Chapter 3

INTERVAL ESTIMATION

3.1 Introduction

In many situations, a point estimate does not provide enough information about the parameter of interest. For example, if we are interested in estimating the mean compression strength of concrete, a single number may not be very meaningful. It would be more desirable to determine an interval within which we would expect to find the value of the parameter. An interval of the form

$$L \le \mu \le U$$

might be more useful.

Such an interval is called an interval estimate. The end points of this interval will be random quantities, since they are functions of the sample data.

Let $X_1,\ldots X_n$ be a random sample from a pdf with unknown parameter θ. Suppose that L and U are statistics, say $\mathbf{L} = \ell(\mathbf{X_1,\ldots X_n})$ and $\mathbf{U = u(X_1,\ldots, X_n)}$**.** If an experiment yields data, $x_1,\ldots, x_n$, then we have observed values $\ell(\mathbf{x_1,\ldots,x_n})$ and $\mathbf{u(x_1,\ldots,x_n)}$.

Definition 3.1

An interval $(\ell(\mathbf{x_1,\ldots,x_n}), \mathbf{u(x_1,\ldots,x_n)})$ is called a **100(1- α)%** ***confidence interval*** for $\boldsymbol{\theta}$ if

$$P[L \le \theta \le U] = 1-\alpha$$

where $0 < \alpha < 1$. The observed values $\ell(\mathbf{x_1,\ldots,x_n})$ and $\mathbf{u(x_1,\ldots,x_n)}$ are called ***lower*** and ***upper confidence limits***, respectively.

The probability level, **1- α ,** is called the ***confidence coefficient***, the percentage 100 **(**1- α) **%** is called the ***confidence level***. When 1- α = 0.95, the interval is called 95% confidence interval for **θ**.

The most common interpretation of the above confidence interval for θ is:

"We are **100 (1- α)%** confident that the single computed interval **(ℓ, u)** contains the parameter θ"

or

"In repeated sampling, from a normally distributed population, **100 (1- α)%** of all intervals of the form (ℓ, u) will in the long run include the population parameter $\boldsymbol{\theta}$".

3.2 Confidence Interval for μ (case (i): σ^2 is known)

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n taken from a normal population with mean μ and variance σ^2 (or from non-normal population but $n \geq 30$), the sample mean and the sample variance are

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{and} \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

where $\bar{X}$ and S^2 are considered as random variables. We know that:

- The mean of the variable $\bar{X}$ is always equal to the mean of the population from which the random samples are chosen and in no way depends on the size of the sample. In symbols we have,

$$\mu_{\bar{X}} = \mu$$

- The variance of $\bar{X}$, however, does depend on the sample size and is equal to the original population variance σ^2 divided by **n**. In symbols we have,

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}.$$

- The square root of the variance of $\bar{X}$, namely $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ is called the ***standard error of the mean***. Consequently, the larger the sample size, the smaller the standard error of $\bar{X}$.
- The sampling distribution of $\bar{X}$ is normal if the population is normal and approximately normal if the conditions of the central limit theorem are met. Therefore, we have

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1).$$

The distribution of $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ is shown in Fig. 3.1. From examination of this figure we see that

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha$$

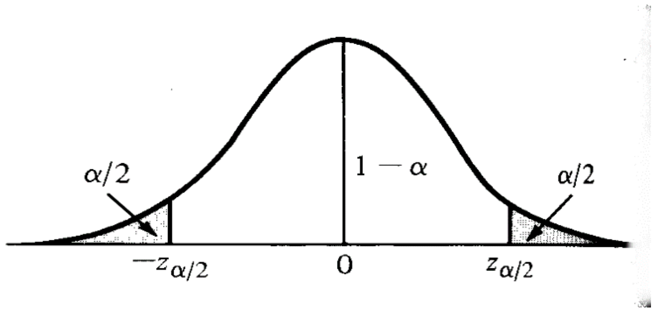

Fig. 3.1 The distribution of Z

Or

$$P\left(-z_{\alpha/2} \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right) = 1-\alpha$$

This can be rearranged as

$$P\left(\bar{X} - z_{\alpha/2}\,\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_{\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha \qquad \textbf{(3.1)}$$

Therefore a **100(1 - α)%** confidence interval for **μ** is

$$\bar{x} - z_{\frac{\alpha}{2}}\,\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}}\,\frac{\sigma}{\sqrt{n}}$$

where $\bar{x}$ is the mean of a sample of size n from a population with known variance σ^2 and $z_{\alpha/2}$ is the value of a standard normal distribution leaving an area of α/2 to the right i.e. $z_{\alpha/2}$ is such that

$$\Phi\,(z_{\alpha/2}) = 1 - \alpha/2$$

For small samples selected from non-normal distributions, we cannot expect our degree of confidence to be accurate. However, for samples of size $n \ge 30$, regardless of the shape of most populations, sampling theory (central limit theorem) guarantees good results.

Clearly, we see that an interval estimate may be constructed, in general, as follows

estimator ± (reliability coefficient) x (standard error)

In particular, when sampling is from a normal distribution with known variance σ2, an interval estimate for the population mean μ may express as

$$\bar{x} \pm z_{\alpha/2}\,.\sigma_{\bar{x}} \qquad \text{i.e.} \qquad \bar{x} \pm z_{\alpha/2}\,.\frac{\sigma}{\sqrt{n}}$$

where $\bar{X}$ is the point estimate of the population mean μ based on a sample of size n from a population with known variance and $z_{\alpha/2}$ is the reliability coefficient (value of a standard normal distribution leaving an area of (1- α /2) to the left) and $\sigma_{\bar{x}}$ is the standard error of $\bar{X}$.

Example 3.1

A sample of size 25 is drawn from a normal population with unknown mean μ and variance 16, have $\bar{X} = 15$. Find a 95% and 99% confidence interval for μ.

Solution

Here we have n = 25, $\sigma^2 = 16$, $\bar{x} = 15$. To find a 95% confidence interval for μ, we put

$$1 - \alpha = 0.95 \;\Rightarrow\; \alpha = 0.05 \;\Rightarrow\; \frac{\alpha}{2} = 0.025$$

Hence

$$\Phi\left(z_{\alpha/2} \right) = 1 - \alpha/2 = 0.975$$

but since

$$\Phi\,(1.96) = 0.975$$

then

$$z_{\alpha/2} = 1.96$$

Thus a 95% confidence interval for μ is give by

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

i.e.

$$15 - 1.96 \times \frac{4}{5} \le \mu \le 15 - 1.96 \times \frac{4}{5}$$

i.e.

$$13.43 \le \mu \le 16.57$$

i.e.

$$P(\,13.43 \le \mu \le 16.57\,) = 0.95$$

If we put

$$1 - \alpha = 0.99 \;\Rightarrow\; \alpha = 0.01 \;\Rightarrow\; \frac{\alpha}{2} = 0.005$$

$$\Phi\left(z_{\alpha/2} \right) = 1 - \alpha/2 = 0.995$$

$$z_{\alpha/2} = 2.58$$

Thus a 99% confidence interval for μ is given by

$$13 \le \mu \le 17$$

Frequently, we are attempting to estimate the mean of a population when the variance is unknown and it is impossible to obtain a sample of size $n \geq 30$. Cost can often be a factor that limits our sample size. As long as our population is approximately bell-shaped, confidence intervals can be computed when σ^2 is unknown and the sample size is small by using the t-distribution in place of the standard normal distribution. The procedure is the same as for large samples except that we use values in place of z values and replace σ by its estimate S.

Maximum Error and Sample size

For large random samples, we find from equation **(3.1)** that the probability is **1-α** that the mean of a large random sample will differ from the population mean by at most $\mathbf{z_{\alpha/2} \cdot \sigma / \sqrt{n}}$ i.e.

$$\mathbf{P(|\bar{X} - \mu| \leq z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha}$$

In other words, if we are going to use $\bar{x}$ as an estimate of μ, the probability is $1-\alpha$ that this estimate will be "Off" either way by at most,

$$\mathbf{E = \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}}$$

The error in estimating the mean μ is $|\bar{X} - \mu|$ and from above we have

$$\mathbf{P(|\bar{X} - \mu| \leq \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}) = 1 - \alpha}$$

The quantity

$$\mathbf{E = \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}}$$

is called the ***maximum error*** occurred in estimating the population mean μ with probability $(1-\alpha)$. The formula for **E** can also be used to determine the sample size that is needed to attain a desired degree of precision. Solving the above equation of **E** for n we obtain

$$\mathbf{n = \left[\frac{z_{\alpha/2} \cdot \sigma}{E}\right]^2}$$

Example 3.2

The life, in hours, of a 150-watt light bulb is known to be approximately normally distributed with standard deviation 25 hours. What sample size should be taken in order to be 95% confident that the error in estimating the mean life is less than

5 hours?

Solution

Since σ=25, α=0.05 $\Rightarrow$ $z_{\alpha/2}$ = $z_{0.025}$ = 1.96 and E = 5, we may find the required sample size from the formula of n as

$$\mathbf{n} = \left[\frac{\mathbf{z}_{\alpha/2}\cdot\sigma}{\mathbf{E}}\right]^2 = \left[\frac{1.96 \times 25}{5}\right]^2 = 96$$

Example 3.3

A medical research worker intends to use the mean of a random sample of size 120 to estimate the mean blood pressure of women in their fifties. If , based on experience, he knows that σ=10.5 mm of mercury, what can he assert with probability 0.99 about the maximum error?

Solution

Since σ=10.5, α=0.01 $\Rightarrow$ $z_{\alpha/2}$ = $z_{0.005}$ = 2.58 and n = 120, we may find the maximum error from the formula of E as

$$\mathbf{E} = \frac{\mathbf{z}_{\alpha/2}\cdot\sigma}{\sqrt{\mathbf{n}}} = \frac{2.58 \times 10.5}{\sqrt{120}} = 2.47 \text{ mm}$$

(B) Confidence Interval for μ (case (ii): when σ^2 is unknown and n < 30)

Let $\mathbf{X_1}$, $\mathbf{X_2}$, ... , $\mathbf{X_n}$ be a random sample of size n (n < 30) taken from a normal population with mean $\boldsymbol{\mu}$ and unknown variance $\boldsymbol{\sigma^2}$, then a **100(1 - α)%** confidence interval for $\boldsymbol{\mu}$ is

$$\bar{\mathbf{x}} - \mathbf{t}_{\alpha/2}\frac{\mathbf{S}}{\sqrt{\mathbf{n}}} \le \mu \le \bar{\mathbf{x}} + \mathbf{t}_{\alpha/2}\frac{\mathbf{S}}{\sqrt{\mathbf{n}}}$$

where $\bar{x}$ and S are the mean and standard deviation, respectively of a sample of size n < 30 from a population that is approximately normal and $t_{\alpha/2}$ is the value of the t-distribution with ν = n-1 degrees of freedom, leaving an area of $\alpha/2$ to the right.

Example 3.4

Nine measurements of reaction time of an individual to certain stimuli were recorded as 0.28, 0.33, 0.30, 0.32, 0.27 0.29, 0.27, 0.31 and 0.33 seconds. Find 95% confidence limits for the actual mean reaction time.

Solution

The sample mean and standard deviation for the given data, using the calculator, are

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 0.30$$

and

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = .000575 \Rightarrow s = \sqrt{.000575} = 0.024$$

Using t-table we find, for $\nu = n-1 = 8$, $t_{\alpha/2} = 2.306$. Hence the 95% confidence limits for μ are

$$0.300 - 2.306\left(\frac{0.024}{3}\right) \le \mu \le 0.30 + 2.306\left(\frac{0.024}{3}\right)$$

which reduces to

$$0.282 \le \mu \le 0.318$$

3.3 Confidence Interval for the Difference of Means of Two Populations (μ_1-μ_2)

Many problems arise where we wish to estimate the difference between two population means by point estimate or by a confidence interval. For example, a farmer may investigate a new variety of wheat by estimating the difference in the average yield of the new variety he has planted in the past.

Case 1:(Use of the Norml Distribution)

Suppose that our two populations have means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, to obtain a point estimate of (μ_1-μ_2), we select two independent random samples, one from each population, of sizes n_1 and n_2, and compute the difference, $(\bar{X}_1 - \bar{X}_2)$, of the sample means. If our independent samples are selected from populations that are approximately normally distributed, or failing this, or if n_1 and n_2 are both greater than 30, we can use the sampling distribution of $(\bar{X}_1 - \bar{X}_2)$ to establish a 100(1 - α)% confidence interval for (μ_1 - μ_2) of the form

$$(\bar{X}_1 - \bar{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\ ,\ (\bar{X}_1 - \bar{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

where $\bar{X}_1$ and $\bar{X}_2$ are means of independent random samples of size $\mathbf{n_1}$ and $\mathbf{n_2}$ from populations with known variances $\boldsymbol{\sigma_1^2}$ and $\boldsymbol{\sigma_2^2}$, respectively, and $\mathbf{Z_{\alpha/2}}$ is defined by

$$\Phi(z_{\alpha/2}) = 1 - \alpha/2$$

If σ_1^2 and σ_2^2 are unknown but $n_1 > 30$ and $n_2 > 30$, we may replace σ_1^2 by S_1^2 and σ_2^2 by S_2^2 without appreciably affecting the confidence interval.

Example 3.5

A standardized mathematics test was given to 50 girls and 75 boys at certain college. The girls made an average grade of 76 with a standard deviation of 6, while the boys made an average grade of 82 with a standard deviation of 8. Find A 95% confidence interval for the difference (μ_1-μ_2), where μ_2 and μ_2 are the mean scores of all boys and all girls respectively, who might take this test.

Solution

Here we have

$$n_1 = 75 \;, \overline{X}_1 = 82 \;,\; S_1 = 8 \;\; (\text{ for boys }),$$
$$n_2 = 50 \;, \overline{X}_2 = 76 \;,\; S_2 = 6 \;\; (\text{ for girls })$$

Since n_1 and n_2 are large, then the confidence interval for the difference (μ_1-μ_2) is given by the above formula.

Let $1 - \alpha = 0.95$, then $\alpha = 0.05$, so

$$\Phi(z_{\alpha/2}) = 1 - \alpha/2 = 0.975 \;\Rightarrow\; z_{\alpha/2} = 1.96$$

Thus

$$\text{Lower limit} = (\overline{X}_1 - \overline{X}_2) - z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = (82 - 76) - 1.96\sqrt{\frac{64}{75} + \frac{36}{50}} = 6 - 2.46 = 3.54$$

$$\text{Upper limit} = (\overline{X}_1 - \overline{X}_2) + z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = 6 + 2.46 = 8.46$$

Therefore (3.54, 8.46) is a 95% confidence interval for the difference of **(μ_1 - μ_2)**.

Case 2: (Use of the t-Distribution)

The above procedure for estimating the difference between two means is applicable if σ_1^2 and σ_2^2 are known or can be estimated from large samples. If the sample sizes are small and σ_1^2 and σ_2^2 are unknown, we can establish confidence intervals for (μ_1 - μ_2) by using the t-distribution, provided that the unknown variances are equal. Thus, if

$$\sigma_1^2 = \sigma_2^2 = \sigma^2,$$

we estimate σ^2 by S_p^2, where S_p^2 obtained by combining or pooling the sample variances according to the formula

$$S_p^2 = \frac{(n_1 - 1)\,S_1^2 + (n_2 - 1)\,S_2^2}{n_1 + n_2 - 2}$$

Therefore, a 100(1-α)% confidence interval for small samples is established in the same way as for large samples except the quantity

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

is replaced by

$$S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

and we use $t_{\alpha/2}$ with $v = n_1 + n_2 – 2$ degrees of freedom in place of $z_{\alpha/2}$. The degrees of freedom for $t_{\alpha/2}$ correspond to the divisor in the formula for S_p^2.

Therefore, a(1-α) 100% confidence interval for the difference ($\mu_1 - \mu_2$) is

$$\left((\overline{X}_1 - \overline{X}_2) - t_{\alpha/2}\,S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\,,\,(\overline{X}_1 - \overline{X}_2) + t_{\alpha/2}\,S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

where $\overline{X}_1$ and $\overline{X}_2$ are the means of small independent samples of sizes n_1 and n_2, respectively, from approximate normal distributions, S_p is the pooled standard deviation, and $t_{\alpha/2}$ is the value of the t-distribution with $v = n_1 + n_2 - 2$ a degrees of freedom, leaving an area of (1-α/2) to the left.

Example 3.6

Seven plants of wheat grown in pots and given a standard fertilizer treatment yield respectively 8.2, 4.4, 4.0, 6.3, 4.7, 11.0 and 9.7 gram dry weight of seed. A further eight plants from the same source are grown in similar conditions, but with a different fertilizer and yield respectively 12.5, 7.3, 10.6, 8.2, 13.0, 6.4, 9.6 and 13.2 g. Construct a 95% confidence interval for the difference of means of the two populations weights of seed production.

Solution

Let μ_1 and μ_2 represent the mean dry weights of seeds produced using two different fertilizers respectively. We wish to find 95% confidence interval for ($\mu_2 – \mu_1$).

Here we have $n_1 = 7$ & $n_2 = 8$. The means and variances may be obtained directly, using any pocket calculator, giving

$$n_1 = 7\ ,\ \bar{x}_1 = 6.9\ ,\ s_1^2 = 7.70 \Rightarrow (n_1 - 1)\, s_1^2 = 46.2$$

$$n_2 = 8\ ,\ \bar{x}_2 = 10.1\ ,\ s_2^2 = 7.06 \Rightarrow (n_2 - 1)\, s_2^2 = 49.42$$

Thus the pooled variance is

$$S_p^2 = \frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2}{n_1 + n_2 - 2} = \frac{46.2 + 49.42}{13} = 7.355$$

$$\therefore\ s_p = 2.712$$

Since $\alpha = 0.05$ & $v = n_1 + n_2 - 2 = 13$, we find from the t-distribution table that

$$t_{\alpha/2} = t_{0.025} = 2.16$$

Hence

$$\text{Lower limit } = (\overline{X}_1 - \overline{X}_2) - t_{\alpha/2}\, S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 3.2 - 3.03 = 0.17$$

$$\text{Upper limit } = (\overline{X}_1 - \overline{X}_2) + t_{\alpha/2}\, S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 3.2 + 3.03 = 6.03$$

Therefore we are 95% confident that the interval (0.17, 6.03) contains the true differences of the mean weights of seeds using the two fertilizers.

3.4 Confidence Interval for the Proportion p

Suppose there is a population of interest, a particular trait is being studied, and each member of the population can be classed as either having or failing to have the trait (e.g. success or failure). Confidence limits are to be found for the parameter **p**, the proportion of the population with the trait. For this purpose, we should draw a random sample from the population of interest, determine the proportion of objects in the sample with the trait, and use this sample proportion as a point estimate of the population proportion **p**. That is,

$$\hat{p} = \frac{X}{n} = \frac{\text{number of objects in sample with trait}}{\text{Sample size}}$$

$$= \text{sample propotion}$$

Note that **X** is a binomial random variable with mean **np** and variance **np(1-p)** .

As the sample size n increases, the sampling distribution of $\hat{p}$ approaches a normal distributed with mean and variance,

$$\mu_{\hat{p}} = p \quad \text{and} \quad \sigma_{\hat{p}}^2 = \frac{p\ (1\ -\ p)}{n}$$

i.e.

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p\,(1-p)}{n}}} \sim N(0\ ,\ 1) \quad \textbf{for n large enough}$$

In order to use the normal approximation, n should be large enough so that

$$np \geq 5 \quad \text{and} \quad n(1-p) \geq 5$$

Although some statisticians recommend both terms np and n(1-p) be greater than 9 or even 10.

Hence a **100(1 - α)%** confidence limits for **p** are given by

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

Since **p**, the parameter we are trying to estimate, is unknown, we must use $\hat{p}$ as an estimate and thus our confidence limits for **p** become

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

Example 3.7

A random sample of 100 patients is selected and treated by a new drug for **AIDS**. After six weeks, 34 of them show signs of improvement. Find a 95% confidence limits for the true proportion of all patients treated by this new drug and show improvement after six weeks.

Solution

Here we have, n =100 and x =34, so the sample proportion is

$$\hat{p} = \frac{34}{100} = 0.34$$

We now use the formula to obtain the confidence interval for the population proportion p. A 95% confidence interval for p is:

$$\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p\,(1-p)}{n}} = 0.34 \pm 1.96 \sqrt{\frac{(0.34)(0.66)}{100}}$$

$$= (0.27, 0.41)$$

<+><+><+><+><+><+><+><+>

EXERCISES

[1] The mean and standard deviation of the maximum leads supported by 60 cables are given by 11.09 tons and 0.73 tons respectively. Find **(i)** 95% and **(ii)** 99% confidence interval for the mean of the maximum loads of all cables produced by the company.

[2] A study of the annual growth of certain cacti showed that 64 of them, selected at random in a desert region, grew on the average 52.80 mm with standard deviation of 4.5 mm. Construct 98% confidence interval for the true average annual growth of the given kind of cactus.

[3] The average amount of money μ for a hospital's accounts receivable must be estimated. The population variance $\boldsymbol{\sigma^2}$ is known to be 100. The sample size needed to be 96% confident that the error of estimating the mean μ will not exceed 2 L.E. is

a. 77 **b.** 96 **c.** 106 **d.** None of the above.

[4] A hospital administrator wishes to estimate the mean weight of babies borne in his hospital. How large a sample of birth records should be taken if he wants a 98% confidence interval that is 0.9 pound wide? Assume that a reasonable estimate of the population standard deviation is 0.85 pounds.

[5] A study is to be conducted to estimate the proportion p of students in secondary schools who smoke regularly. How large a sample is required if we wish to be 96% confident that the error in estimating this proportion is less than 0.05, if

i- it is known that the true proportion lies on the interval 0.30 and 0.10;

ii- it has no idea what the true value might be.

[6] A medical research worker intends to use the mean of a random sample of size 80 to estimate the mean blood pressure of women in their fifties. If , based on experience, he knows that σ=7.5 mm of mercury, what can he assert with probability 0.98 about the maximum error?

[7] The weights of 9 boxes of god food are 10.2, 9.7, 10.3, 10.0, 10.1, 9.8, 9.9, 10.3 and 9.8 ounces. Find a 99% confidence interval for the mean of all such boxes of dog food. Assume an approximate normal distribution.

[8] A sample of 12 measurements of the breaking strengths of cotton threads gave a mean of 7.38 ounces and standard deviations of 1.24 ounces. Find (a) 95% and (b) 99% confidence limits for the actual breaking strength.

[9] Five measurements of the reaction time of an individual to certain stimuli were recorded as 0.28, 0.30, 0.27, 0.33 and 0.31 seconds. Find 99% confidence limits for the actual reaction time.

[10] We wish to estimate the average number of heartbeats per minute for a certain population. The average number of heartbeats per minute for a sample of 49 subjects was found to be 90. If it is reasonable to assume that these 49 patients constitute a random sample and that the population is normally distributed with a standard deviation of 10. find:

a- 90%, **b-** 95% **c-** 99%

confidence interval for μ .

[11] A sample of 150 brand A light bulbs showed a mean lifetime of 1400 hours and standard deviation of 120 hours. A sample of 200 brand B light bulbs showed a mean lifetime of 1200 hours and standard deviation of 80 hours. Find (a) 95% and 99% confidence limits for the difference of the mean lifetime of the populations of brand A and B.

[12] Determinations of saliva pH levels were made in two independent random samples of seventh grade schoolchildren. Sample A children were caries-free while sample B children had a high incidence of carries. The results were as follows:

A: 7.14, 7.11, 7.61, 7.98, 7.21, 7.16, 7.89, 7.24, 7.86, 7.47, 7.82

B: 7.36, 7.04, 7.19, 7.41, 7.10, 7.15, 7.36, 7.57, 7.64, 7.00

Construct a 90% confidence interval for the difference between the population means. Assume that the population variances are equal.

[13] Drug A was prescribed for a random sample of 12 patients complaining of insomnia. An independent random sample of 16 patients with the same complaint received drug B. The numbers of hours of sleep experienced during the second night after treatment began were as follows:

A: 3.5, 5.7, 3.4, 6.9, 17.8, 3.8, 3.0, 6.4, 6.8, 3.6, 6.8, 5.7

B: 4.5, 11.7, 10.8, 4.5, 6.3, 3.8, 6.2, 6.6, 7.1, 6.4, 4.5, 5.1, 3.2

Construct a 95% confidence interval for the difference between the population means. Assume that the population variances are equal.

[14] A national survey was run to estimate the proportion p of individuals 16 years old and under who smoke regularly. Of 1000 individuals interviewed, 200 smoked regularly. Find a 99% confidence interval on p. If you read an article that claimed that this proportion is 0.23, would you be surprised? Explain.

[15] In a simple random sample of 150 of the employees of a large firm, 93 were absent due to sickness three or more days this year. Construct 95 percent confidence intervals for the true proportion of these employees who are absent three or more days yearly due to sickness.

[16] Multiple Choice

1. The critical value $\mathbf{z}_{\alpha/2}$ that corresponds to a degree of confidence 91% is

a. 1.34 **b.** 1.75 **c.** 1.70 **d.** 1.645

2. The margin of error E in estimating the duration of telephone calls directed by a local telephone company with **σ**=3.0 minutes, n=580,97 and confidence level 97% is

a. 0.270 min. **b.** 0.057 min. **c.** 0.011 min. **d.**0.006 min.

3. The minimum sample size you should use to assure that your estimate of $\hat{p}$ will be within the required margin of error E=0.006, around the population **p** and confidence level 95% is:

a. 161 **b.** 38,415 **c.** 82 **d.** 38416

4. A 90% confidence interval estimate for the mean μ with **n=30**, $\bar{\mathbf{x}} = \mathbf{79.1}$ and **s=16.8**, assume that the population has a normal distribution, is:

a. $72.83 < \mu < 85.37$ **b.** $73.89 < \mu < 84.31$
c. $73.92 < \mu < 84.28$ **d.** $70.65 < \mu < 87.5$

5. In a test for acid rain, an SRS of 49 water samples showed a mean pH level of 4.4 with standard deviation of 0.35. A 95% confidence interval estimate for the mean **pH** level is given by

a. 4.4 ± 0.098 **b.** 4.4 ± 0.082 **c.** 4.4 ± 0.98 **d.** None of the above

6. Sixteen measurements of the reaction time of an individual to certain stimuli were recorded and gave mean of 0.296 seconds and standard deviation of 0.1 seconds. A 98% confidence interval estimate for the actual mean reaction time is given by

a. (0.270,0.322) **b.** (0.231,0.361) **c.** (0.238,0.364) **d.** None of the above

7. The width of a 100(1-α)% confidence interval for the mean μ decreases when

a. α increases **b.** n decreases **c.** The S.D. increases **d**. None of the above.

6.. The width of a 100(1-α)% confidence interval for the mean μ increases when

a. α increases **b.** α decreases **c.** n increases **d.** both a & c
e. both b &c

<+><+><+><+><+><+><+><+>