**CONFIDENCE INTERVALS**

**PEN & PAPER SESSIONS**

**A1. SAMPLING VARIABILITY OF MEANS**

1. a) n = 20, = 84.25, s = 3.16

b) Would expect roughly 95% in the range  ± 1.96 × s, i.e. from 78.1 to 90.4.

c) All 20 observations are within this range.

d) SE of is estimated by s/√n = 3.16/√20 = 0.71.

95% CI given by ± 2.093×SE, i.e. from 82.8 to 85.7. We use t rather than z to allow for the extra error introduced by using the sample standard deviation to estimate the population standard deviation. However, it makes very little difference for sample sizes greater than about 20.

There is a 95% chance that this interval will include the true mean height. *[To understand this interpretation imagine we are able to draw random samples of the same size from the population many many times, say 10,000. For each sample we calculate a 95% confidence interval. There will then be 10,000 such intervals. We would expect about 9,500 of these 95% confidence intervals to include the true mean height.]*

1. The interval calculated in b) is the range of values between which you expect most of the heights of individual boys to lie, whereas the interval calculated in d) is the confidence interval for the mean height.
2. 86.5 falls above the upper limit of our confidence interval so there is evidence that Jamaican boys with sickle cell disease do differ from the Tanner and Whitehouse standard.

g) The SE would be smaller 3.16/√ 100, therefore the confidence interval would be smaller

h) T would be 2.861 instead of 2.093 and therefore the confidence interval would be wider.

## Summary points

* The Standard Error of a mean measures the sampling variability of the mean. Use in calculating confidence interval for the true mean.
* The Standard Deviation measures variability in individual data values for a variable. Use to describe data spread/range.
* For a particular result, the 95% CI will be wider than the 90% CI but narrower than the 99% CI. We can never attain 100% certainty, 100% CIs do not exist.
* 95% CI: mean