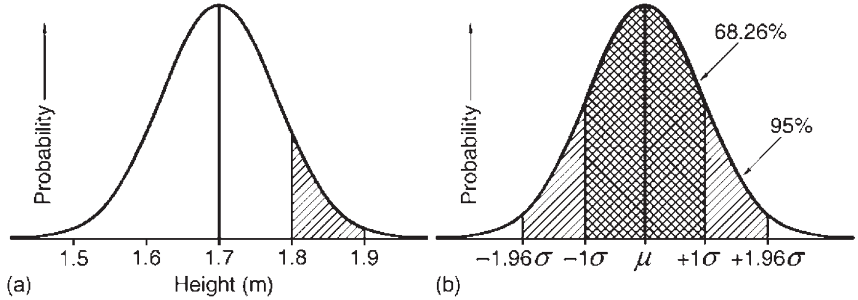

Fig. A2.1 Normal distributions.

allowance for the uncertainty involved in estimating the population standard deviation from a small sample.

The sea cucumber data summarized in Chapter 4, Table 4.1 are used here as an example. First, the variance (s^2) is calculated from the data in the table:

$$\text{variance } (s^2) = (\Sigma x^2 - {}^1/_n(\Sigma x)^2)/(n-1) = (601 - 3481/7)/6 = 17.286 \quad \text{(A2.1)}$$

The standard deviation (s) is calculated as $\sqrt{(s^2)} = \sqrt{17.286} = 4.158$. The standard error of the mean (SE) is then calculated as:

$$\text{SE} = s/\sqrt{n} = 4.158/2.646 = 1.571 \quad \text{(A2.2)}$$

The confidence limits are calculated as $\bar{x} \pm t_{0.05} \times$ SE where $t_{0.05}$ is the value read from the table given in Appendix 7. The value of t for 6 degrees of freedom is 2.447 at the 95% level and the confidence limits are 8.43 ± (2.447 × 1.571) or 8.43 ± 3.844. In terms of percentages the value of 3.844 is 46% of the mean value of 8.43. Thus there is 95% confidence (or 0.95 probability) that the true stock size lies somewhere between (1315 – 46%) and (1315 + 46%), or between 710 and 1920 sea cucumbers.