Chapter 7 Sampling and Sampling Distributions

Solutions

* 1. To get a simple random sample of all iPad purchasers, you could randomly select customers from all of the stores that sell iPads and ask their ages and whether or not they have purchased an iPad.
  2. To get a stratified random sample you could first create strata based on ethnicity; for example under white, black, Hispanic, Asian, and then randomly select people in each group and ask his/her age when he/she purchased an iPad.
  3. To get a cluster sample, you could make clusters of stores that sell iPads, and randomly select certain stores and gather data about the ages of people who purchase iPads there.

1. There could be some nonresponse bias due to certain people choosing not to stop at the booth. There could also be some selection bias since the booth is only open during the weekend. Some people may avoid the mall on the weekend and therefore the booth will not get information about people who shop and eat at the food courts during the week. Also, since they are trying to determine the monthly expenditure, and they only gather information from one weekend, this could cause some bias as well.
2. Natalie’s analysis is not based on a representative sample because she is only surveying students in her accounting class. Since students other than accounting students are likely to apply to the MBA program, Natalie should include other majors/concentrations as well. In addition, since she will likely be competing with students from other schools as well, she should also include information from students at other schools who are likely to apply to the Berkeley MBA program.
3. 1. This could lead to nonresponse bias since people who do not mail the envelope back may have different preferences than people who do respond.
   2. This could lead to selection bias, since customers who frequent the store in the morning are more likely to prefer an earlier opening time than people who frequent the store at other times of the day.
   3. This could also lead to selection bias and nonresponse bias since not everyone in Grover Beach reads the newspaper and not everyone can get online to the store’s website.
4. 1. For a simple random sample you could have the clerks randomly select customers who visit the Vons in Grover Beach throughout the day. The clerks could ask them and record how likely they would be to visit the store between 6am and 7am.
   2. For a stratified random sample, you could create groups of customers based on age or time of day visiting the store and randomly sample customers in each group to ask how likely they would be to shop at Vons between 6am and 7am.
   3. For cluster sampling, you could cluster people by neighborhood that they live in and randomly select neighborhoods to do door-to-door surveys.
5. 1. Both sample means have a normal distribution because the population is normally distributed.
   2. Yes, because for both sample sizes is normally distributed.



* 1. 1. For



* + 1. For



1. 1. Since we do not know whether or not the population is normally distributed and *n* = 16< 30, we cannot assume that the distribution of with is approximately normal. However, for , by the Central Limit Theorem, is approximately normal.



* 1. We can only use the normal approximation for the sample mean with .







* 1. .



1. 1. .



* 1. .



* 1. .



1. 1. .



* 1. .



* 1. .



1. 1. .



* 1. The probability that the mean weight of a 12-pack of beer is less than 325 ml is much less than that of a single bottle because the variation in is less when the sample size is bigger.



1. 1. .



* 1. .



* 1. Janice. A higher probability found in part a suggests that Janice's findings are more likely if a representative sample is used.

1. The probability that the total weight of the 16 selected persons is greater than 3,200 pounds is the same as the probability that their average weight is greater than 3,200/16 = 200 pounds per person.



Therefore, there is a 0.0021 probability that a random sample of 16 individuals will exceed the weight limit of 3,200 pounds.

* 1. The sample mean with has a normal distribution because the sample is taken from a normally distributed population; in addition, .







* 1. .



1. 1. . Therefore, there is a 3.14% chance of getting a sample average of 35 or more without a discount.



* 1. We feel reasonably confident that the manager's discount strategy has worked since there is only a small chance of 3.14% of getting 35 or more customers if the manager had not offered the discount.

1. 1. It is appropriate to use the normal distribution approximation because the approximation conditions are satisfied: .



* 1. It is appropriate to use the normal distribution approximation because the approximation conditions are satisfied: .



1. 1. The sampling distribution of the sample proportion is approximately normal when but not when When the approximation condition is not satisfied because . When the approximation conditions are satisfied: .



* 1. As shown in part (a), you can only use the normal approximation for the sample size of 50. For the sample with *n* = 20, you cannot assume that is approximately normally distributed.



1. 1. .



1. 1. The sampling distribution of has because the approximation conditions are met: .



* 1. .



1. 1. The sampling distribution of has The normal approximation criteria are met because . Therefore, it is appropriate to use the normal distribution approximation for the sample proportion.



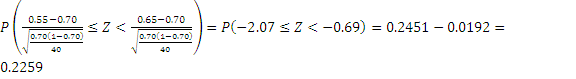
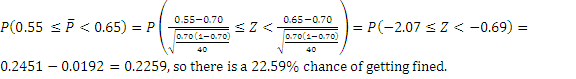
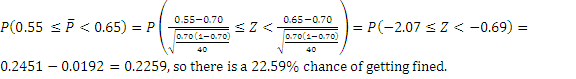




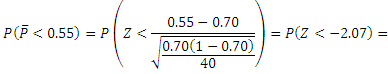
Given , with 26 successes and with 22 successes.



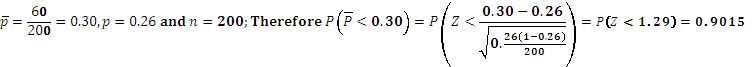
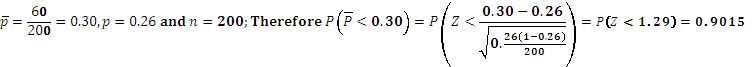
1. The probability that the dealer will be fired, but still continue to operate, is given by



1. The probability that the dealer will be dissolved is given by







* 1. is the proportion of French people who approve, therefore is the population proportion of disapproval.



1. You would choose 50 balls because with larger sample sizes the standard deviation of is *reduced*. The proportion of green balls is 60/100 = 0.6. Therefore your probability of getting 70% green balls is slightly higher with a smaller sample because of the increased standard deviation. (If you are unsure about this you can



calculate to confirm that the probability is higher for



.



* 1. .



* 1. .



* 1. The difference between the two probabilities is from the sample size. The standard deviation is reduced with larger sample sizes, which brings the sample proportion closer to the population proportion. Therefore, when there is a greater probability that the sample proportion



will be closer to the population proportion of 23%, so there is a greater probability that the sample proportion will be more than 20%.

* 1. It is not necessary to apply the finite population correction because the sample constitutes less than 5 percent of the population: *n* = 100 < 125 = 2500(0.05).



* 1. . Yes, it is necessary to apply the finite population correction because the sample constitutes for more than 5 percent of the population: *n* = 70 > 25 = 500(0.05).



* 1. (approximately)



1. 1. ; there is no need to apply the finite population correction factor because the sample size does not account for 5 percent of the population size: *n* = 100 < 150 = 0.05(3,000).



1. 1. . Yes, it is necessary because the sample size accounts for more than 5 percent of the population size: *n* = 80 >0.05(600) = 30.



1. 1. Yes, it is necessary because the sample size is greater than 5 percent of the population size: *n* =32 > 0.05(500) = 25.
   2. The sampling distribution of the sample mean is approximately normal because the sample size is greater than 30.
   3. .



* 1. .



1. 1. No, it is not necessary because the sample size does not account for 5 percent of the population size: *n* =12< 0.05(500) = 25.
   2. We cannot assume the sampling distribution of the sample mean is normally distributed because we do not know if the population has a normal distribution and the sample size is not sufficiently large to assume so (.



* 1. The normal approximation is not justified (see part b).

1. 1. . Therefore, since *np ≥* 5 and  *≥* 5, the sampling distribution of the sample proportion is approximately normal. the sample accounts for more than 5 percent of the population size (*n* =20 > 0.05(250) = 12.5), we need to



apply the finite population correction;



* 1. .



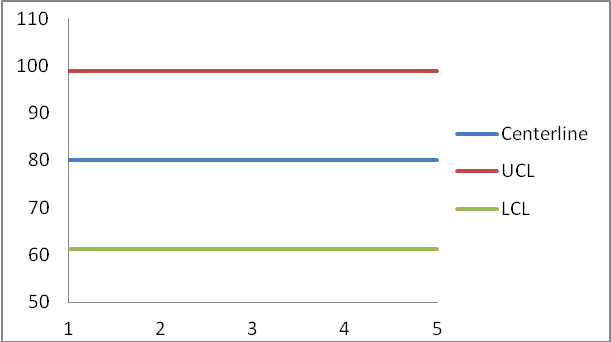
1. First define:



the sample accounts for more than 5 percent of the population size (*n* =120 > .05(1000) = 50), we need to apply the finite population correction; we then find E(Then, , so we compute .







Centerline: *µ* = 80

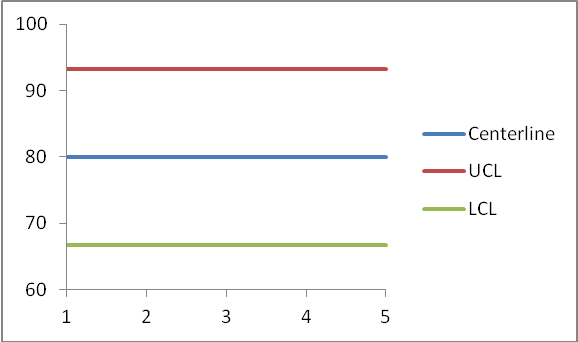
UCL =



LCL =







Centerline: *µ* = 80

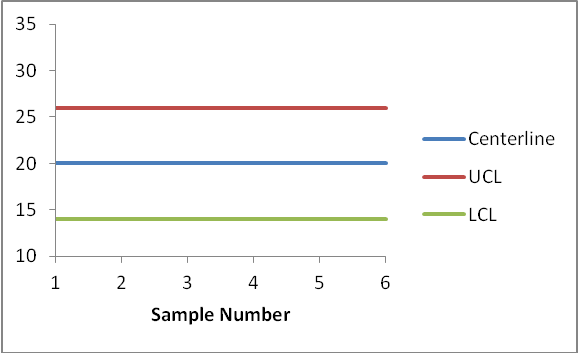
UCL =



LCL =



* 1. The larger sample size gives narrower controls limit due to the smaller standard deviation.



Centerline: *µ* = 20

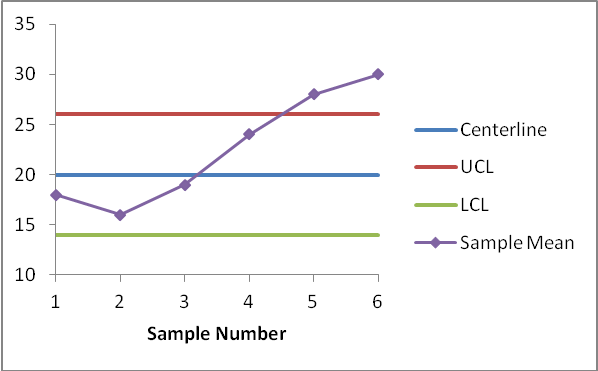
UCL =



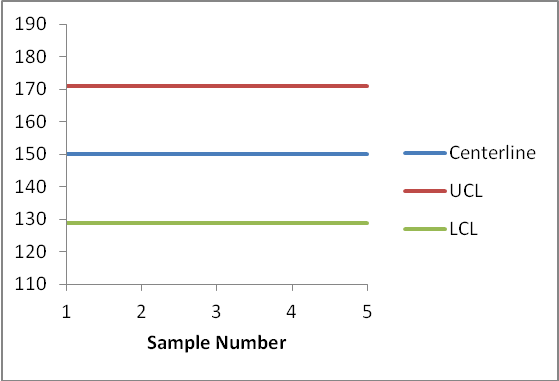
LCL =







* 1. The last two points are outside the upper control limit. There is also an upward trend, suggesting the process is becoming increasingly out of control. The process should be adjusted.



Centerline: *µ* = 150

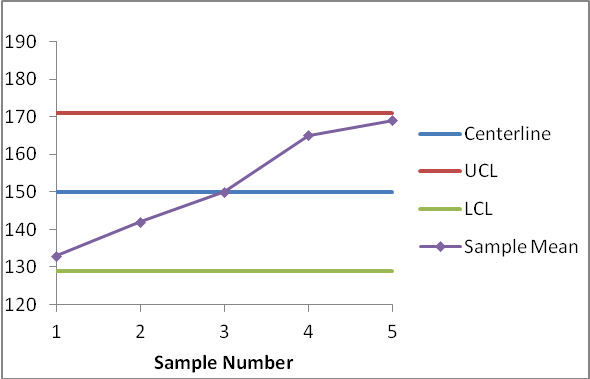
UCL =



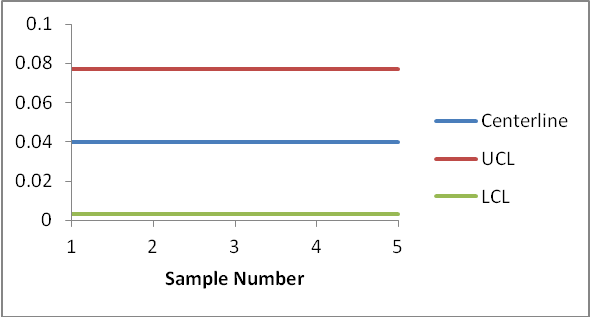
LCL =







* 1. There are no points outside of the control limits. However, there is a positive trend, suggesting that the process may soon have a mean outside of the upper control limit if it is not adjusted.



Centerline:



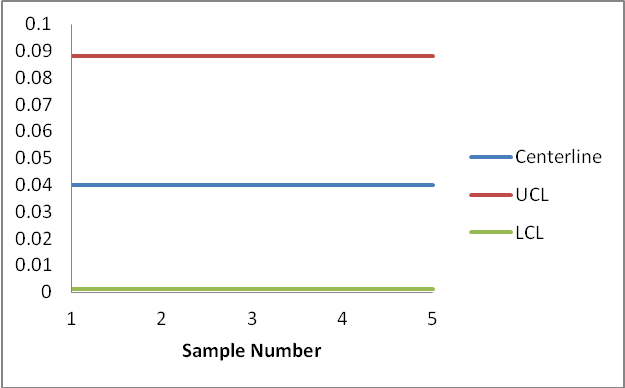
UCL =



LCL =







Centerline:



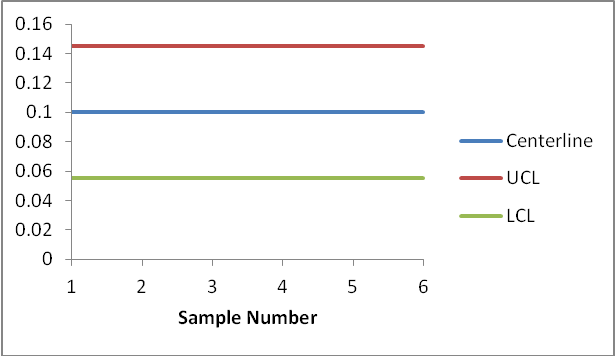
UCL =



LCL =



* 1. The control limits have a larger spread with smaller sample sizes due to the increased standard deviation for the smaller sample size.



Centerline:



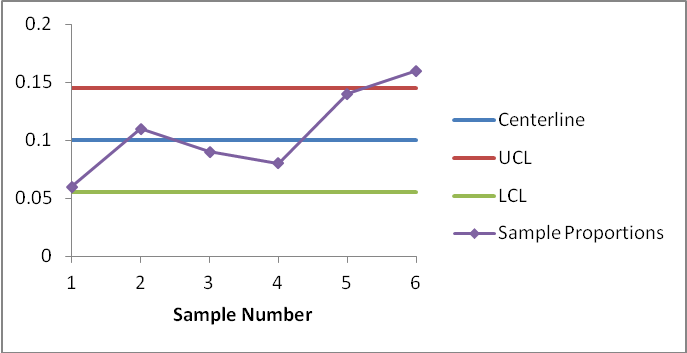
UCL =



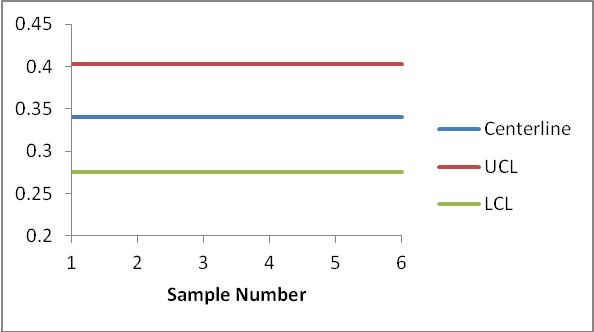
LCL =







* 1. No, the production process seems to be out of control due to the 6th sample proportion, which is above the upper control limit.



Centerline:



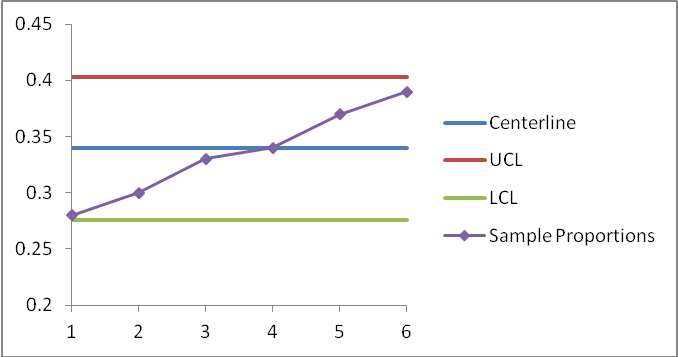
UCL =



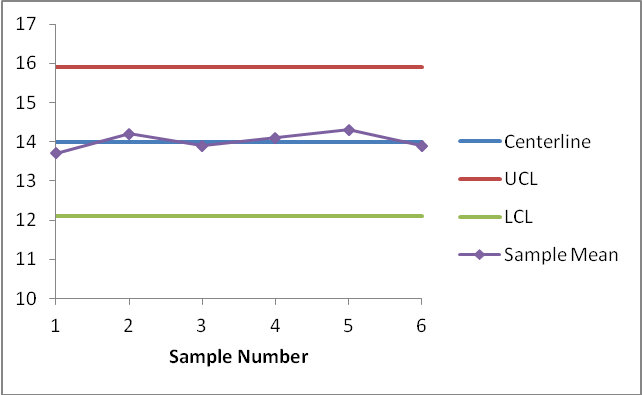
LCL =







* 1. There are no points outside the control limits, so the process is under control. However, the positive trend suggests that the process may become out of control if the upward trend continues.



Centerline: *µ* = 14

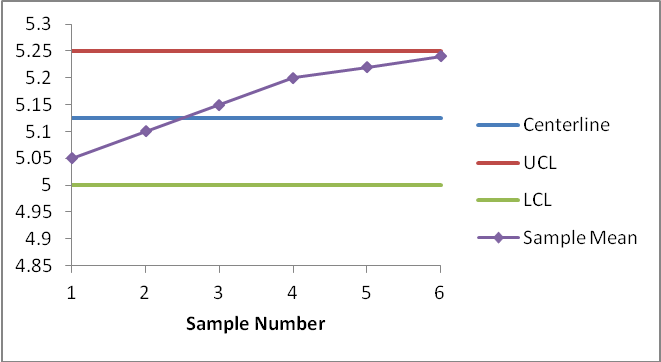
UCL =



LCL =



* 1. All the sample means lie within the control limits. Therefore, we can conclude that the production process is in control and operating properly.



Centerline: *µ* = 5.125

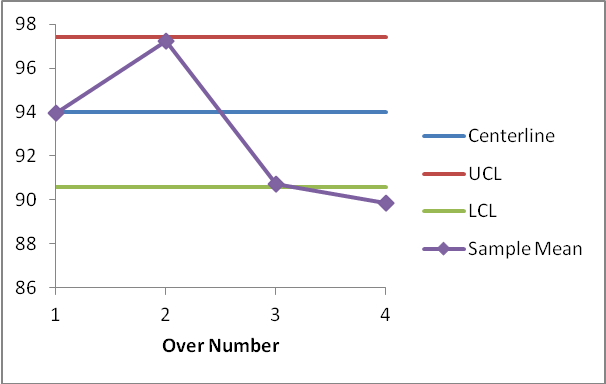
UCL =



LCL =



* 1. There are no points outside the control limits. It appears that the process is under control, but the positive trend suggests the process may become out of control if the trend continues.



Centerline: *µ* = 94

UCL =



LCL =



To plot the average speed, take the average of each over:

Over 1  =



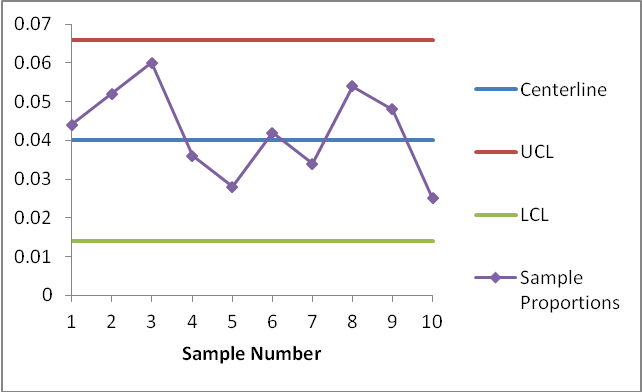
Similarly,

Over 2 = 97.23

Over 3 = 90.70

Over 4  = 89.85

* 1. Kalwant’s average speed is out of the control limits on 1 out of 4 of his overs, which rather justifies his coach’s concern that he is not very consistent.



Centerline:



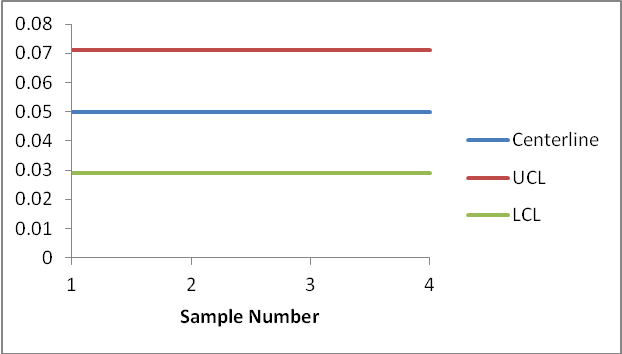
UCL =



LCL =



* 1. All sample proportions are within the control limits and there is no apparent trend, suggesting that the machine is operating properly.



Centerline:



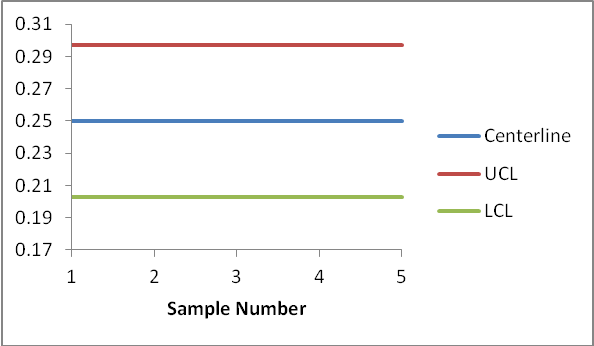
UCL =



LCL =



* 1. Since 0.062 is within the control limits, the process is in control.



Centerline:



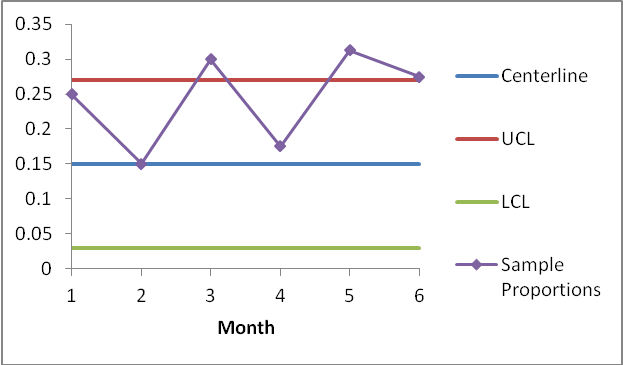
UCL =



LCL =



* 1. 240/750 = 0.32, which is above the upper control limit of 0.297, so the university should be concerned about this.



Centerline:



UCL =



LCL =



* 1. Find the proportion of complaints each month:

|  |  |
| --- | --- |
| Month | Sample Proportion |
| 1 | 20/80 = 0.25 |
| 2 | 12/80 = 0.15 |

|  |  |
| --- | --- |
| 3 | 24/80 = 0.30 |
| 4 | 14/80 = 0.175 |
| 5 | 25/80 = 0.3125 |
| 6 | 22/80 = 0.275 |

We plot each sample proportion on the control chart (shown above) to see that 3 out of 6 months were out of the control limits, which is a good justification for why Dell chose to direct customers away from India call centers.

1. 1. There could be selection bias because people who go to the beach often go there to walk or exercise and are more likely to follow a consistent walking regimen.
   2. There could be nonresponse bias due to differences in people who mail back the envelope.
   3. There could be selection bias because some people may not have access to the website, or may not be computer savvy. Also nonresponse bias because

people who choose not to respond may have different preferences than the people who do choose to respond.

* 1. There could be selection bias because patients in hospitals are probably ill and may not be capable of walking vigorously three times a week.

1. 1. As one example, use a random number table or a random number generator (in Excel, for instance) to randomly select individuals into the sample from the list of all residents of Miami. Then conduct the survey by contacting those selected.
   2. To get a stratified random sample, you could create strata based on ethnicity; for example under white, black, Hispanic, Asian, and then

randomly select adults in each group and ask whether or not they walk regularly.

* 1. To get a cluster sample, you could choose a number of representative neighborhoods in Miami and randomly select adults within these neighborhoods and ask whether or not they walk regularly.











* 1. Even though finance graduates have a lower mean starting salary than accounting graduates, there is greater standard deviation for finance graduates. Therefore, 100 finance graduates have a slightly higher probability (1.07% compared to 0.39%) of earning an average salary of more than $52,000.

1. 1. .



* 1. .



1. 1. The sampling distribution of the sample mean will be approximately normal for any sample size because the population is normally distributed; in addition,



* 1. .



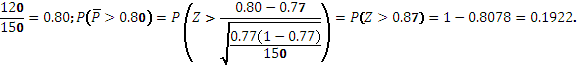
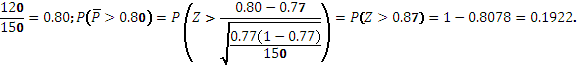
















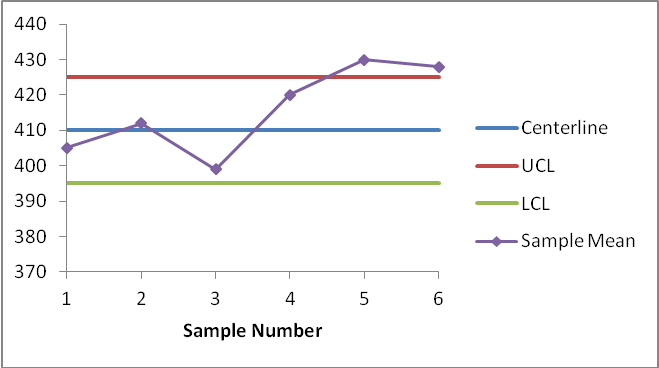
* 1. There is a smaller probability that more than 40 percent have pushed back their retirement date with the larger sample size of 200 adults (0.0694 compared to 0.1469 with a sample size of 100). This is because with larger sample sizes, the standard deviation of is reduced; is more likely to be



closer to the population proportion of 0.35, and therefore has less probability of being greater than 0.40.







Centerline: *µ* = 410

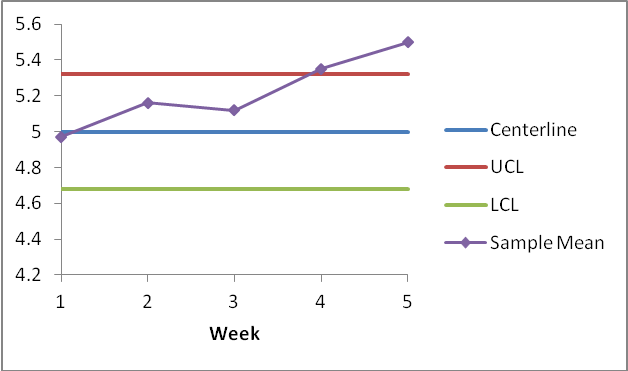
UCL =



LCL =



* 1. Two of the sample means are above the upper control limit, indicating that the advertised amount of sodium content is not accurate.



Centerline: *µ* = 5

UCL =



LCL =



The five weekly sample means are:

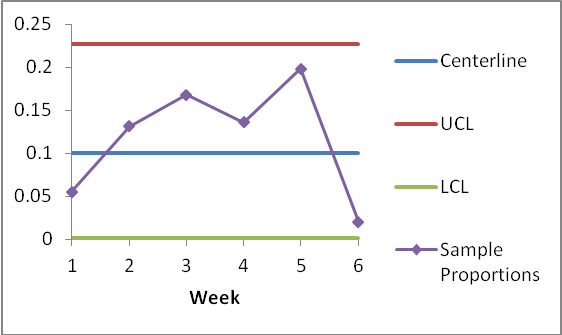
|  |  |
| --- | --- |
| Week 1 | 4.97 |
| Week 2 | 5.16 |
| Week 3 | 5.12 |

|  |  |
| --- | --- |
| Week 4 | 5.35 |
| Week 5 | 5.50 |

* 1. The last two points are outside the upper control limit, and there is a positive trend, suggesting that the process is out of control and will continue getting out of control and needs to be adjusted.







Centerline:



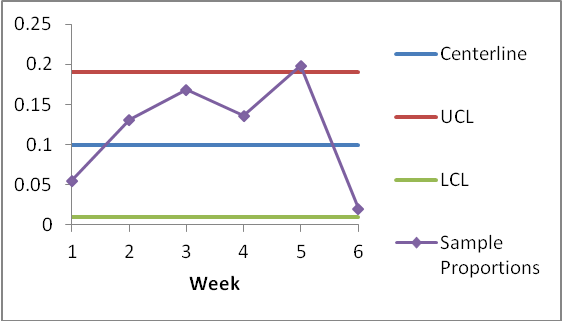
UCL =



LCL =



All of the sample proportions appear randomly within the control limits, and therefore no adjustments are needed.



Centerline:



UCL =



LCL =



Week 5 is slightly above the upper control limit when the sample size is larger. In week 6 the sample proportion is almost below the lower control limit. The firm may want to inspect the machine.

**Case Study 7.1**



1. The 1st sample with will be more representative of the U.S. as a whole due to the smaller standard deviation.



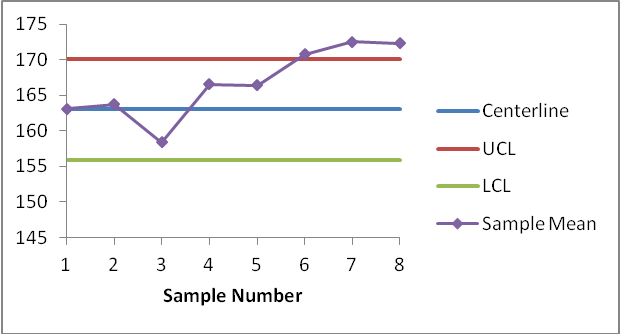
**Case Study 7.2**

1. Since there is an expected 8 percent *unemployed* college graduates, the expected number of college graduates who are employed is (1 - 0.08)(220) ≈202.

Similarly, the expected number of high school graduates who are employed is (1 - 0.245)(140) ≈ 106.



**Case Study 7.3**



Centerline: µ = 163 grams

UCL =



LCL =



1. Sample Means are:

|  |  |
| --- | --- |
| #1 | 163.1 |
| 2 | 163.73 |
| 3 | 158.4 |
| 4 | 166.52 |

|  |  |
| --- | --- |
| 5 | 166.37 |
| 6 | 170.75 |
| 7 | 172.54 |
| 8 | 172.28 |

The process does not appear to be in control; the last 3 sample means are above the upper control limit, and there is a positive trend, suggesting that the process is going to continue to get further out of control. The process should be adjusted.