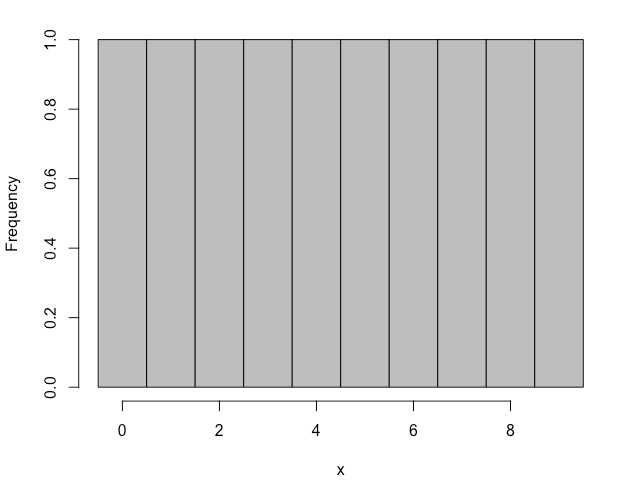
**CHAPTER 9 SOLUTIONS**

1. First, we create the dataset by typing in the following R command:

**x = c(0, 1, 2, 3, 4, 5, 6, 7, 8, 9)**

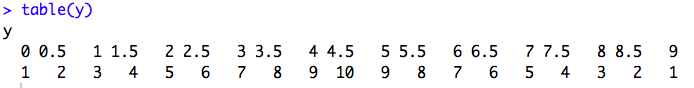
Then, we create a histogram by typing in the R command **hist(x, breaks = -0.5:9.5, col = "grey", main = "")**

****

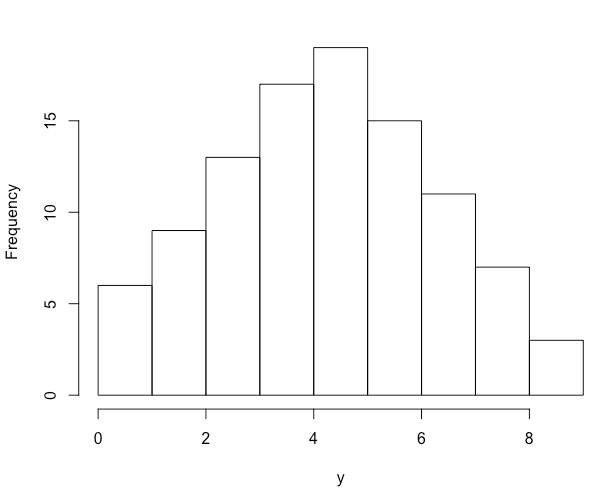
The shape is uniform.

b) The relevant samples of size 2 are: 00,01,02,03,04,05,06,07,08,09,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99

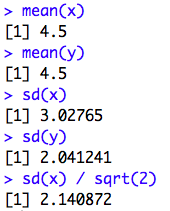
c) The frequency distribution is presented below.



d) Although the population was uniformly distributed, the sampling distribution of means is approximately normally distributed.



e) Although the mean of the sampling distribution is equal to the mean of the population, the standard error is not exactly equal to the standard deviation of the population divided by the square root of 2. This result does not contradict the Central Limit Theorem because the conditions have not been met: given that the population is not normally distributed, the sample size should be at least 30.



f) The probability of selecting a mean between 2.5 and 6.5 inclusive is 70/100 = .70.

* 1. To start, we generate the approximation of the sampling distribution of the mean by typing the R commands:

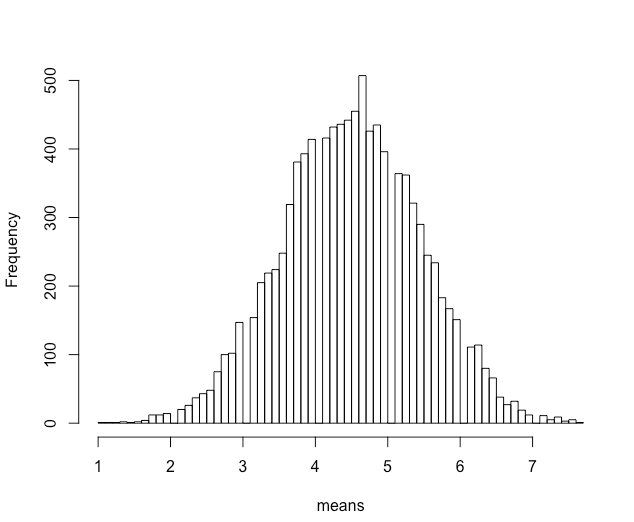
**x = c(0, 1, 2, 3, 4, 5, 6, 7, 8, 9)**

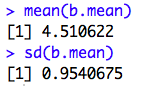
**set.seed(1234)**

**b.mean = boot.mean(x, B = 10000, n = 9)$bootstrap.samples**

We then create the histogram using the R command **hist(b.mean, breaks = 50, xlab = "means", main = "")** and obtain the mean and standard deviation of the bootstrapped sample means using the R commands **mean(b.mean)** and **sd(b.mean)**.

The sampling distribution generated has mean 4.51 and standard deviation .95. The mean of the actual sampling distribution, as predicted from the CLT, is the mean of the population {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}, which is 4.5, agreeing with the result we obtained by simulation. The standard deviation of the actual sampling distribution, as predicted from the CLT, is the standard deviation of the population divided by the square root of 9, which yields 1.01, agreeing, within rounding, with the result we obtained by simulation.





a) Empirically, one would find all equally likely samples of size 8 and compute the mean of each, and the list of those means is the sampling distribution. However, for most populations, this would be too difficult and time-consuming to do without a computer. R can be used to approximate the sampling distribution of means by bootstrapping a large number of samples.

b) For the sampling distribution, the mean is 70 and the standard deviation, or standard error, is .88 to two decimal places.

c) Increasing the sample size has no effect on the mean of the sampling distribution, but it would decrease the standard error.

d) .0118. Because the sampling distribution of the means is normally distributed with mean 70 and standard error .88, the probability of selecting a sample mean that is larger than 72 is the same as the probability of selecting a *z*-score that is higher than 2.26.

a) The distribution will be normally distributed.

b) 500

c) 18.26

d) Approximately .046.

e) It will be normally distributed with mean 500 and standard error approximately 33.33.

a) According to the Central Limit Theorem, it will be approximately normally distributed.

b) 500

c)  14.14

d)  Approximately .046

e) Given that the uniform distribution is symmetric, it is likely that the CLT will apply and that the sampling distribution of means will be approximately normally distributed with mean equal to 500 and standard deviation equal to 33.33.

1. The sampling distribution of the means is approximately normally distributed with mean 100 and standard error 10/√64 = 1.25.
2. .0548 (*z* = (98-100)/1.25 = -1.6. Area to the left of -1.6 = .0548.)
3. False. The standard deviation of the sampling distribution of the mean is also called the standard error.
4. False. The mean of the sample is probably close to the population mean, but is usually not exactly equal to it.
5. True. The standard error is given by the population standard deviation divided by the square root of the sample size. So when you have a larger sample, you are dividing by a larger number, resulting in a smaller standard error.
6. False. The CLT tells us that mean of the sampling distribution of means is the same as the population mean.
7. True. A sample is a representation of its population. This is especially true for large samples.
8. False. The sampling distribution of the means is normal, like the population, regardless of sample size.
9. True. For a normal population the sampling distribution of means is normally distributed.
10. True. The CLT tells us that mean of the sampling distribution of means is the same as the population mean no matter what the sample size.
    1. The sample is not random because not every person in New York City is equally likely to be a member of a statistics class at New York University.
    2. The statement is false because not all sampling distributions of means are normally distributed. For example, if the population is not normally distributed and the sample size is small, the sampling distribution of means may not be normally distributed.
    3. In inferential statistics, one selects a single sample and bases the inferences on that. The sampling distribution of means is not empirically created when conducting an inferential analysis.
    4. Using this procedure, the researchers will not obtain a simple random sample because the twin of any selected subject is also selected. Thus, it is not true that at each stage of the selection process all subjects remaining in the population have an equal chance of being selected.
    5. Positively skewed.
    6. Approximately normal (from the CLT).
    7. Score or X.
    8. Sample mean or.