Practice Problems

I. The weight of oranges produced by the MoneyWorth Orchards is distributed normally with mean μ and standard deviation σ. The oranges which weigh more than 266 gms. are exported and those with weight less than 100 gms. are sold to the local fruit juice factory. When MoneyWorth Orchards estimated a two-sided 95% confidence interval for μ based on the known value of σ, the width of the interval turned out to be 39.2. Considering that the standard error is inversely proportional to the square root of the sample size, they increased the sample size by 300 more oranges, and the width was exactly halved (i. e., it became 19.6). They have exported 12.30% of their production.

1. What is the value of σ?
2. What is the value of μ?
3. What percentage of the production was sold to the local fruit juice factory?
4. What should be sample size required if MoneyWorth orchards wants a 95% confidence interval with a width of ± 15 gms?

Solution:

Let X: Weight of oranges produced by the MW orchards

$\boxed{\text{Given:}}$

- $X \sim N(\mu, \sigma^2)$
- If $X > 266$ gms ; exported
- If $X < 100$ gms ; sold to local factory
- width of 95% CI is 39.2
- When 'n' changes to 'n+300' the width of 95% CI is 19.6 (n: sample size)
- $P[X > 266 \text{ gms}] = 0.123$

(1) Width of 95% CI for μ :

$$2 \times 1.96 \times \frac{\sigma}{\sqrt{n}} = 39.2$$

$$\Rightarrow \frac{\sigma}{\sqrt{n}} = 10 \quad \text{—— (1)}$$

when n increases by 300, we have

$$2 \times 1.96 \times \frac{\sigma}{\sqrt{n+300}} = 19.6$$

$$\Rightarrow \frac{\sigma}{\sqrt{n+300}} = 5 \quad \text{—— (2)}$$

Solving (1) & (2) we get

$$\boxed{\begin{aligned} \sigma &= 100 \text{ gms} \\ n &= 100 \end{aligned}}$$

* $1.96 = z_{\alpha/2}$; $\alpha = 0.05$

(2) $P(X > 266) = 0.123$

$\Rightarrow P\left(\frac{X-\mu}{\sigma} > \frac{266-\mu}{\sigma}\right) = 0.123$

$\Rightarrow P\left(Z > \frac{266-\mu}{100}\right) = 0.123$

$\therefore 1.16012 = \frac{266-\mu}{100}$

$\Rightarrow \boxed{\mu = 149.988 \text{ gms}}$

(3) $P(X < 100) = ?$

$\Rightarrow P(X < 100) = P\left(Z < \frac{100-\mu}{\sigma}\right)$

$= P\left(Z < \frac{100 - 149.988}{100}\right)$

$= P(Z < -0.49988)$

$= \boxed{0.3086}$

$\therefore$ 30.86% of produce was sold in local fruit juice factory.

④ Width = ± 15 gms or 30 gms

or

$$2 \times 1.96 \times \frac{\sigma}{\sqrt{n}} = 30$$

$$\Rightarrow \quad 2 \times 1.96 \times \frac{100}{\sqrt{n}} = 30$$

$$\Rightarrow \quad n \approx 171$$

II. Many public polling agencies conduct surveys to determine the current consumer sentiment concerning the state of the economy. One such agency randomly sampled 484 consumers and found that 257 were optimistic about the state of the economy.

1. Develop a 95% confidence interval for the proportion of consumers who are optimistic about the state of the economy.
2. Based on the above, is it possible to conclude that the majority of the consumers are optimistic about the state of the economy?
3. If the true proportion of consumers optimistic about the economy was 0.5, what is the probability that 257 or more in a sample of 484 are optimistic about the state of the economy?

Solution:

Given: π : population proportion of consumers who are optimistic about the state of the economy.

p : sample proportion of "

$n = 484$

$p = \frac{257}{n} = 0.531$

(1) 95% CI for π

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

$p = 0.531$, $\alpha = 0.05$, $z_{\alpha/2} = 1.96$, $n = 484$

$\Rightarrow$ 95% CI for π is

$$0.531 \pm 1.96 \sqrt{\frac{0.531 \times 0.469}{484}}$$

or

$$\boxed{[0.4865, \; 0.5755]}$$

(2) No. The lower limit is smaller the 0.5.

(3) $\pi = 0.5$

$$P\left(p \geq \frac{257}{484}\right)$$

$$= P\left[Z \geq \frac{0.531 - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \right]$$

$$= P\left[Z \geq \frac{0.531 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{484}}} \right]$$

$$= P\left[Z \geq 1.364 \right]$$

$$= 0.0863$$

$\approx$ 9% probability.

III. Dhanvin Karthik has recently joined MoGames Unlimited as the senior analyst. MoGames is in the mobile games space and has a number of popular games. The revenues of MoGames depend on the time spent on each game by the users. Even though the games are free downloadable, the users will have to register and in the process, MoGames collects a large amount demographic data on the users. Dhanvin was asked by the CMD to analyze the data on a particular game called CandyCrash. He extracted the data for a specific day (8 Sep 2016) with respect to users in the age group of 22 – 25 years from the database. From this group, he randomly selected 41 users and the data is summarized below:

	Male	Female
Sample size	25	16
Sample mean	180 minutes	140 minutes
Sample standard deviation (s)	45 minutes	55 minutes

σ : unknown

1. What is the lower limit of a 95% two-sided confidence interval for the average time of playing for male users (μ_M) (rounded off)?

$$\bar{X}_m - t_{(0.025,\,25-1)} \cdot \frac{s_m}{\sqrt{n}} = 180 - 2.064 \times \frac{45}{\sqrt{25}} = 161 \text{ mins}$$

2. What is the upper limit of a 95% two-sided confidence interval for the average time of playing for male users (μ_M) (rounded off).

$$\bar{X}_m + t_{(0.025,\,24)} \cdot \frac{s_m}{\sqrt{n}} = 180 + 2.064 \times \frac{45}{\sqrt{25}} = 199 \text{ mins}$$

3. The CMD wanted a 99% confidence interval (two sided) for the average time of playing for Female users (μ_F). What is the increase in the total width of the confidence interval, in minutes, as compared to the 95% interval calculated for the same (μ_F)?

Width of 95% CI: $2 \times t_{(0.025,\,16-1)} \cdot \frac{s_f}{\sqrt{n_F}}$

Width of 99% CI: $2 \times t_{(0.005,\,16-1)} \cdot \frac{s_f}{\sqrt{n_F}}$

$\therefore$ increase in width is $2 \times \frac{s_f}{\sqrt{n_f}} \cdot \left\{ t_{(0.005,\,15)} - t_{(0.025,\,15)} \right\}$

$$= 2 \times \frac{55}{\sqrt{16}} \left\{ 2.9467 - 2.13145 \right\}$$

$$= 22.42.$$

4. If we calculate separate confidence intervals for male and female users (using the same confidence level), one of the two is always likely to have a smaller width. Why is this true?

The margin of error will be smaller for males as $\frac{s_m}{\sqrt{n_m}} < \frac{s_f}{\sqrt{n_f}}$.

IV. The Indian army, after the recent clashes, decided to evaluate a new gun called Bi-fours (it is so called because its range is claimed to be at least16 kms). They have hired the services of Kanaka, who recently graduated from IIMB. Kanaka selected a simple random sample of size 25 and calculated the sample average. Based on the sample average and the σ as given by the manufacturers, she calculated a 90%, two sided confidence interval for the population mean, μ. The value of σ given by the manufacturers was 1500 meters.

1. What is the standard error of the sample mean?

$$\frac{\sigma}{\sqrt{n}} = \frac{1500}{\sqrt{25}} = 300 \text{ meters}$$

2. What is the range of the above mentioned 90% confidence interval?

$$\left[\bar{x} - 2 \times 1.64 \times \frac{1500}{5}, \; \bar{x} + 2 \times 1.64 \times \frac{1500}{5}\right]$$

3. What percentage of the sample means will fall within ±1000 meters of the population mean (μ) $P[\mu - 1000 \le \bar{x} \le \mu + 1000] = ?$

$$P[-1000 \le \bar{x} - \mu \le 1000]$$

$$= P\left[\frac{-1000}{\sigma/\sqrt{n}} \le Z \le \frac{1000}{\sigma/\sqrt{n}}\right]$$

$$= P\left[\frac{-1000}{300} \le Z \le \frac{1000}{300}\right]$$

$$= P[-3.33 \le Z \le 3.33]$$

$$\approx 1.$$

4. What should be sample size (rounded off to the next integer), if she wants to achieve a confidence level of 95% and at the same time retain the width of the 90% confidence interval that she got with a sample size of 25?

width from 90% CI is

$$2 \times 1.64 \times \frac{\sigma}{\sqrt{n}} = 984$$

width from 95% CI is

$$2 \times 1.96 \times \frac{1500}{\sqrt{n}} = 984$$

solving for 'n' we get

$n \approx 36$.

V. Narayan Modi is contesting for elections from the Hafizpet constituency in Rangareddy District. In order to ascertain his chances in the election, he selected a simple random sample of 72 voters in the constituency and elicited their opinion through a secret ballot. 28 of the 72 voters were in his favour.

1. What is the lower limit of a 95 percent confidence interval for the proportion of votes in his favour (π) based on this sample data?

Solution: $n = 72$ voters

$p = \frac{28}{72}$ (sample proportion in his favour)

$= 0.39$

95% CI for π is

$$p \pm 1.96 \times \sqrt{\frac{p(1-p)}{n}}$$

$\Rightarrow$ lower limit is

$$0.39 - 1.96 \times \sqrt{\frac{0.39 \times 0.61}{72}}$$

$= 0.39 - 0.1127$

$= 0.2773$

2. If Narayan Modi wants to estimate the confidence interval (same 95% confidence level) within ± 0.005, what is maximum sample size required?

Solution:

$$n = \frac{z^2_{\alpha/2}pq}{B^2}$$

$z_{\alpha/2} = 1.96$

$p = 0.39$

$q = 1 - 0.39 = 0.61$

$B = 0.005 \times 2 = 0.01$

$$\therefore n = \frac{(1.96)^2 \times 0.39 \times 0.61}{(0.01)^2}$$

≈ 9139

** I'm unsure of the wording of this question. Is $B = 0.005$ or 0.01 or ± 0.005 from existing 95% CI?

Please confirm question meaning before moving forward.

3. Test the null hypothesis that the population proportion is greater than or equal to 50% using $\alpha = 0.10$. What is the p value associated with this test?

Solution:

$$H_0 : \pi \geq 0.5$$

Vs

$$H_1 : \pi < 0.5$$

$\alpha = 0.1$

The test statistic for testing this is:

$$Z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

; p = Sample proportion

p_0 : proportion under null hypothesis.

the z-value is $$z = \frac{0.39 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{72}}}$$

$$= -1.8667$$

p-value is $P(z < -1.18667)$

$$= 0.0309$$

VI. NS Software Systems and Services (popularly known as N4S) has recently recruited Komala, a quality expert from Institute of International Management, Bilekahalli (IIMB). She decided to train the employees of N4S in quality awareness as well as the techniques to be used through a new innovative method, developed specially by her. Dr. Raghav who is the VP, Corporate Training agreed to experiment with the new method of training, provided the effectiveness of the training can be measured. Komala and Dr. Raghav agreed to use a special instrument (as some of the OB experts prefer to call it), which has to be administered to the trainees before and after the training. Based on the scores of the instrument, before and after the training, a "Quality Capability Motivation Improvement Index (QCM)" can be computed for each of the trainees. They have agreed to use this index as a measure for the effectiveness of the training and a positive value of the index indicates effectiveness of the training.

It was proposed that a sample of 20 employees would be selected for training under the new method. For obvious reasons, they decided to do a one sided hypothesis test. *The sample standard deviation of the QCM obtained from these 20 observations was 12.5*.

1. Dr. Raghav informed Komala that he is willing to go ahead with the new method if the sample average of QCM is greater than equal to 4.8327. Based on this decision, calculate the Type I error that Dr. Raghav is willing to tolerate.

Let X_1 : Scores before training

X_2 : Scores after training

μ_1 : population mean score before training

μ_2 : population mean score after training.

Prob. of type I error:

$$P(\bar{X}_1 - \bar{X}_2 > 4.8327)$$

$$\Rightarrow P\left(\frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{S_{\bar{x}_1 - \bar{x}_2}/\sqrt{n}} > \frac{4.8327 - 0}{12.5/\sqrt{20}}\right)$$

$$\Rightarrow P(T_{n-1} > 1.7289) = 0.0500$$

↑ type I error level.

2. Based on the above, formulate the null and alternate hypothesis.
3. At this stage, Dr. Rahav decided to change the α value to 0.025 and suitably modified the decision rule. The average QCM of the 20 trainees turned out to be 5.3. Should Dr. Raghav go ahead with the new method of training?

Let μ be the population mean of QCM.

$$H_0 : \mu \leq 4.8327$$

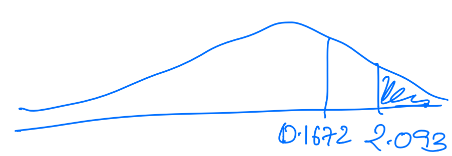
$$H_1 : \mu > 4.8327$$

If $\alpha = 0.025$ and $\bar{X} = 5.3$

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{5.3 - 4.8327}{12.5/\sqrt{20}}$$

$$= 0.1672$$

$$T_{0.025, n-1} = 2.093$$



Our evaluate test statistic falls outside the rejection region

$\Rightarrow$ we fail to reject H_0.

He should probably not go ahead with the new plan.

4. Assume that the true average QCM is 0.4673. What is the Type II error that Dr. Raghav will be committing, *based on the decision in question 3 above*?

Solution: We fail to reject H_0 $\therefore$ there is a possibility that we fail to reject a false H_0 $\therefore$ we may committ a type II error.

Probability of type II error is

$$P\left[\text{fail to reject } H_0 \mid H_0 \text{ is false}\right]$$

$$= P\left[\bar{x} \leq 4.8327 \mid \mu = 0.4673\right]$$

$$= P\left[t_{19} \leq \frac{4.8327 - 0.4673}{12.51/\sqrt{20}}\right]$$

$$= P\left[t_{19} \leq 1.5618\right] = 0.0674$$