Confidence Intervals

We sample 500 individuals from a large population and measure their score on a test. (SAT) We find $\bar{x} = 461$. We want to know the mean test score, μ, for all 475,000 students in the population; (high school seniors in California). Assume the standard deviation is σ (population standard deviation)

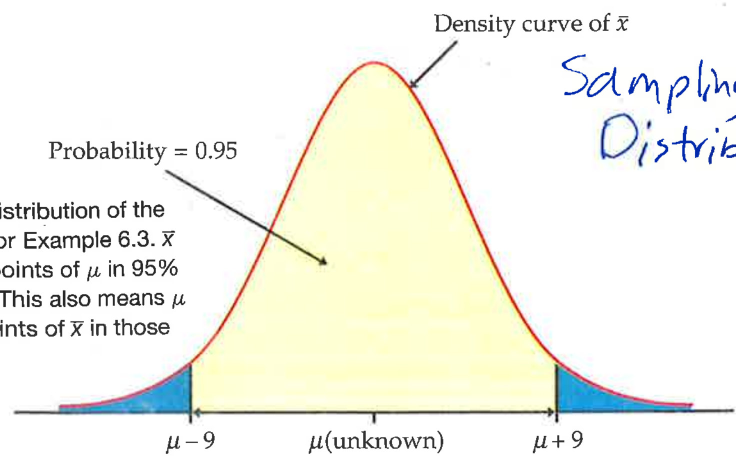
$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{500}}\right)$$

The use of ~~z~~-statistics require that we know σ, (Unrealistic). If you don't know σ use t-statistics. Say $\sigma = 100$, Then

$$\bar{x} \sim N(\mu, 4.5).$$

Recall

95% of the time $\bar{x}$ lies within $2 \times 4.5 = 9$ points of μ.



Sampling Distribution

FIGURE 6.2 Distribution of the sample mean for Example 6.3. $\bar{x}$ lies within ±9 points of μ in 95% of all samples. This also means μ in within ±9 points of $\bar{x}$ in those samples.

We dont know where 461 lies on this graph because we don't know μ.

9 is the margin of error of this study

our estimate is 461 = $\bar{x}$

95% of the time μ is within 9 points of our estimate.

(461-9, 461+9) = (452, 470)

is our confidence interval.

We are 95% confident that the population mean is between 452 and 470.

If we start over with a new random sample of 500 students we will get a new confidence interval.

95% of the time our confidence interval will contain μ,

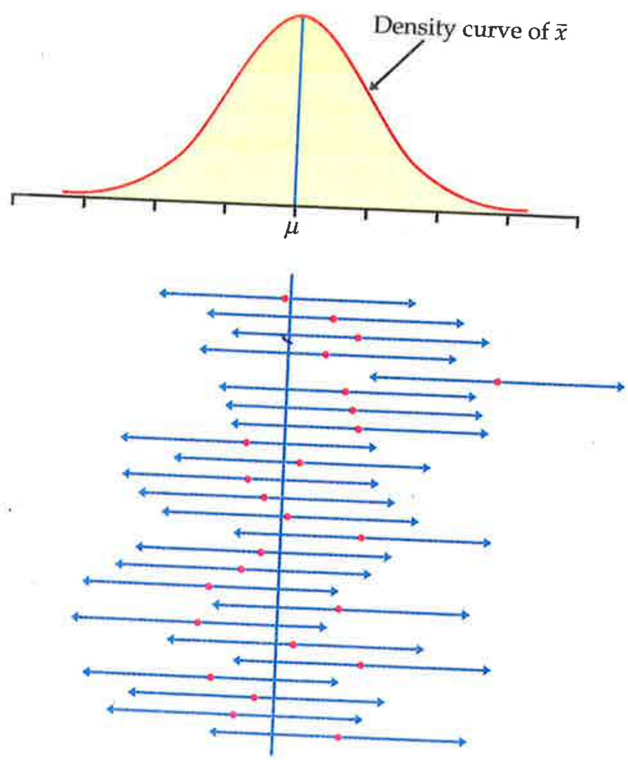


FIGURE 6.3 Twenty-five samples from the same population gave these 95% confidence intervals. In the long run, 95% of all samples give an interval that covers μ. The sampling distribution of $\bar{x}$ is shown at the top.

z^* is the critical value of the test statistic in this case z-score.

CONFIDENCE INTERVAL FOR A POPULATION MEAN

Choose an SRS of size n from a population having unknown mean μ and known standard deviation σ. The **margin of error** for a level C confidence interval for μ is

$$m = z^* \frac{\sigma}{\sqrt{n}}$$

(need to know σ)

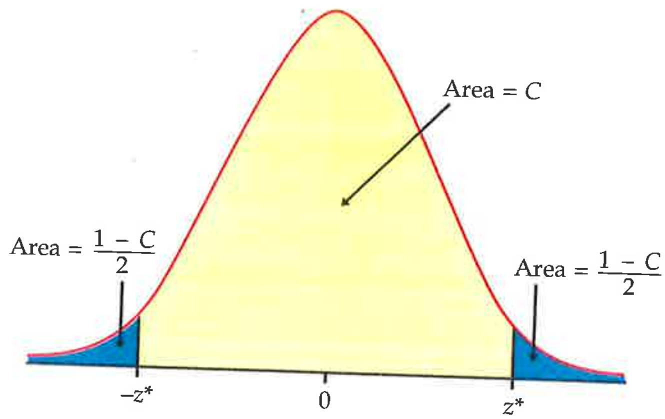
Here z^* is the value on the standard Normal curve with area C between the critical points $-z^*$ and z^*. The level C **confidence interval** for μ is

$$\bar{x} \pm m$$

This interval is exact when the population distribution is Normal and is approximately correct when n is large in other cases.

FIGURE 6.4 To construct a level C confidence interval, we must find the number z^*. This is how they are related. The area between $-z^*$ and z^* under the standard normal curve is C.

z^*	1.645	1.960	2.576
C	90%	95%	99%



$\bar{x}$ lies between $\mu - z^* \frac{\sigma}{\sqrt{n}}$ and $\mu + z^* \frac{\sigma}{\sqrt{n}}$ with probability C so that's the same as saying μ lies between

$$\bar{x} - z^* \sigma/\sqrt{n} \text{ and } \bar{x} + z^* \sigma/\sqrt{n}$$

confidence interval

Width of confidence intervals
(ie margin of error)
Depends on sample size and confidence level

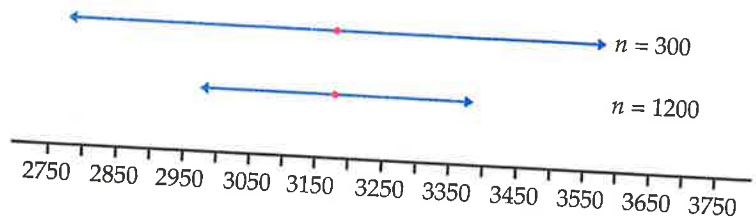


FIGURE 6.5 Confidence intervals for $n = 1200$ and $n = 300$, for Examples 6.4 and 6.5. A sample size 4 times as large results in a confidence interval that is half as wide.

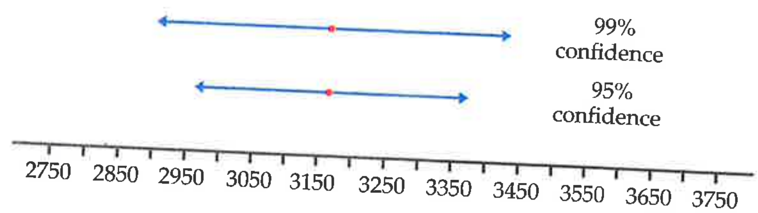


FIGURE 6.6 Confidence intervals for Examples 6.4 and 6.6. The larger the value of C, the wider the interval.

SAMPLE SIZE FOR DESIRED MARGIN OF ERROR

The confidence interval for a population mean will have a specified margin of error m when the sample size is

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