1. **(10 marks)** A sample of size $n = 100$ produced the sample mean of $\bar{X} = 16$. Assuming the population standard deviation $\sigma = 3$, compute a 95% confidence interval for the population mean μ.

Sol:

1.

A 95% confidence interval for μ is

$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

where $\alpha = 0.05$, so $z_{\alpha/2} = z_{0.025} = 1.96$ from the table of Normal distribution.

Then, the 95% confidence interval for μ is

$$16 \pm (1.96)\frac{3}{\sqrt{100}} = 16 \pm 0.588 = [15.412, 16.588]$$

2. **(10 marks)** Assuming the population standard deviation $\sigma = 3$, how large should a sample be to estimate the population mean μ with a margin of error not exceeding 0.5?

Sol:

2.

$n \geq (z_{\alpha/2}\sigma/\Delta)^2$, where $\alpha = 0.05$, the critical value $z_{\alpha/2} = z_{0.025} = 1.96$, and the required margin $\Delta = 0.5$.

Then we need a sample size of

$$n \geq \left(\frac{(1.96)(3)}{0.5}\right)^2 = 138.3,$$

i.e., we need a sample of size at least 139.

3. **(10 marks)** We observed 28 successes in 70 independent Bernoulli trials. Compute a 90% confidence interval for the population proportion p.

Sol:

3.

A 90% confidence interval for p is

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

where the sample proportion $\hat{p} = 28/70 = 0.4$ and the sample size $n = 70$.

Then, the 90% confidence interval for p is

$$0.4 \pm (1.645)\sqrt{\frac{(0.4)(0.6)}{70}} = 0.4 \pm 0.096 = [0.304, 0.496].$$

4. **(10 marks)** 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 minutes. 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 up to ± 15 seconds with 92% confidence?

Sol:

4.

Let μ denote the mean assembly time (in minutes). It is given that $\sigma = 3.6$ min.

a) We want a 92% confidence interval for μ based on the following information: $n = 120$, $\overline{X} = 16.2$ min, $\alpha = 1 - 0.92 = 0.08$, and $\sigma = 3.6$ min. Since σ is known, we will use the z-interval:

$$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 16.2 \pm (1.75)\frac{3.6}{\sqrt{120}} = 16.2 \pm 0.575 = [15.6, 16.8],$$

where $z_{0.04} = 1.75$ obtained from the normal table. Thus, the mean assembly time for a worker is estimated to be between 15.6 min and 16.8 min, with 92% confidence.

b) We need to compute the sample size n based on the following information: $\alpha = 0.08$, $\sigma = 3.6$, and margin for error $E = 15$ sec $= 0.25$ min. Thus,

$$n = \left\lceil \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 \right\rceil = \left\lceil \left(\frac{1.75(3.6)}{0.25}\right)^2 \right\rceil = 636.$$

Thus, 636 workers are need in the study to achieve the desired precision of inference.

5. **(10 marks)** 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.

Sol:

5.

a) It is given that margin of error $E = 0.02$ and $\alpha = 0.10$. Using $p = 0.5$ as the conservative guess in the sample size formula gives,

$$n = \left\lceil p(1-p)\left(z_{\alpha/2}/E\right)^2 \right\rceil = \left\lceil \left(z_{\alpha/2}/\{2E\}\right)^2 \right\rceil = \lceil (1.645/0.04)^2 \rceil = 1692.$$

b) From the data, $\hat{p} = 400/1000 = 0.40$. Since $n = 1000$ is large, the 90% confidence interval for p is:

$$\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} = 0.40 \pm 1.645\sqrt{0.40(0.60)/1000} = [0.375, 0.425].$$

Thus, between 37.5% to 42.5% consumers are estimated to be happy with their purchase.

6. **(10 marks)** In order to ensure efficient usage of a server, it is necessary to estimate the mean number of concurrent users. According to records, the sample mean and sample standard deviation of number of concurrent users at 100 randomly selected times is 37.7 and 9.2, respectively.

a) Construct a 90% confidence interval for the mean number of concurrent users.

b) Do these data provide significant evidence, at 1% significance level, that the mean number of concurrent users is greater than 35?

Sol:

6.

Let μ denote the mean number of concurrent users in the population. It is given that $n = 100$, $\overline{X} = 37.7$ and $S = 9.2$.

a) We want a 90\% confidence interval for μ. Since n is large, we will use the large-sample z-interval: $\overline{X} \pm z_{\alpha/2}\frac{S}{\sqrt{n}}$. From the normal table, $z_{0.05} = 1.645$. Thus, the desired confidence interval is:

$$37.7 \pm (1.645)\frac{9.2}{\sqrt{100}} = 37.7 \pm 1.5 = [36.2, 39.2].$$

b) We need to test the null $H_0 : \mu = 35$ against the one-sided alternative $H_1 : \mu > 35$, at level $\alpha = 0.01$. Since n is large, we will do a large-sample z-test. The rejection region is $Z > z_\alpha = 2.33$, using the normal table. The test statistic is

$$Z = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{37.7 - 35}{9.2/\sqrt{100}} = 2.93$$

Since $Z = 2.93 > 2.33$, H_0 is rejected. Thus, there is significant evidence at 1% significance level that the mean number of concurrent users is greater than 35.

7. **(10 marks)** 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 grams) 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 for μ.

   b) Do these data give evidence at 5% significance level that the scale is not accurate? Answer this question by performing an appropriate test of hypothesis.

Sol:

7.

a. From the given data, we have: $n = 4$, $\overline{X} = 0.99$ and $S = 0.032$. Since n is small and the data are normally distributed we will use the t-interval:

$$\overline{X} \pm t_{n-1,\alpha/2}\frac{S}{\sqrt{n}} = 0.99 \pm 3.182\frac{0.032}{2} = 0.99 \pm 0.05 = [0.94, 1.04].$$

Thus, the mean weight is estimated to be between 0.94 to 1.04 grams, with 95% confidence.

b. We need to test the null hypothesis $H_0 : \mu = 1$ against the two-sided alternative $H_1 : \mu \neq 1$. Since the null value of $\mu = 1$ falls in the 95% confidence interval computed in the previous part, it follows that the 5% level test of does not reject H_0. Thus, there is no evidence at 5% significance level that the scale is inaccurate.

10. **(10 marks)** 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.

Sol:

10.

Let μ denote the mean installation time (in minutes). We want a 95% confidence interval for μ. The following quantities are given: $1 - \alpha = 0.95$, $n = 64$, $\overline{X} = 42$ min, and $\sigma = 5$ min. Since σ is known, we will use the z-interval:

$$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 42 \pm (1.96)\frac{5}{\sqrt{64}} = 42 \pm 1.225 = [40.8, 43.2],$$

where $z_{0.025} = 1.96$ is obtained from the normal table.Thus, the mean installation time is estimated to be between 40.8 min and 43.2 min, with 95% confidence.

11. (**10 marks**) An exit poll of 1000 randomly selected voters found that 515 favored candidate A. Is the race too close to call? Answer this question by performing an appropriate test of hypothesis at 1% level of significance.

Sol:

Sol:

11.

Let p denote the proportion of voters who favor candidate A. From the data, $\hat{p} = 515/1000 = 0.515$. We would like to test the null hypothesis $H_0 : p = 0.5$ against the two-sided alternative $H_1 : p \neq 0.5$, at $\alpha = 0.01$. Since $n = 1000$ is large, we will do a large-sample z-test. The rejection rejection is $|Z| > z_{\alpha/2} = 2.576$, using the normal table. The test statistic is:

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} = (0.515 - 0.5)/\sqrt{0.5 \cdot 0.5/1000} = 0.949$$

Since $|Z| = 0.949 < 2.576$, H_0 is not rejected. Thus, the race is too close to call.

12. **(10 marks)** Bags of a certain brand of tortilla chips claim to have a net weight of 14 ounces. The net weights actually vary slightly from bag to bag and are normally distributed with mean μ. A representative of a consumer advocacy group wishes to see if there is any evidence that the mean net weight is less than advertised. For this, the representative randomly selects 16 bags of this brand and determines the net weight of each. He finds the sample mean to be $\overline{X} = 13.82$ and the sample standard deviation to be $S = 0.24$. Use these data to perform an appropriate test of hypothesis at 5% significance level.

Sol:

Sol:

12.

We need to test the null hypothesis $H_0 : \mu = 14$ against the one-sided alternative hypothesis $H_1 : \mu < 14$, at level $\alpha = 0.05$. Since $n = 16$ is small and the data are normally distributed, we will do a t-test. The rejection region is $T < -t_{n-1,\alpha} = -t_{15,0.05} = -1.753$, using the t-table. The test statistic is

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{13.82 - 14}{0.24/4} = -3$$

Since $T = -3 < -1.753$, H_0 is rejected. Thus, there is evidence at 5% significance level that the mean weight is less than advertised.

13. **(10 marks)** The time needed for college students to complete a certain maze follows a normal distribution with a mean of 45 seconds. To see if the mean time 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.

Sol:

Sol:

13.

We need to test the null $H_0 : \mu = 45$ against the two-sided alternative $H_1 : \mu \neq 45$, at level $\alpha = 0.05$. It is given that $n = 9$, $\overline{X} = 49.2$ and $S = 3.5$. Since n is small and the data are normally distributed, we will do a t-test. The rejection region is $|T| > t_{n-1,\alpha/2} = t_{8,0.025} = 2.306$, using the t-table. The test statistic is

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{49.2 - 45}{3.5/\sqrt{9}} = 3.6$$

Since $|T| = 3.6 > 2.306$, H_0 is rejected. Thus, there is significant evidence at 5% significance level that the mean time to complete the maze is changed after the exercise.

Example:

A statistician chooses 27 randomly selected dates, and when examining the occupancy records of a particular motel for those dates, finds a standard deviation of 5.86 rooms rented. If the number of rooms rented is normally distributed, find the 95% confidence interval for the population standard deviation of the number of rooms rented.

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$$

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} \leq \sigma \leq \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}}$$

Solution:

For a sample size of $n = 27$, we will have $df = n - 1 = 26$ degrees of freedom. For a 95% confidence interval, we have $\alpha = 0.05$, which gives 2.5% of the area at each end of the chi-square distribution. We find values of $\chi^2_{0.975} = 13.844$ and $\chi^2_{0.025} = 41.923$. Evaluating $\dfrac{(n-1)s^2}{\chi^2}$, we obtain 21.297 and 64.492. This leads to the inequalities $21.297 \leq \sigma^2 \leq 64.492$ for the variance, and taking square roots, $4.615 \leq \sigma \leq 8.031$ for the standard deviation.

Example:

Suppose a sample of 30 ECC students are given an IQ test. If the sample has a standard deviation of 12.23 points, find a 90% confidence interval for the population standard deviation.

Solution:

Solution: We first need to find the critical values:

$\chi^2{}_{1-\alpha/2} = \chi^2{}_{0.95,29} \approx 17.708$ and $\chi^2{}_{\alpha/2} = \chi^2{}_{0.05,29} \approx 42.557$

Then the confidence interval is:

$$\frac{(n-1)s^2}{\chi^2{}_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2{}_{1-\alpha/2}}$$

$$\frac{(30-1)12.23^2}{42.557} < \sigma^2 < \frac{(30-1)12.23^2}{17.708}$$

$$101.9249 < \sigma^2 < 244.9472$$

$$10.10 < \sigma < 15.65$$

So we are 90% confident that the standard deviation of the IQ of ECC students is between 10.10 and 15.65 bpm.

Example:

In a study on cholesterol levels a sample of 12 men and women was chosen. The plasma cholesterol levels (mmol/L) of the subjects were as follows: 6.0, 6.4, 7.0, 5.8, 6.0, 5.8, 5.9, 6.7, 6.1, 6.5, 6.3, and 5.8. We assume that these 12 subjects constitute a simple random sample of a population of similar subjects. We wish to estimate the variance of the plasma cholesterol levels with a 95 percent confidence interval.

Solution:

- Value of s^2

  s = .3918680978

- Values of χ^2 from table

  $\chi^2_{.975} = 21.920$

  $\chi^2_{.025} = 3.816$

Rcode:

```
data<-c(6.0, 6.4, 7.0, 5.8, 6.0, 5.8, 5.9, 6.7, 6.1, 6.5, 6.3, 5.8)

sd<-sd(data)

df=length(data)-1

LowerLimit=df*sd^2/qchisq(0.975,df)

LowerLimit

UpperLimit=df*sd^2/qchisq(0.025,df)

UpperLimit
```

Solution: (0.07706035, 0.442682)