Research Methods 2

Week 6: Exercise Sheet 1

Question 1.

Haemoglobin concentrations in healthy adult females can be taken to have a Normal distribution and plausible values for the population mean and SD are, respectively, 13.5 and 1 g/dl. Use Minitab to generate a sample of 90 concentrations from this population and store it in C1. By clicking on the **Stat** menu and choosing **Basic Statistics** and then **Display Descriptive Statistics...** you should compute the SD of this sample. Use a calculator to compute the SE of the sample mean, as the SD divided by √90. Compare your result with the value given by Minitab under the heading SE mean.

Question 2.

As you learnt in the document for this week, an alternative way to calculate the SE is to compute the mean of each of a collection of samples, and then find the SD of these means.

Do this, by repeating the generation of a sample of size 90 ten times. For each sample compute the mean and type the value of this mean into a spare column, say C5. Thus at the end of the exercise C5 will contain ten sample means. Compute the SD of these values and compare it to the SE computed in Question 1.

Question 3.

It is likely that there will not be very close agreement between the results in questions 1 and 2 because it is unlikely that just ten sample means will adequately describe the distribution of the sample means. Try question 2 again but generate not 10 but 1000 samples, each of size 90.

[To do this by hand would take forever, so you should get the computer to do it for you. A special piece of code, known as a Minitab macro, has been written to allow you to do this. Therefore you should go back to the home page for this week, read the explanation about macros and then download and use the macro sampmns]

Draw the histogram of the 1000 sample means. Run the macro again, but this time generate 1000 samples of size 9. Draw the histogram of these sample means and compare with the previous histogram.

Question 4

The following are the number of polyps in the colons of 9 patients with familial adenomatous polyposis coli.

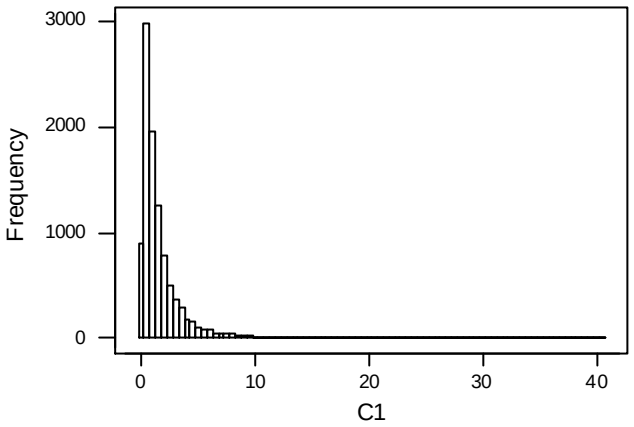
26, 16, 40, 14, 16, 11, 26, 45, 32

Type the data into Minitab and compute the mean of the sample. Also compute the standard error of the mean.

(Data due to Giardiello et al., NEJM, 1993, 328, 1313-16, selected from Piantodosi, S., (1997) Clinical Trials: a Methodologic Perspective, Wiley)

Question 5 { this question is a bit harder than the others}.*

The histogram shows a sample of size 10000 drawn from a population which is *not* Normal.



Download the macro 'lgnmns'. If you have saved the macro in Minitab's macro subdirectory, then typing, e.g.,

%lgnmns 90 200 c6

at the MTB> prompt in the Session window will generate 200 samples, each of size 90, from the population underlying the above histogram and compute the mean of each sample. The 200 sample means will be stored in column c6.

{The amendment for macros stored in other locations is the same as for the sampmns macro}

Use the macro to generate three columns. The first holding 1000 means, each mean being from a sample of size 10. The second is the same but for samples of size 50 and the third is based on samples of size 250. Draw the histograms of the three columns and comment.

End of Exercise Sheet 1