CONFIDENCE INTERVAL ASSIGNMENT

*1. Suppose we want to estimate the average weight of an adult male in DeKlab County, Georgia. We draw a random sample of 1,000 men from a population of 1000,000 men and weigh them. We find that the average man in our sample weighs 180 pounds, and the standard deviation of the sample is 30 pounds. What is the 95% confidence interval?

$\mu_{SD} = 180, \sigma_{SD} = 30, n = 1000$

$$\text{Standard Error (SE)} = \frac{\sigma}{\sqrt{n}} = \frac{30}{\sqrt{1000}} = 0.95$$

$Z \Rightarrow C \text{ level} \Rightarrow 95\%$

from Z-table

$\Rightarrow 2 \ (1.96)$

$$\mu_{pop} \in \mu_{SD} \pm Z * SE$$

$$= 180 \pm 2 * 0.95$$

$$= 180 \pm 1.9$$

$$* \ 95\% \Rightarrow \underbrace{178.1 \longleftrightarrow 181.9}_{\text{Confidence Interval.}}$$

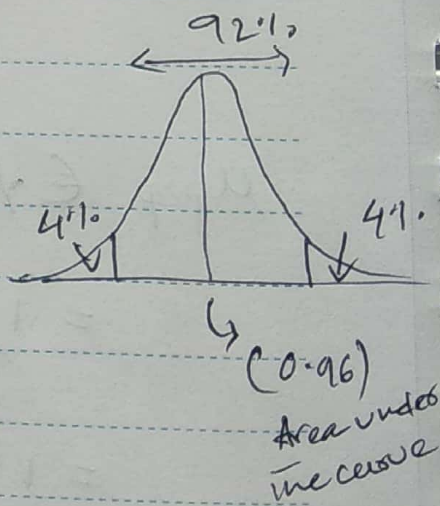
*2. The operations manager of a large production plant would like to estimate the mean amount of time a worker takes to assemble a new electronic component. Assume that the standard deviation of this assembly time is 3.6 minutes.

a) After observing 120 workers assembling similar devices, the manager noticed that their average time was 16.2 minute. Construct a 92% confidence interval for the mean assembly time.

b) How many workers should be involved in this study in order to have the mean assembly time estimated upto ±15 seconds with 92% confidence?

a) $n = 120$, $\sigma_{pop} = 3.6$, $\mu_s = 16.2$

$Z \Rightarrow$ Confidence Level $= 92\% = 1.76$



$$\mu_{pop} \in \mu_s \pm Z \times SE$$

$$= \overset{16.2}{\cancel{3.6}} \pm (1.76) \times \frac{3.6}{\sqrt{120}}$$

$$= 16.2 \pm (1.76) \times 0.328$$

$$= 16.2 \pm 0.577$$

$$= 15.623 \longleftrightarrow 16.777$$

Scanned with CamScanner

b) ~~MH~~

$\pm 15 sec. = Z * SE$

$\pm 0.25 min = Z * \frac{\sigma_{pop}}{\sqrt{n}}$

$= Z * \frac{3.6}{\sqrt{n}}$

$= Z_{0.92} * \frac{3.6}{\sqrt{n}}$

$\Rightarrow \sqrt{n} = \frac{Z_{0.92} \times 3.6}{0.25}$

$\Rightarrow n = \left(\frac{Z_{0.92} \times 3.6}{0.25}\right)^2$

$= \left(\frac{1.76 \times 3.6}{0.25}\right)^2$

$= 642.31.$

Scanned with CamScanner

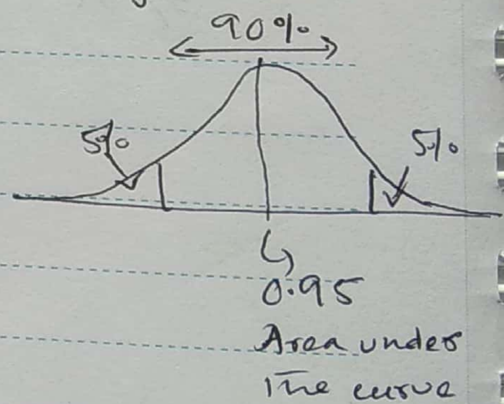
*3. Suppose a consumer advocacy group would like to conduct a survey to find the proportion 'p' of consumers who bought the newest generation of an MP3 player were happy with their purchase.

a) How large a sample 'n' should they take to estimate 'p' with 2% margin of error and 90% confidence?

b) The advocacy group took a random sample of 1000 consumers who recently purchased this MP3 player and found that 400 were happy with their purchase. Find a 95% confidence interval for 'p'.

(As conservative guess)

a) $ME = Z * SE$, $P_0 = 0.5$



$$\Rightarrow 0.02 = Z_{0.90} * \sqrt{\frac{P_0(1-P_0)}{n}}$$

$$= 1.65 * \sqrt{\frac{P_0(1-P_0)}{n}}$$

$$\Rightarrow \sqrt{n} = \frac{1.65 * \sqrt{P_0(1-P_0)}}{0.02}$$

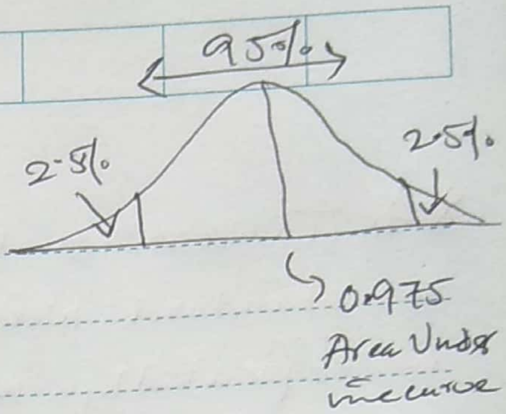
$$\Rightarrow n = \left(\frac{1.65 \times \sqrt{(0.5)(0.5)}}{0.02}\right)^2$$

$$= \frac{(1.65)^2 \times 0.25}{(0.02)^2}$$

$$= \frac{0.6806}{0.0004} = 1701.5$$

If we take Z-value 1.64 then

$$= \left(\frac{1.64 \times \sqrt{(0.5)(0.5)}}{0.02}\right)^2$$



b) $n = 1000$

$\hat{p} = \frac{400}{1000} = 0.4$

95% Confidence Interval for 'p' is

$$= \hat{p} \pm Z * SE$$

$$= \hat{p} \pm Z_{0.95} * \sqrt{\frac{\hat{p}_a(1-\hat{p}_a)}{n}}$$

$$= 0.4 \pm 1.96 \times \sqrt{\frac{(0.4)(0.6)}{1000}}$$

$$= 0.4 \pm 1.96 \times 1.55$$

$$= 0.4 \pm 3.038$$

$$= \overset{2.638}{\cancel{3.438}} \longleftrightarrow 3.438$$

4. To assess the accuracy of a laboratory scale, a standard weight that is known to weigh 1 gram is repeatedly weighed 4 times. The resulting measurements (in gram) are : 0.95, 1.02, 1.01, 0.98. Assume that the weighings by the scale when the true weight is 1 gram are normally distributed with mean μ.

a) Use these data to compute a 95% confidence interval μ.

b) Do these data gives evidence at 5% significance level that the scale is not accurate?

a) Here, $n = 4$

$$\mu_{SD} = \frac{0.95 + 1.02 + 1.01 + 0.98}{4}$$

$$= 0.99$$

$$\sigma_{SD} = \sqrt{\frac{(.95-0.99)^2 + (1.02-0.99)^2 + (1.01-0.99)^2 + (.98-.99)^2}{4}}$$

$$= \sqrt{\frac{(0.04)^2 + (0.03)^2 + (0.02)^2 + (-0.01)^2}{4}}$$

$$= \sqrt{\frac{0.0016 + 0.0009 + 0.0004 + 0.0001}{4}}$$

$$= \sqrt{0.00075} = 0.0273$$

Degree of freedom (df) $= n-1 = 3$

Since, $n < 30$, t-value for 95% confidence with df=3 is, $t = 3.182$ (from t-table)

Scanned with CamScanner

Cofidence Interval (CI) $= \mu_{SD} \pm t * \frac{\sigma_{SD}}{\sqrt{n}}$

$= 0.99 \pm 3.182 * \frac{0.0273}{\sqrt{4}}$

$= 0.99 \pm 3.182 * 0.0137$

$= 0.99 \pm 0.0436$

$\Rightarrow 0.9464 \longleftrightarrow 1.034$

b) We need to test null hypothesis $H_0 : \mu = 1$ against two sided alternative hypothesis $H_1 : \mu \neq 1$

Since, null value of $\mu = 1$ falls in the 95% confidence interval computed in the previous part, it follows that the 5% level of test does not reject H_0.

There is no evidence at 5% significance level that the scale is inaccurate.

*5. The time needed for college students to compute a certain maze follows a normal distribution with a mean of 45 seconds. To see if the mean time μ (in seconds) is changed by vigorous exercise, we have a group of nine college students exercise vigorously for 30 minutes and then complete the maze. The sample mean and standard deviation of the collected data is 49.2 seconds and 3.5 seconds respectively. Use these data to perform an appropriate test of hypothesis at 5% level of significance.

$n = 9$, $\mu_{pop} = 45$, $\mu_{sp} = 49.2$, $\sigma_{sp} = 3.5$

Null Hypothesis $H_0 : \mu = 45$

Alternative " $H_1 : \mu \neq 45$

at level $\alpha = 0.05$

Since, $n < 30$, we would do a t-test

Degree of freedom, $df = n-1 = 8$

The rejection region is $|T| > t_{n-1,\alpha/2} = t_{8,0.025}$

$= 2.306$ (from T-table)

$$T = \frac{x - \mu}{s/\sqrt{n}} = \frac{49.2 - 45}{3.5/\sqrt{9}}$$

$$= 3.6$$

Since, $|T| = 3.6 > 2.306$, H_0 is rejected

$\boxed{\text{Thus, there is a significance evidence at 5\% significance level that the mean time to complete the maze is changed after exercise.}}$

*6. Installation of a certain hardware takes a random amount of time with a standard deviation of 5 minutes. A computer technician installs this hardware on 64 different computers, with the average installation time of 42 minutes. Compute a 95% confidence interval for the mean installation time.

$\sigma = 5$, $n = 64$, $\mu = 42$

95% CI = ?

$$= \mu_s \pm z * \frac{\sigma}{\sqrt{n}}$$

$$= 42 \pm (1.96) * \frac{5}{\sqrt{64}}$$

$$= 42 \pm 1.225$$

$$= 40.8 \longleftrightarrow 43.2 \quad \Big\{ (40.8, 43.2)$$

The mean installation time is estimated to be between 40.8 min and 43.2 mins, with 95% confidence.

#8. What is the smallest sample size required to provide a 95% confidence interval for a mean, if it is important that the interval be no longer than 1 cm? You may assume that the population is normal with variance 9cm^2.

$$\sigma = \sqrt{9} = 3,$$

$$\cancel{n = \mu_{SD} \pm z * \frac{\sigma}{\sqrt{n}}}$$

$$\cancel{n = \left(\frac{z * \sigma}{\mu_{SD}}\right)^2}$$

=

$$\text{Margin of Error (ME)} = z * \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow 0.5 = 1.96 * \frac{3}{\sqrt{n}}$$

(ME is half of the width of CI)

$$\Rightarrow \sqrt{n} = \frac{1.96 * 3}{0.5}$$

$$\Rightarrow n = (11.76)^2$$

$$= 138.29$$

$$\approx 139$$

Q9. The recommended retail price of a brand of designer jeans is $150. The price of the jeans in a sample of 16 retailers is on average $141 with a sample standard deviation of 4. If this is a 'random' sample and the prices can be assumed to be normally distributed, construct a 95% confidence interval for the average sale price.

$\mu_{SD} = 141$, $\sigma = 4$, $n = 16$

Since $n < 30$, hence we will use T-Test.

$$t_{n-1, \alpha/2} = t_{16-1, 0.05/2}$$
$$= t_{15, 0.025}$$
$$= 2.131$$

Degree of freedom
$df = n-1$
$= 15$

$$T = \frac{x - \mu}{\sigma/\sqrt{n}}$$
$$= \frac{140 - 150}{4/\sqrt{16}}$$
$$= -9$$

95% CI for the average sale price is

141 ± 2.13

$= (138.87, 143.13)$

*10. Alcohol abuse has been described by college presidents as the number ~~of~~ one problem on campus, and is an important cause of death in young adults. A survey of 17,096 students in U.S. four-year colleges collected information on drinking behaviours and alcohol-related problems. The researches defined 'frequent binge-drinking' as having five or more drinks in a row three or more times in the past two weeks. According to their definition, 3,314 students were classified as frequent binge-drinkers. Construct 90% confidence interval around the true proportion of binge-drinkers.

$n = 17{,}096$

$\cancel{\mu_s = \frac{3314}{17{,}096}}$ The sample proportion, $\hat{p} = \frac{3314}{17{,}096}$

$= 0.1938$

90% CI of binge-drinkers is

$$\hat{p} \pm Z^*_{\alpha/2}\, SE$$

$$\Rightarrow 0.1938 \pm Z_{0.10/2} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\Rightarrow 0.1938 \pm 1.64 * \sqrt{\frac{0.1938(1-0.1938)}{17{,}096}}$$

$$\Rightarrow 0.1938 \pm 1.64 * 0.003$$

$$\Rightarrow 0.1938 \pm 0.0049$$

$$\Rightarrow (0.1889,\ 0.1987)$$

11. A random sample of 100 items is taken, producing a sample mean of 49. The population Std. deviation is : 4.49 construct a 90% confidence interval to estimate the population mean.

$n = 100$, $\mu = 49$, $\sigma = 4.49$

90% CI to estimate the population mean is

$$\mu \pm Z * SE$$

$$\Rightarrow \mu \pm Z * \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow 49 \pm 16.4 * \frac{4.49}{\sqrt{100}}$$

(Z-score of 90%)

$$\Rightarrow 49 \pm 16.4 * 0.449$$

$$\Rightarrow 49 \pm 7.3636$$

$$\Rightarrow (41.6364, 56.3636)$$

12. Click fraud has become a major concern as more companies advertise on the internet. When Google places an ad for a company with its search results, the company pays a fee to Google each time someone clicks on the link. That's fine when it's a person who's interested in buying a product or service, but not so good when it's a computer program pretending to be a customer. An analysis of 1200 clicks coming into a company's site during a week identified that 175 of these clicks are fraudulent. Compute the confidence interval with 95% confidence for the proportion of fraudulent clicks.

$$n = 1200$$

The sample proportion $\hat{p} = \frac{175}{1200} = 0.1458$

95% CI for the proportion of fraudulent clicks is;-

$$\hat{p} \pm Z * SE$$

$$\Rightarrow \hat{p} \pm Z_{\alpha/2} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\Rightarrow 0.1458 \pm Z_{\alpha/2} * \sqrt{\frac{0.1458(1-0.1458)}{1200}}$$

(0.05)

$$\Rightarrow 0.1458 \pm 1.96 * 0.0102$$

$$\Rightarrow 0.1458 \pm 0.0199$$

$$\Rightarrow (0.1259,\ 0.1657)$$

13.

*√ Of the 59 basket-ball players, 15 are left-handed. find the true percentage of left-handed players at 95% confidence.

$n = 59$

The sample proportion, $\hat{p} = \cancel{\frac{59}{15}} = \cancel{3.9333}$

$= \frac{15}{59} = 0.2542$

95% CI of left handed players is

$\hat{p} \pm Z_{\alpha/2} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$\Rightarrow \cancel{3.9333 \pm 1.96 * \sqrt{\frac{3.9333(1-3.9333)}{59}}}$

$\Rightarrow \overset{0.2542}{\cancel{3.9333}} \pm 1.96 * \sqrt{\frac{(0.2542)(1-0.2542)}{59}}$

$\Rightarrow \overset{0.2542}{\cancel{3.9333}} \pm 1.96 * 0.0567$

$\Rightarrow \overset{0.2542}{\cancel{3.9333}} \pm 0.1111$

$\Rightarrow \overset{0.1431,\ 0.3653}{(\cancel{3.8222}, \cancel{4.0444})}$

14. A group of investors wants to develop a chain of fast food restaurants. In determining the potential costs for each facility, they must consider, among other expenses, the average monthly electric bill. They decide to sample fast food restaurants currently opening to estimate the monthly costs of electricity. They want to be 90% confident of their results and want the error of the interval estimate to be no more than \$100. They estimate the such bill have the standard deviation 475. How large the sample should they take?

$n = ?$

$\sigma = 475$,

$$ME = Z * SE$$

$$= Z * \frac{\sigma}{\sqrt{n}}$$

$$Z_{\alpha/2} = Z_{0.1/2} = 1.64$$

$$\Rightarrow 100 = 1.64 * \frac{475}{\sqrt{n}}$$

$$\Rightarrow \sqrt{n} = 1.64 * \frac{475}{100}$$

$$\Rightarrow n = (1.64 * 4.75)^2$$

$$= 60.68$$

$$\simeq 61$$

#15. A marketing director of a large department store wants to estimate the average no. of customers who enter the store every five minutes. She randomly selects five-minutes intervals and counts the no. of arrivals at the store. She obtains the figures 68, 42, 51, 57, 56, 80, 45, 39, 36 and 79. Using this data, the analyst computes a 95% confidence interval to estimate the mean value for all five-minute intervals. What interval value does she get?

$n = 10$

$$\mu_{SD} = \frac{68+42+51+57+56+80+45+39+36+79}{10}$$

$$= \frac{553}{10} = 55.3$$

$$\sigma_{SD} = \sqrt{\frac{(68-55.3)^2+(42-55.3)^2+(51-55.3)^2+(57-55.3)^2+(56-55.3)^2+(80-55.3)^2+(45-55.3)^2+(39-55.3)^2+(36-55.3)^2+(79-55.3)^2}{10}}$$

$$= \sqrt{\frac{(12.7)^2+(-13.3)^2+(-4.3)^2+(1.7)^2+(0.7)^2+(24.7)^2+(-10.3)^2+(-16.3)^2+(-19.3)^2+(23.7)^2}{10}}$$

$$= \sqrt{\frac{2268.21}{10}} = 15.06$$

Degree of freedom, df $= n-1 = 9$

Since, $n < 30$, t-value for 95% confidence with df = 9 is = 2.262

Confidence Interval (CI) $= \mu_{SD} \pm t * \frac{\sigma_{SD}}{\sqrt{n}}$

$$= 55.3 \pm 2.262 * \frac{15.06}{\sqrt{10}}$$

$$= 55.3 \pm 2.262 * 3.162$$

$$= 55.3 \pm 7.152$$

$$= (48.148, 62.452)$$