6.3A Sampling Distribution of the Sample Mean

Purpose: Investigate how the sampling distribution of the sample mean is affected by the sample size and the distribution from which the sample is taken.

Reading Assignment: Read through Section 6.3.

Problem Description: Do variations on Exercises 6.39 and 6.38.

Problem 1: This is a variation on textbook Exercise 6.39 (page 283). We will essentially generate 500 random samples of sizes n=1,2,4,8,16,32 from the uniform distribution on the interval from 0 to 10 and compute the sample mean for each of the 500 samples.

Step 1: Name columns c32-c37 as **n=1, n=2, n=4, n=8, n=16**, and **n=32** respectively.

Step 2: Generate 500 random samples of size 32 from said uniform distribution on (0,10) into c1-c32 as follows.

```
MTB > random 500 c1-c32;
SUBC> uniform 0 to 10.
```

Look at the data in the worksheet.

Step 3: Column c32, which we named 'n=1', contains 500 random samples of size n=1 from the uniform distribution on the interval from 0 to 10. Columns c1 and c2 together contain 500 independent random samples of size 2 from this uniform distribution. Compute the corresponding sample means and put them into c33, which has been named 'n=2'. This can be done using the "row-mean" command, **rmean**, as follows.

```
MTB > rmean c1-c2, put into c33
```

Step 4: Similarly, compute sample means using the first 4, 8, 16 and 32 columns, and put the resulting row means into columns c34-c37, respectively.

```
MTB > rmean c1-c4, c34
MTB > rmean c1-c8, c35
MTB > rmean c1-c16, c36
MTB > rmean c1-c32, c37
```

Step 5: Generate dotplots <u>on the same scale</u> for the data in c32-c37.

```
MTB > dotplot c32-c37;
SUBC> same.
```

Step 6: Generate dotplots <u>not</u> on the same scale for the data in c32-c37.

```
MTB > dotplot c32-c37
```

STOP AND THINK: How do the plots illustrate the Central Limit Theorem? Do the sample means appear to be unbiased estimates of the population mean (which is 5)? And does the spread of the sampling distribution tend to be smaller for larger sample size?

Step 6: Generate descriptive statistics for the data in columns c32-c37.

```
MTB > desc c32-c37
```

STOP AND THINK: Formulae for the mean and standard deviation of the uniform distribution are given on textbook page 220. Our samples have been drawn from a uniform distribution with c=0 and d=10, so the mean of the distribution is 5 and the standard deviation is 10/sqrt(12) = 2.88675.

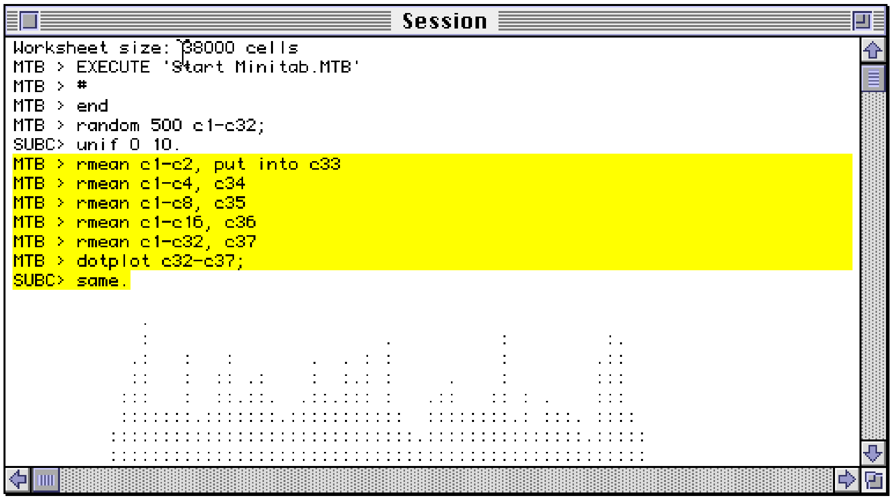
We have learned, for a random sample from a population (or distribution), that the mean of the sample mean is the same as the mean of the population (or distribution) from which the sample is drawn. (See the Properties of the Sampling Distribution of "x-bar", textbook page 274.). Do the simulations support this? (Look at the means (of the means) in the descriptive statistics, and remember that simulations are approximate -- not exact!)

We have also learned, for a random sample from a population (or distribution), that the standard deviation of the sample mean is smaller than the standard deviation of the distribution from which the sample is drawn, by a factor of the square root of the sample size. Do the simulations support this? (Look at the standard deviations of the sample means in the descriptive statistics. Are their values close to the population standard deviation divided by the square root of the sample size?)

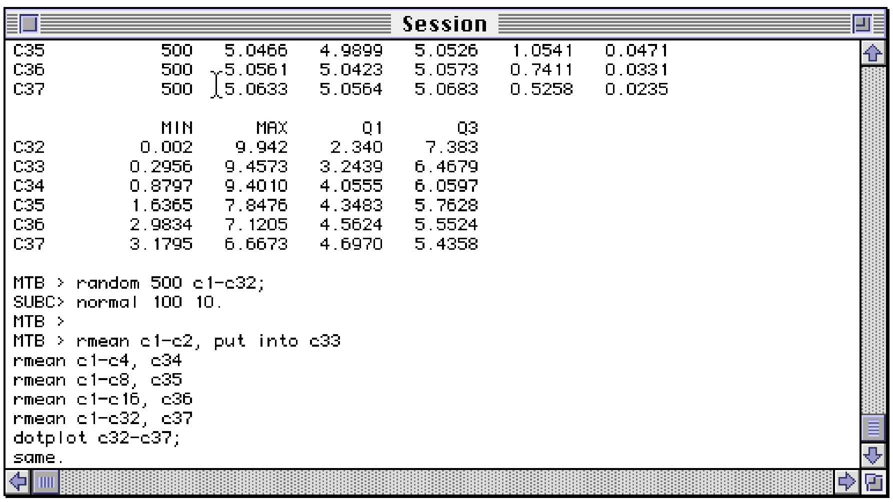
Problem 2: This is a variation on textbook Exercise 6.38 (page 283). Repeat Problem 1 of this lab, except sample from the normal distribution with mean 100 and standard deviation 10, instead of sampling from the uniform distribution. Thus, replace the Minitab commands shown under Step 2 of Problem 1 with the following commands. (You need not rename the columns.)

```
MTB > random 500 c1-c32;
SUBC> normal 100 10.
```

You can avoid retyping the commands from Steps 3-5 by copying and pasting them as follows. Scroll up in the Minitab Session window to where you entered the same commands for Problem 1, select them as shown below, then copy them (i.e. Edit -> Copy).



Next, go to the end of your session (hit the "return" key), then paste the commands into your Session window at the last Minitab prompt as shown below.



(You could then edit what you pasted in, but this time it isn't necessary.) Finally, press the "return" key to get Minitab to execute these commands. This is a great time saver, so you are encouraged to try it!

Don't forget to also generate the second set of dotplots, as well as the descriptive statistics!

When you finish both problems, copy your Minitab output into a Word document, then paginate, save and print it.

Lab report: Write a report on the sampling distribution of the sample mean. Draw upon what you did and observed in both problems in this lab. Include discussion of the following items and questions. In the first problem, the samples were from a uniform distribution -- not a normal distribution. Was the Central Limit Theorem illustrated in the first problem? In the second problem, the samples were from a normal distribution. What should the sampling distributions of the sample means look like? Did they look as they should? Given a random sample of size n from a population (or distribution) with mean mu and standard deviation sigma, we know that the mean and standard deviation of the sample mean are mu and sigma over the square root of n, respectively. This should be approximately true in the simulation, but it isn't likely to work out exactly, since the simulations only approximate the theoretical sampling distributions. Did the means and standard deviations of the sample means roughly follow these rules?

(As always, append annotated Minitab output to your report, and cross-reference it in your output.)

Lab 6.3a, 6/2002