Q.1. Assume that the alcohol content (in percentage) in a cough syrup taken from any particular brand is normally distributed with the true standard deviation 0.75 .
(a) Compute a 95% CI for the true average alcohol content of a certain brand if the average content for 20 specimens from the brand was 4.85 .
(b) Compute a 98% CI for the true average alcohol content of a another brand based on 16 specimens with a sample average alcohol content of 4.56 .
(c) How large a sample size is necessary if the width of the 95% interval is to be 0.40?
(d) What sample size is necessary to estimate the true average alcohol content to within 0.2 with 99%confidence?

Sol.1. Given that population is normally distributed with $\sigma = 0.75$.
(a) A 95% CI for mean of a normal population is given by:

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

Given that 1- α = 0.95, α = 0.05, n = 20, $\bar{x}$ = 4.85
Desired 95% CI is: $4.85 \pm z_{0.025}\frac{0.75}{\sqrt{20}}$ = 4.85±0.33 =(4.52, 5.18)

(b) Given that α=0.02, n=16, $\bar{x}$=4.56
Desired 98% CI is: $4.56 \pm z_{0.01}\frac{0.75}{\sqrt{16}}$ = (4.12, 5.00)

(c) Width of 100(1-α)% interval is given by:

$$w = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Therefore, $n = \left[\frac{2(1.96)(0.75)}{0.40}\right]^2$ = 54.02 ≈ 55

(d) Now, given that the true average alcohol content must lie within the interval (-0.2, 0.2).
Width=2(0.2)= 0.4
Therefore, $n = \left[\frac{2(2.58)(0.75)}{0.40}\right]^2$ = 93.61 ≈ 94

Q.2. A random sample of n=15 light bulbs of a certain type yielded the following observations on lifetime (in years):

2.0	1.3	6.0	1.9	5.1	0.4	1.0	5.3
15.7	0.7	4.8	0.9	12.2	5.3	0.6	

(a) Assume that the lifetime distribution is exponential. Obtain a 95% CI for expected (true average) lifetime.
(b) What is a 95% CI for the standard deviation of the lifetime distribution?

Sol.2. We know that h$(X_1, X_2, ..., X_n; \lambda) = 2\lambda \sum X_i$ has a chi-squared distribution with 2n degrees of freedom.
The relevant number of df here is 2(15)=30.The $\nu = 30$ row of chi-squared distribution table shows

that 46.979 captures upper-tail area 0.025 and 16.791 captures lower-tail area 0.025 .

Thus for n=15:

P(16.791<$2\lambda \sum X_i$<46.979)=0.95

Desired 95% CI for $\mu = \frac{1}{\lambda}$ is:

$$\Big(\frac{2\sum x_i}{46.979}, \frac{2\sum x_i}{16.791}\Big)$$

Here $\sum x_i$=63.2
Therefore, desired interval is (2.69,7.53).

(b) We know theta when X has an exponential distribution, then V(X)=$\frac{1}{\lambda^2}$. So, the standard deviation is $\frac{1}{\lambda}$ which is same as the mean.
Thus the interval calculated in part'a' is also a 95% CI for the standard deviation of the lifetime distribution.

Q.3. A company manufactures taps. Out of 143 manufactured taps, 10 are found to be defective. Calculate a lower confidence bound at the 95% confidence level for the true proportion of such taps that are defective. Also, interpret the 95% confidence level.

Sol.3. Here, α=0.05, n=143
For One sided bound,we need $z_\alpha = z_{0.05}$=1.645
Sample proportion of defective taps=$\frac{10}{143} = 0.07$

Desired 95% lower confidence bound for the true proportion of defective taps:

$$\frac{\hat{p} + \frac{z_\alpha^2}{2n}}{1 + \frac{z_\alpha^2}{n}} - z_\alpha \frac{\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z_\alpha^2}{4n^2}}}{1 + \frac{z_\alpha^2}{n}}$$

$$= \frac{\hat{.07} + \frac{1.645^2}{2(143)}}{1 + \frac{1.645^2}{143}} - 1.645 \frac{\sqrt{\frac{(0.07)(.93)}{143} + \frac{1.645^2}{4(143)^2}}}{1 + \frac{1.645^2}{143}}$$

$$= 0.078 - 0.036 = 0.042$$

We are 95% confident that more than 4.2% of all manufactured taps are defective.

Q.4. In a survey of 2003 adults, 25% said they believed in God.
(a) Calculate and interpret a confidence interval at the 99% confidence level for the proportion of all adults who believe in God.
(b) What sample size would be required for the width of a 99% CI to be almost .05 irrespective of the value of p̂?

Sol.4. With such a large sample size, we use simplified CI formula that is given by:

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Using, $\hat{p}$=0.25, n=2003, and $z_{\alpha/2} = z_{0.005}$=2.576
Desired 99% CI for p is:

$$0.25 \pm 2.576\sqrt{\frac{(0.25)(0.75)}{2003}}$$

$$= 0.25 \pm 0.025 =$$

$$(0.225, 0.275)$$

(b) Here, $\hat{p} = \hat{q} = 0.5$ Therefore,

$$n = \frac{4z^2\hat{p}\hat{q}}{w^2}$$

$$n = \frac{4(2.576)^2(0.5)(0.5)}{0.05^2} = 2654.31 \approx 2655$$

So, a sample of atleast 2655 is required.

Q.5. Let X_1, X_2, . . . ,X_n be a random sample from a continuous probability distribution having median $\tilde{\mu}$. Show that

(a) $\mathrm{P}\left[\min X_i < \tilde{\mu} < \max X_i\right] = 1 - \frac{1}{2^{n-1}}$, so that $(\min X_i,\ \max X_i)$ is a $100(1-\alpha)\%$ confidence interval for $\tilde{\mu}$ with $\alpha = 1/2^{n-1}$.

(b) For each of six normal male infants, the amount of the amino acid alanine (mg/100 mL) was determined while the infants were on an isoleucine-free diet, resulting in the following data:

2.84, 3.54, 2.80, 1.44, 2.94, 2.70

Compute a $100(1-\frac{1}{2^5})\%$ CI for the true median amount of alanine for infants on such a diet.

(c). Let $X_{(2)}$ denote the second smallest of the x_is and $X_{(n-1)}$ denote the second largest of the x_is. What is the confidence coefficient of the interval $(X_{(2)}, X_{(n-1)})$ for $\tilde{\mu}$?

Sol.5. (a) $\tilde{\mu}$ is the median, therefore

$$\mathrm{P}\left[X_i \leq \tilde{\mu}\right] = \mathrm{P}\left[X_i \geq \tilde{\mu}\right] = \frac{1}{2}, \text{ for all } i = 1..n.$$

Furthermore,

$$\begin{aligned}
\mathrm{P}\left[\min X_i \leq \tilde{\mu} \leq \max X_i\right] &= 1 - \mathrm{P}\left[(\tilde{\mu} < \min X_i)\ or\ (\tilde{\mu} > \max X_i)\right] \\
&= 1 - \mathrm{P}\left[\tilde{\mu} < \min X_i\right] - \mathrm{P}\left[\tilde{\mu} > \max X_i\right] \\
&= 1 - \mathrm{P}\left[\tilde{\mu} < X_1\right]...\mathrm{P}\left[\tilde{\mu} < X_n\right] - \mathrm{P}\left[\tilde{\mu} > X_1\right]...\mathrm{P}\left[\tilde{\mu} > X_n\right] \\
&= 1 - \frac{1}{2^n} - \frac{1}{2^n} = 1 - \frac{1}{2^{n-1}}.
\end{aligned}$$

(b) We have,

$$\min x_i = 1.44,\ \max x_i = 3.54 \text{ and } n = 6.$$

Therefore (from (a) part) (1.44, 3.54) is a $100(1-\frac{1}{2^5})\%$ confidence interval for $\tilde{\mu}$.

(c) We have

$$\mathrm{P}\left[X_{(2)} \leq \tilde{\mu} \leq X_{(n-1)}\right] = 1 - \mathrm{P}\left[\tilde{\mu} < X_{(2)}\right] - \mathrm{P}\left[\tilde{\mu} > X_{(n-1)}\right]$$

where,

$$\begin{aligned}\mathrm{P}\left[\tilde{\mu} < X_{(2)}\right] &= \mathrm{P}\left[\text{at most one } X_i \text{ is below } \tilde{\mu}\right] = \frac{1}{2^n} + \binom{n}{1}\frac{1}{2}\frac{1}{2^{n-1}} = \frac{1}{2^n}(1+n)\\ \mathrm{P}\left[\tilde{\mu} > X_{(n-1)}\right] &= \mathrm{P}\left[\text{at most one } X_i \text{ is above } \tilde{\mu}\right] = \frac{1}{2^n}(1+n)\end{aligned}$$

Hence,

$$\mathrm{P}\left[X_{(2)} \leq \tilde{\mu} \leq X_{(n-1)}\right] = 1 - \frac{1}{2^{n-1}}(1+n).$$

Therefore, $(X_{(2)}, X_{(n-1)})$ is a $100(1 - \frac{1}{2^{n-1}}(1+n))\%$ confidence interval for $\tilde{\mu}$.

Q.6. Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform distribution on the interval $[0, \theta]$, and let $U = \frac{1}{\theta}\max X_i$.

(a) Show that the rv U has density function

$$f_U(x) = \begin{cases} nx^{n-1} & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

(b) Verify that

$$\mathrm{P}\left[\left(\frac{\alpha}{2}\right)^{1/n} \leq U \leq \left(1 - \frac{\alpha}{2}\right)^{1/n}\right] = 1 - \alpha,$$

and

$$\mathrm{P}\left[\alpha^{1/n} \leq U \leq 1\right] = 1 - \alpha.$$

Use above identities to obtain two $100(1-\alpha)\%$ CI for θ.

(c) If the waiting time for a bus during morning hours is uniformly distributed and observed waiting times are $x_1 = 4.2$, $x_1 = 3.5$, $x_3 = 1.7$, $x_4 = 1.2$, and $x_5 = 2.4$, derive a 95% CI for θ by using both intervals.

Sol.6. (a)The CDF for U is given as

$$\begin{aligned}F_U(x) &= \mathrm{P}\left[U \leq x\right] = \mathrm{P}\left[\max X_i \leq x\theta\right]\\ &= \mathrm{P}\left[X_1 \leq x\theta\right] \ldots \mathrm{P}\left[X_n \leq x\theta\right]\\ &= x^n\end{aligned}$$

Therefore,

$$f_U(x) = F_U'(x) = \begin{cases} nx^{n-1} & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

(b)Integrating $f_U(x)$ between $\left(\frac{\alpha}{2}\right)^{1/n}$ and $\left(1-\frac{\alpha}{2}\right)^{1/n}$, one obtain

$$\mathrm{P}\left[\left(\frac{\alpha}{2}\right)^{1/n} \leq U \leq \left(1 - \frac{\alpha}{2}\right)^{1/n}\right] = \int_{\left(\frac{\alpha}{2}\right)^{1/n}}^{\left(1-\frac{\alpha}{2}\right)^{1/n}} f_U(x)dx = \left(1 - \frac{\alpha}{2}\right) - \frac{\alpha}{2} = (1-\alpha)$$

i.e.,

$$\mathrm{P}\left[\left(\frac{\alpha}{2}\right)^{1/n} \leq \frac{1}{\theta}\max X_i \leq \left(1 - \frac{\alpha}{2}\right)^{1/n}\right] = 1 - \alpha$$

or

$$\mathrm{P}\left[\frac{\max X_i}{\left(1-\frac{\alpha}{2}\right)^{1/n}} \leq \theta \leq \frac{\max X_i}{\left(\frac{\alpha}{2}\right)^{1/n}}\right] = 1 - \alpha \qquad (6(a))$$

Which implies that $\left[\frac{\max X_i}{\left(1-\frac{\alpha}{2}\right)^{1/n}}, \frac{\max X_i}{\left(\frac{\alpha}{2}\right)^{1/n}}\right]$ is an $100(1-\alpha)\%$ CI for θ.

Again, integrating $f_U(x)$ between $\alpha^{1/n}$ and 1, we get

$$\mathrm{P}\left[\alpha^{1/n} \leq U \leq 1\right] = 1 - \alpha$$

or, equivalently

$$\mathrm{P}\left[\max X_i \leq \theta \leq \frac{1}{\alpha^{1/n}} \max X_i\right] = 1 - \alpha. \tag{6(b)}$$

The last equation gives a second $100(1-\alpha)\%$ CI for θ, as $\left[\max X_i, \frac{1}{\alpha^{1/n}} \max X_i\right]$.

(c) For the given data $\min X_i = 1.2$, $\max X_i = 4.2$ and $n = 5$. Therefore 95% (i.e., $\alpha = 0.05$) CIs for θ from (6(a)) and (6(b)) are evaluated $[4.24, 8.78]$ and $[4.2, 7.646]$, respectively.

Q.7. Find the probability that a random sample of 25 observations, from a normal population with variance $\sigma^2 = 6$, will have a sample variance $S^2 > 9.1$.

Sol.7. We want to evaluate P$[S^2 > 9.1]$. We know that the rv

$$X = \frac{(n-1)S^2}{\sigma^2}$$

has chi-squared distribution with df $\nu = (n-1) = 24$. Thus

$$P(S^2 > 9.1) = P(\frac{24}{6}S^2 > \frac{24}{6}9.1) = P(\frac{24}{6}S^2 > 36.4).$$

From table $\chi^2_{0.05,24} \approx 36.4$, therefore $P(S^2 > 9.1) = 0.05$.