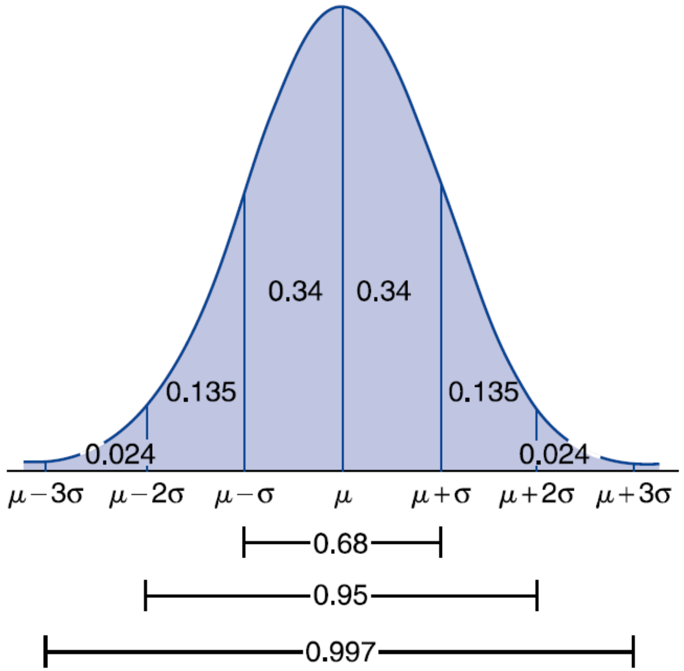


$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

$$\mathrm{SD}(\overline{X}) = \frac{\sigma}{\sqrt{n}}$$

standard error of $\overline{X}$ as an estimator of the mean.

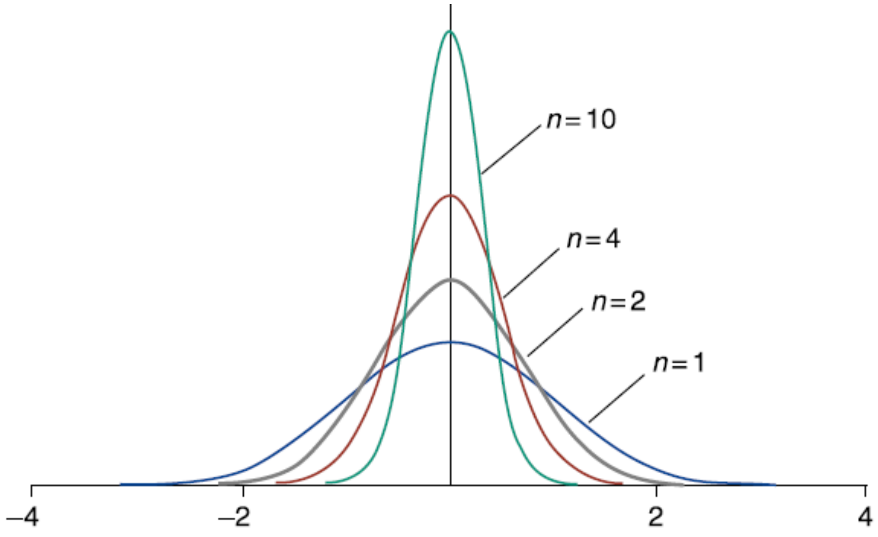
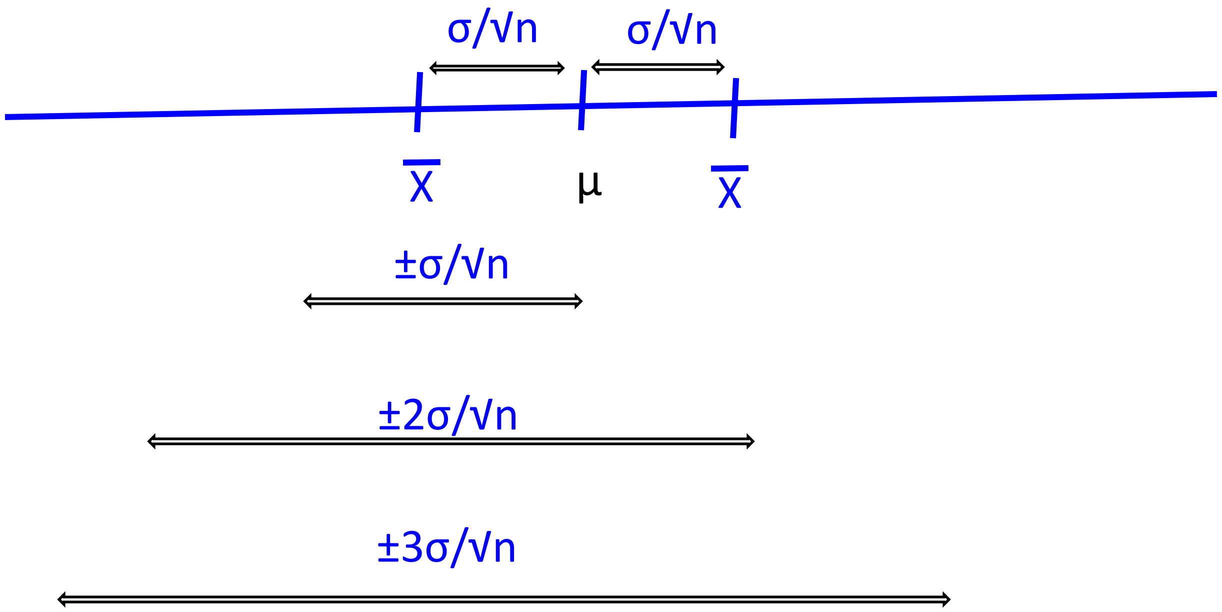


FIGURE 7.1

Densities of sample means from a standard normal population.



Review

$$\bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \quad is\ written\ as\ \overline{x} \pm Z_{\alpha/2}\, s.e\,(\bar{x})$$

100(1-α) % confidence interval estimator of population mean μ

$$\overline{\boldsymbol{x}} \pm \boldsymbol{Z}_{\alpha/2}\frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}$$

The Length of this interval

$$\mathbf{2}\,\boldsymbol{Z}_{\alpha/2}\frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}$$

Will be less than or equal to *b* when the sample size *n* is such that

$$\boldsymbol{n} \geq \left[\mathbf{2}\,\boldsymbol{Z}_{\alpha/2}\ \frac{\boldsymbol{\sigma}}{\boldsymbol{b}}\right]^{\mathbf{2}}$$

Review

A 100(1-α) upper confidence bound for μ is $\overline{x} + Z_{\alpha}\frac{\sigma}{\sqrt{n}}$

A 100(1-α) lower confidence bound for μ is $\overline{x} - Z_{\alpha}\frac{\sigma}{\sqrt{n}}$

95% Confidence interval level is commonly called as margin of error

For a Binomial Distribution

$$\frac{X/n - p}{\sqrt{p(1-p)/n}} = \frac{X - np}{\sqrt{np(1-p)}}$$

$P\{X = i\}$ as $P\{i - 0.5 \le X \le i + 0.5\}$ *Continuity Correction*

Review

Confidence interval estimate of p

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

Length of confidence interval

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

$$\boldsymbol{\leq \frac{Z_{\alpha/2}}{\sqrt{n}}}$$

$$\mathbf{SD}\ (\boldsymbol{\hat{p}}) = \sqrt{\frac{\boldsymbol{\hat{p}(1-\hat{p})}}{\boldsymbol{n}}} \leq \frac{\mathbf{1}}{\mathbf{2}\sqrt{\boldsymbol{n}}}$$

A drug manufacturer claims that a certain drug cures a blood disease, on the average, 80% of the time. To check the claim, government testers used the drug on a sample of 100 individuals and decided to accept the claim if 75 or more were cured.

(a) What is the probability that, the claim will be rejected when the cure probability is, in fact, 0.8?

(b) What is the probability that the claim will be accepted by the government when the cure probability is as low as 0.7?

a) $\mu = (100)(0.8) = 80$ and $\sigma = [(100)(0.8)(0.2)]^{1/2} = 4$, Claim will be rejected with $P(X < 75)$

$z = (74.5 - 80)/4 = -1.38$. Therefore

$$P(\text{Claim is rejected when } p = 0.8) = P(Z < -1.38) = 0.0838.$$

b) If the cure probability if 0.7, then $\mu = (100)(0.7) = 70$ and $\sigma = [(100)(0.7)(0.3)]^{1/2} = 4.583$

Claim will be accepted if $P(X > 74.5)$

$z = (74.5 - 70)/4.583 = 0.8728.$

Therefore $P(\text{Claim is accepted when } p = 0.7) = P(Z > 0.8728) = 1-0.8078 = 19.12\%$

A certain supplier manufactures a type of rubber mat that is sold to automotive companies. The material used to produce the mats must have certain hardness characteristics. Defective mats are occasionally discovered and rejected. The supplier claims that the proportion defective is 0.05. A challenge was made by one of the clients who purchased the mats, so an experiment was conducted in which 400 mats are tested and 17 were found defective.
a. Compute a 95% two-sided confidence interval on the proportion defective.

$n = 400, x = 17$, so $\hat{p} = \frac{17}{400} = 0.0425$.

(a) $z_{0.025} = 1.96$. So,

$$0.0425 \pm (1.96)\sqrt{\frac{(0.0425)(0.9575)}{400}} = 0.0425 \pm 0.0198,$$

which yields $0.0227 < p < 0.0623$.

b. Compute an appropriate 95% one-sided confidence interval on the proportion defective. Interpret both intervals from (a) and (b) and comment on the claim made by the supplier.

(b) $z_{0.05} = 1.645$. So, the upper bound of a left-sided 95% confidence interval is $0.0425 + (1.645)\sqrt{\frac{(0.0425)(0.9575)}{400}} = 0.0591$.

Using both intervals, we do not have evidence to dispute suppliers’ claim.

A multiple-choice quiz has 200 questions, each with 4 possible answers of which only 1 is correct. What is the probability that sheer guesswork yields from 25 to 30 correct answers for the 80 of the 200 problems about which the student has no knowledge?

The probability of a correct answer for each of the 80 questions is *p = 1/4.*

If X **represents the number of correct answers due to guesswork,**

then we have to find P(25 ≤ *X* ≤ *30) with mean of 20 and SD of 3.873.*

$$P(25 \le X \le 30) = P\{[(24.5-20)/3.873] \le Z \le [(30.5-20)/3.873]\} = 0.1196$$

The weights of adobe bricks used for construction are normally distributed with a mean of 3 pounds and a standard deviation of 0.25 pound. Assume that the weights of the bricks are independent and that a random sample of 20 bricks is selected. What is the probability that all the bricks in the sample exceed 2.75 pounds?

Let X denote the weight of a brick, then

$P(X > 2.75) = P(Z > (2.75-3)/0.25 = P(Z > -1) = 0.84134$

Let Y denote the number of bricks in the sample of 20 that exceed 2.75 pounds.

Then by independence, Y has a binomial distribution with $n = 20$ and $p = 0.84134$.

Therefore, $P(Y=20) = (0.84134)^{20} = 0.0315$

The average zinc concentration recovered from a sample of measurements taken in 36 different locations in a river is found to be 2.6 grams per milliliter. Find the 95% and 99% confidence intervals for the mean zinc concentration in the river. Assume that the population standard deviation is 0.3 gram per milliliter. How large a sample is required if we want to be 95% confident that our estimate of μ is off by less than 0.05.

The point estimate of μ is sample mean = 2.6. For 95% confidence interval, the z-value leaving an area of 0.025 to the right, and therefore an area of 0.975 to the left, is $z_{0.025} = 1.96$. Hence, the 95% confidence interval is

$$2.6 - (1.96)\left(\frac{0.3}{\sqrt{36}}\right) < \mu < 2.6 + (1.96)\left(\frac{0.3}{\sqrt{36}}\right),$$ which reduces to $2.50 < \mu < 2.70$

To find a 99% confidence interval, we find the z-value leaving an area of 0.005 to the right and 0.995 to the left, again, $z_{0.005} = 2.575$, and the 99% confidence interval is

$$2.6 - (2.575)\left(\frac{0.3}{\sqrt{36}}\right) < \mu < 2.6 + (2.575)\left(\frac{0.3}{\sqrt{36}}\right),$$ which reduces to $2.47 < \mu < 2.73$

$$n = \left[\frac{(1.96)(0.3)}{0.05}\right]^2 = 138.3.$$

In a psychological testing experiment, 25 subjects are selected randomly and their reaction time, in seconds, to a particular stimulus is measured. Past experience suggests that the variance in reaction times to these types of stimuli is 4 sec^2 and that the distribution of reaction times is approximately normal. The average time for the subjects is 6.2 seconds. Give an upper 95% bound for the mean reaction time.

The upper 95% bound is given by

$$\begin{aligned} \bar{x} + z_\alpha \sigma/\sqrt{n} &= 6.2 + (1.645)\sqrt{4/25} = 6.2 + 0.658 \\ &= 6.858 \text{ seconds.} \end{aligned}$$

Hence, we are 95% confident that the mean reaction time is less than 6.858 seconds.

Scholastic Aptitude Test (SAT) mathematics scores of a random sample of 500 high school seniors in the state of Texas are collected, and the sample mean and standard deviation are found to be 501 and 112, respectively. Find a 99% confidence interval on the mean SAT mathematics score for seniors in the state of Texas.

Since the sample size is large, it is reasonable to use the normal approximation. Using Table, we find $z_{0.005}$ = 2.575. Hence, a 99% confidence interval for μ is

$$501 \pm (2.575)\left(\frac{112}{\sqrt{500}}\right) = 501 \pm 12.9,$$

which yields $488.1 < \mu < 513.9$.

In a random sample of n = 500 families owning television sets in the city of Hamilton, Canada, it is found that x = 340 subscribe to HBO.

a) Find a 95% confidence interval for the actual proportion of families with television sets in this city that subscribe to HBO.

B) How large a sample is required if we want to be 95% confident that our estimate of p is within 0.02 of the true value? Calculate assuming knowledge of sample from part a and also without this knowledge and comment on the outcome.

The point estimate of p is $\hat{p} = 340/500 = 0.68$. $z_{0.025} = 1.96$

the 95% confidence interval for p is

$$0.68 - 1.96\sqrt{\frac{(0.68)(0.32)}{500}} < p < 0.68 + 1.96\sqrt{\frac{(0.68)(0.32)}{500}},$$

which simplifies to $0.6391 < p < 0.7209$.

Let us treat the 500 families as a preliminary sample, providing an estimate is 0.68.

$$n = \frac{(1.96)^2(0.68)(0.32)}{(0.02)^2} = 2089.8 \approx 2090.$$

If estimator is not known then: $$n = \frac{(1.96)^2}{(4)(0.02)^2} = 2401.$$

we see that information concerning *p*, provided by a preliminary sample or from experience, enables us to choose a smaller sample while maintaining our required degree of accuracy

A large diagnostic Laboratory purchases a special kind of sample swabs from a manufacturer. The manufacturer indicates that the defective rate of the swabs is 3%. Suppose that the lab receives 10 shipments in a month, and it randomly tests 20 swabs per shipment. What is the probability that there will be exactly 3 shipments each containing at least one defective swab among the 20 that are selected and tested from the shipment?

A large diagnostic Laboratory purchases a special kind of sample swabs from a manufacturer. The manufacturer indicates that the defective rate of the swabs is 3%. Suppose that the lab receives 10 shipments in a month, and it randomly tests 20 swabs per shipment. What is the probability that there will be exactly 3 shipments each containing at least one defective swab among the 20 that are selected and tested from the shipment?

Sol:

In this case, each shipment can either contain at least one defective item or not.

The number of defective is a binomial distribution with $b(x; 20, 0.03)$. Hence,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - b(0; 20, 0.03)$$

$$= 1 - (0.03)^0 (1 - 0.03)^{20-0} = 0.4562.$$

Assuming independence from shipment to shipment and denoting by Y the number of shipments containing at least one defective item, Y follows another binomial distribution $b(y; 10, 0.4562)$. Therefore,

$$\boxed{P(Y = 3) = \binom{10}{3} 0.4562^3 (1 - 0.4562)^7 = 0.1602.}$$

Suppose X is a random variable with mean μ and variance $\sigma^2 = 1.0$. Suppose also that a random sample of size n is to be taken and sample mean is to be used as an estimate of μ. When the data are taken and the sample mean is measured, we wish it to be within 0.05 unit of the true mean with probability 0.99. What sample size is required?

$$P\ (|\bar{X} - \ \mu|) > 0.05 = 0.99$$

It is known that $P(-2.575 < Z < 2.575) = 0.99$

Therefore

$2.575 = 0.05/(1/n^{1/2})$

$n \sim 2653$

It is desired to know what proportion of students in IIT Bombay are comfortable in communicating in English. What is the minimum size of the sample required if we want to be 95% confident that our estimate of proportion is within 0.02 of the true value?

$$n = \frac{(1.96)^2}{(4)(0.02)^2} = 2401.$$

The concentration of an active ingredient in the output of a chemical reaction is strongly influenced by the catalyst that is used in the reaction. It is felt that when catalyst A is used, the population mean concentration exceeds 65%. The standard deviation is known to be σ = 5%. What is the probability that a sample of outputs from 30 independent experiments gives the average concentration of 64.5% or less.

Now, $n_A = 30$, Sample Mean for A = 64.5% and $\sigma_A = 5\%$,

Hence

$$P(X_A \leq 64.5) = P(Z<[(64.5-65)/(5/\sqrt{30})]) = P(Z<-0.55) = 0.2912$$

Assume that in a digital communication channel, the number of bits received in error can be modelled by a binomial random variable and assume that the probability that a bit is received in error is 1×10^{-5}. If 16 million bits are transmitted, what is the probability that 150 or fewer errors occur?

Assume that in a digital communication channel, the number of bits received in error can be modelled by a binomial random variable and assume that the probability that a bit is received in error is 1 × 10−5. If 16 million bits are transmitted, what is the probability that 150 or fewer errors occur?

$$
\begin{aligned}
P(X \le 150) &= P(X \le 150.5) \\
&= P\left(\frac{X - 160}{\sqrt{160(1 - 10^{-5})}} \le \frac{150.5 - 160}{\sqrt{160(1 - 10^{-5})}} \right) \\
&\approx P(Z \le -0.75) = 0.227
\end{aligned}
$$

A statistician wishes to estimate, with 99 percent confidence, the proportion of people who trust DNA testing. A previous study shows that 91 percent of those who were surveyed trusted DNA testing. The statistician wishes to be accurate within 3 percent of the true proportion. What is the minimum sample size necessary for the statistician to carry out this analysis?

A statistician wishes to estimate, with 99 percent confidence, the proportion of people who trust DNA testing. A previous study shows that 91 percent of those who were surveyed trusted DNA testing. The statistician wishes to be accurate within 3 percent of the true proportion. What is the minimum sample size necessary for the statistician to carry out this analysis?

Solution- $$E = Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

Thus, the minimum samples required is $0.91\ (1-0.91)\ \left(\frac{2.575}{0.03}\right)^2 = 603.38 \sim 604$

How large a sample is needed to ensure that the length of the 80 percent confidence interval estimate of p is less than 0.02?

How large a sample is needed to ensure that the length of the 80 percent confidence interval estimate of p is less than 0.02?

Solution-

To guarantee that the length of 80 percent confidence interval estimator is less than 0.02, we need to choose n so that

$$n > \left(\frac{Z\ 0.1}{0.02}\right)^2$$

since $z_{0.1} = 1.285$, this gives $\quad n > (64.3)^2 = 4134.49$

That is, the sample size needs to be at least 4135 to ensure length of the 80 percent confidence interval will be less than 0.02.

In a population of adults ages 18 to 65, BMI (body mass index) is normally distributed with a mean of 27 and a standard deviation of 5. What BMI marks the bottom 25% of the distribution for this population?

In a population of adults ages 18 to 65, BMI (body mass index) is normally distributed with a mean of 27 and a standard deviation of 5. What BMI marks the bottom 25% of the distribution for this population?

Sol: You want to find the value of X (BMI) where 25% of the population lies below it. In other words, you want to find the 25th percentile of X. First, you need to find the 25th percentile for Z (using the Z-table) and then change the z-value to an x-value by using the z-formula:

$$Z = (X-\mu)/\sigma$$

To find the 25th percentile for Z (or the cutoff point where 25% of the population lies below it), look at the Z-table and find the probability that's closest to 0.25. (***Remember:*** The probabilities for the Z-table are the values *inside* the table. The numbers on the outsides that tell which row/column you're in are actual z-values, not probabilities.) Searching Table, you see that the closest probability to 0.25 is 0.2514.

Next, find what z-score this probability corresponds to. After you've located 0.2514 inside the table, find its corresponding row (–0.6) and column (0.07). Put these numbers together and you get the z-score of –0.67. This is the 25th percentile for Z. In other words, 25% of the z-values lie below –0.67.

To find the corresponding BMI that marks the 25th percentile, use the z-formula and solve for x. You know that $z = -0.67$, $\mu = 27$, and $\sigma = 5$:

$X = 23.65$

So 25% of the population has a BMI lower than 23.65.

In a chemical processing plant, it is important that the yield of a certain type of batch product stay above 80%. If it stays below 80% for an extended period of time, the company loses money. Occasional defective batches are of little concern. But if several batches per day are defective, the plant shuts down and adjustments are made. It is known that the yield is normally distributed with standard deviation 4%. What is the probability of yield of less than 80% when the mean yield is 85%?

In a chemical processing plant, it is important that the yield of a certain type of batch product stay above 80%. If it stays below 80% for an extended period of time, the company loses money. Occasional defective batches are of little concern. But if several batches per day are defective, the plant shuts down and adjustments are made. It is known that the yield is normally distributed with standard deviation 4%. What is the probability of yield of less than 80% when the mean yield is 85%?

(a) $\mu = 85$ and $\sigma = 4$. So, $P(X < 80) = P(Z < -1.25) = 0.1056$.

In a Nernst cell, it was measured that the average time required for the electron to move from Zn anode to Cu cathode was 28 nanoseconds with a standard deviation of 5 nanoseconds. In a given week the experiment was repeated 40 times. What is the probability that the average transport time was greater than 30 nanoseconds?

In a Nernst cell, it was measured that the average time required for the electron to move from Zn anode to Cu cathode was 28 nanoseconds with a standard deviation of 5 nanoseconds. In a given week the experiment was repeated 40 times. What is the probability that the average transport time was greater than 30 nanoseconds?

Solution:

$$P(\bar{X} > 30) = P\left(\frac{\bar{X} - 28}{5/\sqrt{40}} \geq \frac{30.5 - 28}{5/\sqrt{40}}\right) = P(Z \geq 3.16) = 0.0008.$$

The heights of students in IITB are approximately normally distributed with a mean of 175 centimetres and a standard deviation of 7 centimetres. Suppose 25 random samples of size 50 are drawn from this population and the means recorded. Determine the no. of samples means that fall between 172 and 176.

$\mu_{\bar{x}} = \mu = 175$, $\sigma_x = \sigma/\sqrt{n} = 7/\sqrt{50} = 0.989$

$Z_1 = (172-175)/0.989 = -3.03$, $Z_1 = (176-175)/0.989 = 1.01$

$P(172 < \bar{X} < 176) = P(-3.03 < Z < 1.01) = P(-3.03 < Z < 0) + P(0 < Z < 1.01) = 0.9988-0.5+(0.8438-0.5) = 0.8426$

Therefore, the no. of samples means that fall between 172 and 176 = 25(0.8426) = 21.0