**Problem Set 5**

**Problem 1.** The time taken to assemble a car in a certain plant is a random variable, not necessarily normally distributed. A random sample of assembly times of size 49 is taken. The sample mean is 21 hours and the sample standard deviation is 2 hours.

a) Construct a 99% confidence interval about our estimate for the mean.

**Sample standard error = sample std. dev/sqrt(n) = 2/7**

**For a 99% CI, we use NORMSINV(.995) = 2.576 (approx.)**

**Lower confidence limit = 21 – 2.576 x 2/7 = 20.26 (approx)**

**Upper confidence limit = 21 + 2.576 x 2/7 = 21.74 (approx)**

b) Can we conclude, at the 99% level of confidence, that the true mean time to assemble a car is greater than 20 hours?

**Since 20 is less than our lower confidence limit, yes.**

**Problem 2.** Repeat problem 1 but for a 95% confidence.

**Sample standard error = sample std. dev/sqrt(n) = 2/7**

**For a 95% CI, we use NORMSINV(.975) = 1.96 (approx.)**

**Lower confidence limit = 21 – 1.96 x 2/7 = 20.44 (approx)**

**Upper confidence limit = 21 + 1.96 x 2/7 = 21.56 (approx)**

**Since 20 is less than our lower confidence limit, yes.**

**Problem 3.** Repeat problem 1, but this time assume that your sample size is 484 and that it yields the same sample mean and sample standard deviation as stated in problem 1.

**Sample standard error = sample std. dev/sqrt(n) = 2/sqrt(484)**

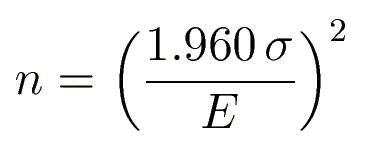
**For a 99% CI, we use NORMSINV(.995) = 2.576 (approx.)**

**Lower confidence limit = 21 – 2.576 x 2/22 = 20.77 (approx)**

**Upper confidence limit = 21 + 2.476 x 2/22 = 21.23 (approx)**

**Since 20 is less than our lower confidence limit, yes.**

**Problem 4.** Based on the information given in the statement of problem 1, how large a sample size would you need to take to guarantee that a 95% confidence interval was within plus or minus 0.5 hours?



**Here, σ = 2 and E = 0.5. Plugging in we get n = 61.46. Rounding up, we get n = 65.**

**Problem 5**. Recall Problem 7 from Homework 4. That is:

A piston in a heart valve must fit into a sleeve. Due to inevitable variations in manufacturing processes, the diameters of pistons and sleeves vary somewhat. The diameter of a sleeve, denoted D, is normally distributed with a mean of 0.0650 inches and a standard deviation of 0.0002 inches. The diameter of a piston, denoted d, is normally distributed with a mean of 0.0600 inches and a standard deviation of 0.0002 inches.

A critical factor that determines how long the heart valve will last is the clearance between the piston and the sleeve, defined as

Clearance = C = D - d.

The nominal design clearance is 0.005 inches, but the heart valve will function acceptably so long as the actual clearance is held to within plus or minus 0.0002 inches of the nominal design standard. In homework #4 we determined that the fraction of the piston/sleeve assemblies meeting the required tolerance is only 52%.

To improve the yield, management has decided to adopt a strategy of sorting pistons and sleeves in batches of size five. That is, they will wait until they have produced 5 pistons and five sleeves. They will then measure and sort the pistons from largest to smallest and the same thing with the sleeves. The largest piston will be matched with the largest sleeve, the next largest piston with the next largest sleeve, and so on. Whatever else you might think of this manufacturing tactic (it is actually used in industry), one thing of interest is whether or not it improves the process yield. To answer this question, management took a 20 samples of 5 pistons and 5 sleeves each (for a total of 100 piston-sleeve pairs). Each group of five was sorted and matched up as described immediately above. Afterwards, the resulting clearances were measured, and it was found that 79 out of the 100 assemblies met the required tolerance.

1. What would be management’s estimate of the process yield from implementing selective assembly as described?

**Estimate of process yield = 79%.**

1. Construct a 99.9% confidence interval about the estimate.

**Sample proportion, P = 0.79.**

**Sample standard error = sqrt[P x (1 – P)/n] = sqrt[0.79 x 0.21 / 100] = 0.04 (approx.)**

**For a 99.9% CI, we use NORMSINV(.9995) = 3.29 (approx.)**

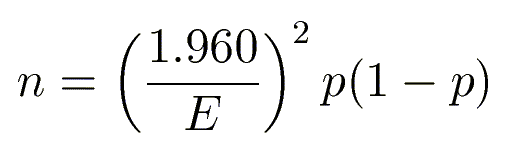
**Lower confidence limit = .79 – 3.29 x 0.04 = 0.656 (approx)**

**Upper confidence limit = .79 + 3.29 x 0.04 = 0.924 (approx)**

1. How sure can management be that implementing selective assembly as described will improve the process yield?

**Very sure. The lower confidence limit is quite a bit higher than the original 0.52.**

1. Suppose, before collecting any data at all about the proposed new process, you wanted to determine a conservative estimate for how large a sample size you would need. How large should your sample size be if you want to have a 95% confidence level that your estimate will be within plus or minus 1 percentage point of the true mean?



**Here, E = 0.01 and p = 0.5 (for a VERY conservative estimate).**

**Plugging in, we get n = 9604 (rounding up).**

1. Repeat part d but now assume plus or minus 5 percentage points.

**Here E = 0.05 and p = 0.5 as before.**

**Plugging in, we get 385 (rounding up).**

**Problem 6.** Skateboards are assembled in a particular manufacturing plant on an assembly line by three workers. Worker 1 performs the first three assembly steps, worker 2 performs the next three, and worker 3 performs the final three.   
  
Suppose we have collected the following 10 observations of the time (in seconds) it takes worker 1 to perform his three assembly steps: (see also accompanying excel spreadsheet)

|  |
| --- |
| 12.26 |
| 12.25 |
| 15.59 |
| 11.09 |
| 19.17 |
| 10.99 |
| 12.10 |
| 15.24 |
| 13.50 |
| 15.17 |

1. What is the sample mean and the sample standard deviation?

**Sample mean = 13.74**

**Sample Standard Deviation = 2.56**

1. Do the data appear to be drawn from a normal distribution (Hint: plot a histogram). Do you think 10 observations is enough for the CLT to apply? How many do you think you would need?

**This doesn’t look very normally distributed to me. It looks more like a uniform distribution, but it is hard to tell for sure with so few observations.**

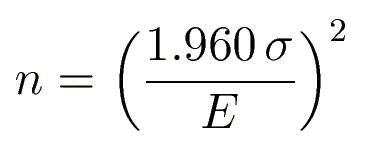
**If it is uniform, we would most likely need around 30 observations for the CLT to apply.**

1. What are your estimates for the mean and standard deviation of the sampling distribution?

**Estimate for Mean of the sampling distribution = 13.74**

**Estimate for std. dev. Of the sampling distribution = sample std. error = 2.56/sqrt(10) = 0.81 (approx.)**

1. How large a sample size will you need to collect if you want to have a 95% confidence level that your estimate of the true average time is within plus or minus 0.5 seconds of the true mean?



**Here, σ = 2.56 and E = 0.5. Plugging in & rounding up, we get n = 101.**

1. Suppose you take a sample of size 100 and obtain the same sample mean and sample standard deviation you calculated in part a. Construct a 95% confidence level about the mean.

**Estimate for Mean of the sampling distribution = 13.74**

**Estimate for std. dev. Of the sampling distribution = sample std. error = 2.56/sqrt(100) = 0.256 (approx.)**

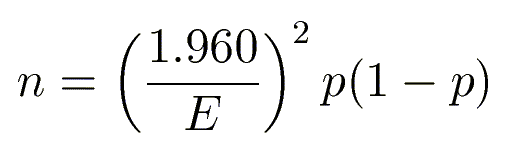
**For a 95% CI, we use NORMSINV(.975) = 1.96 (approx.)**

**Lower confidence limit = 13.74 – 1.96 x 0.256 = 13.24 (approx)**

**Upper confidence limit = 13.74 + 1.96 x 0.256 = 14.24 (approx)**

**Problem 7.** Senator Bumpus is facing a stiff primary challenge from T.Parte. A polling firm has collected data (see accompanying excel spreadsheet) from a randomly selected sample of Senator Bumpus’ constituents.

1. Without even looking at the data, what is a conservative estimate of the sample size needed if we want a 95% confidence level that the true mean is within plus or minus 3 percentage points of our estimate of Senator Bumpus’ support?



**Here E = 0.03 and p = 0.5 (for a conservative estimate). Plugging in we get 1068 (rounding up).**

1. Using the data in the spreadsheet, derive a point estimate for the level of Senator Bumpus’ support.

**175 out of 353 respondents prefer Bumpus.**

**This gives us a point estimate of 175/353 = 49.6% (approx.)**

1. Derive a 95% confidence level about the point estimate.

**Sample proportion, P = 0.496.**

**Sample standard error = sqrt[P x (1 – P)/n] = sqrt[0.496 x 0.504 / 353] = 0.027 (approx.)**

**For a 95% CI, we use NORMSINV(.975) = 1.96 (approx.)**

**Lower confidence limit = .496 – 1.96 x 0.027 = 0.444 (approx)**

**Upper confidence limit = .496 + 1.96 x 0.027 = 0.548 (approx)**