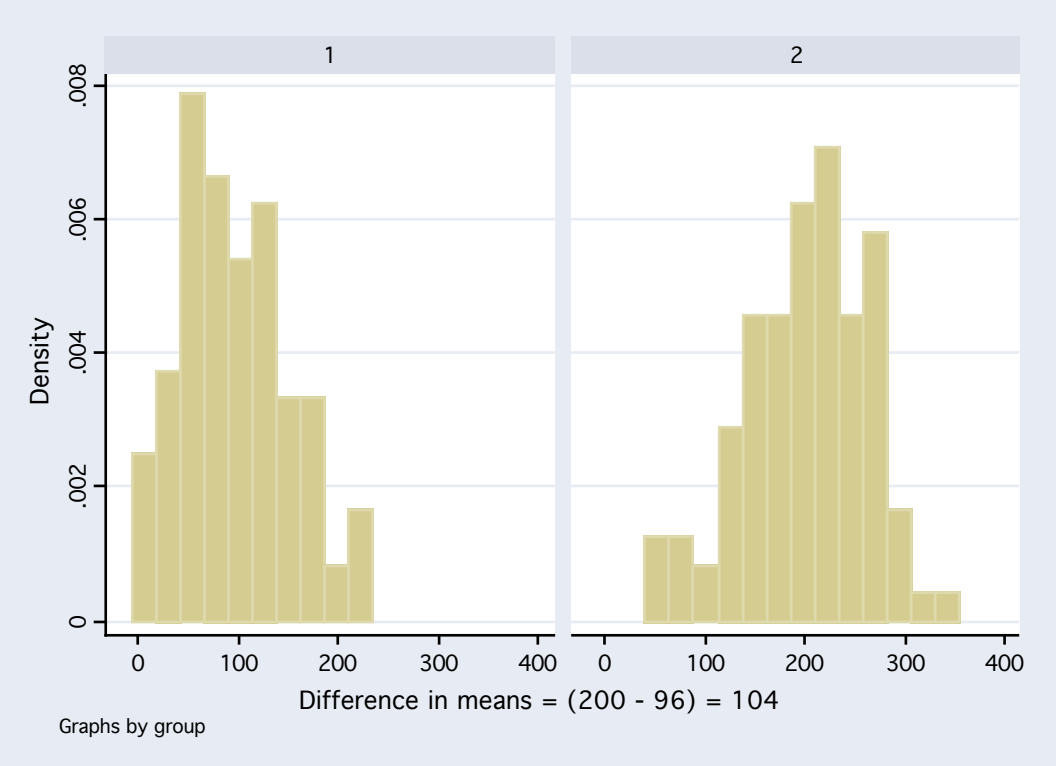
Statistics and clinical evidence
Tony Brady

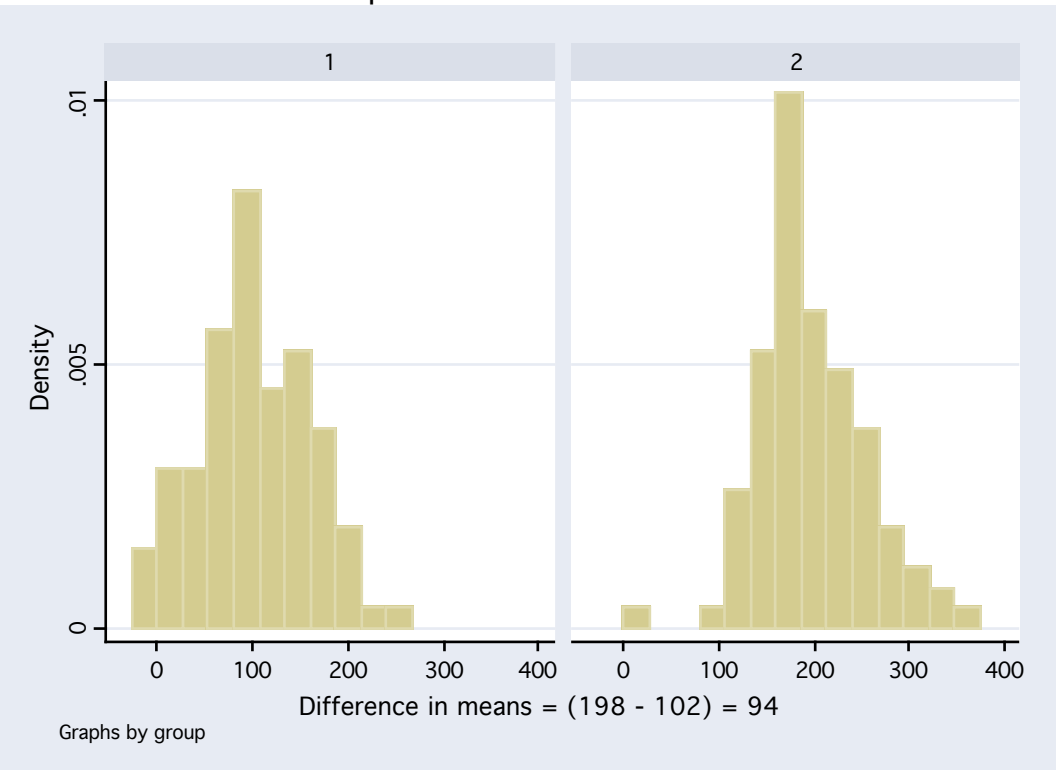
Simulations with population mean in group 1 = 100 and population mean in group 2 = 200

Result from one sample:



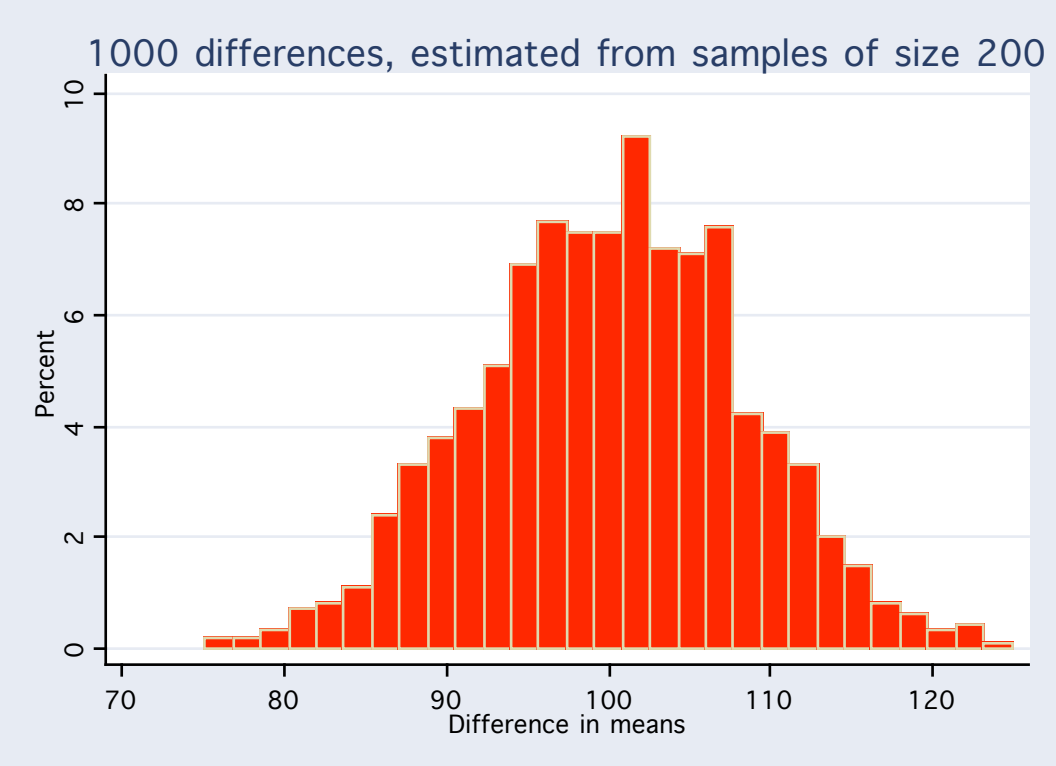
Graph shows distribution of raw data (e.g. blood pressure) in each group. The estimated difference in means from this sample is 104.

Results from another sample:



In the second sample the estimated difference in means is now 94.

If you were able to repeatedly take samples and estimate the difference in means (not normally possible in real life!) you would observe the **sampling distribution** of the mean difference. Here is the sampling distribution from 1000 samples like the two shown above:

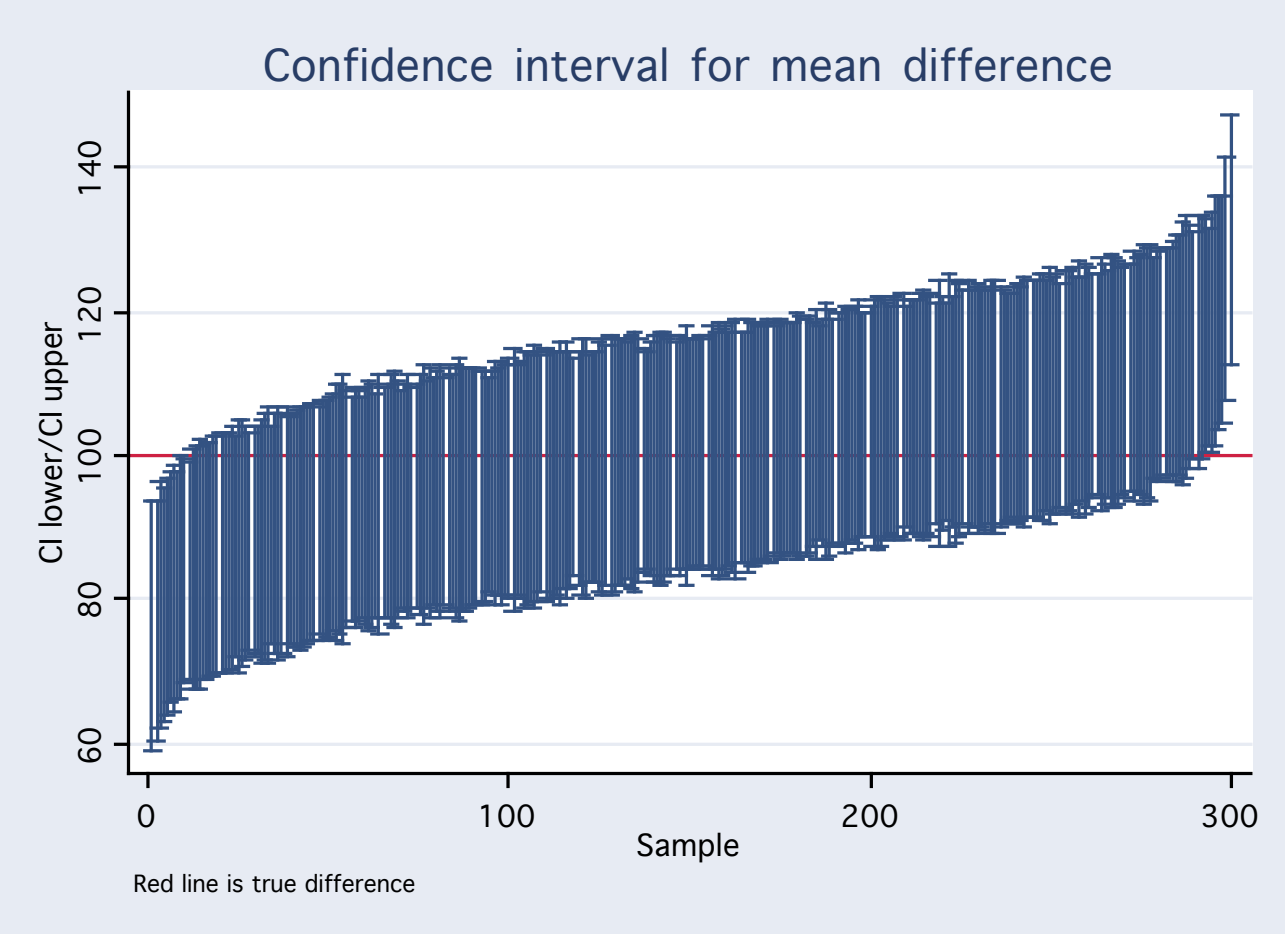


Notice that this sampling distribution is Normally distributed and has a mean very close to the population difference in means (200 - 100 = 100 in this simulation). The standard deviation of this distribution is called the **standard error** of the difference in means. We use the standard error in hypothesis tests (e.g. two sample t-test) and to construct confidence intervals when we only have one sample.

We can do this because, remarkably, we can estimate the standard error when we only have one sample (the usual situation in practice - we only have one trial to analyse!). The standard error turns out to be related to the sample size (100 in each group here) and the standard deviation of the raw data. The formula for standard error varies depending on which sampling distribution you are interested in - e.g. difference in means or just the mean in one group. But it is always based on the sample size and the standard deviation.

Confidence intervals

We can use the same 1000 samples to look at confidence intervals. For each sample the confidence interval for the difference in means has been calculated. The graph below just the first 300 confidence intervals so you can see more clearly:



The red line shows the population difference in means (100) which we only know because this is a simulation. You can see that some of the 95% confidence intervals at each end of the graph do not include 100. In fact around 5% of confidence intervals exclude the population difference. This is why, when we have a confidence interval from only one sample, we are only 95% certain that it will contain the population mean difference.

We can of course construct a 99% confidence interval (it will be wider than the 95% one) - and we can be more confident that it includes the population difference (since only 1% of such confidence intervals will exclude the population difference).