Solution Section 3.1 – Distribution of the Sample Mean / Proportion

Exercise

Assume that SAT scores are normally distributed with mean $\mu = 1518$ and standard deviation $\sigma = 325$.

a) If 1 SAT score is randomly selected, find the probability that it is less than 1500.

b) If 100 SAT scores are randomly selected; find the probability that they have a mean less than 1500.

c) If 1 SAT score is randomly selected, find the probability that it is greater than 1600.

d) If 64 SAT scores are randomly selected, find the probability that they have a mean greater than 1600.

e) If 1 SAT score is randomly selected; find the probability that it is between 1550 and 1575.

f) If 25 SAT scores are randomly selected; find the probability that they have a mean between 1550 and 1575.

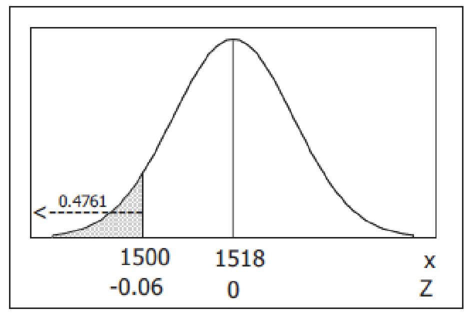
Solution

a) Normal distribution with: $\mu = 1518, \quad \sigma = 325$

$$x = 1500 \quad \rightarrow \quad z = \frac{x-\mu}{\sigma} = \frac{1500-1518}{325} = -0.06$$

$$P(x < 1500) = P(z < -0.06)$$

$$= 0.4761$$



b) Normal distribution, since the original distribution is so

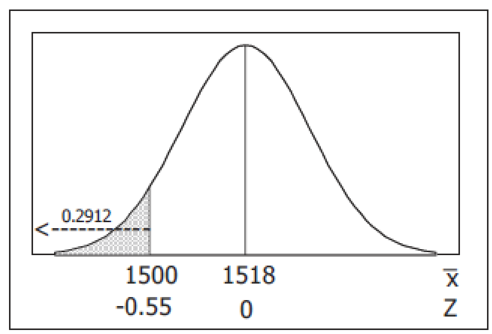
$$\mu_{\overline{x}} = \mu = 1518$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{325}{\sqrt{100}} = 32.5$$

$$\overline{x} = 1500 \quad \rightarrow \quad z = \frac{\overline{x}-\mu}{\sigma} = \frac{1500-1518}{32.5} = -0.55$$

$$P(x < 1500) = P(z < -0.55)$$

$$= 0.2912$$



c) Normal distribution with: $\mu = 1518, \quad \sigma = 325$

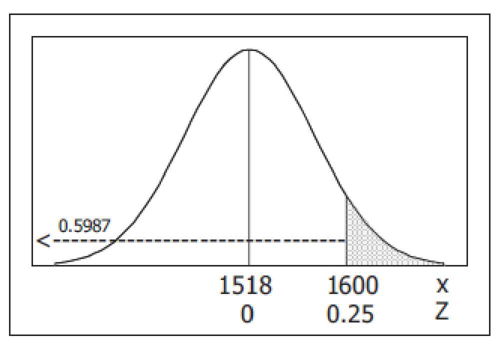
$$x = 1600 \quad \rightarrow \quad z = \frac{x-\mu}{\sigma} = \frac{1600-1518}{325} = 0.25$$

$$P(x > 1600) = P(z > 0.25)$$

$$= 1 - P(z < 0.25)$$

$$= 1 - 0.5987$$

$$= 0.4013$$

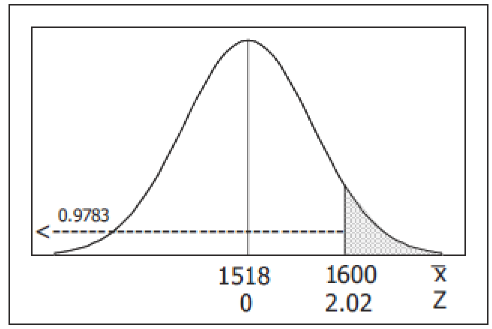


d) Normal distribution, since the original distribution is so

$$\mu_{\overline{x}} = \mu = 1518$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{325}{\sqrt{64}} = 40.625$$

$$\overline{x} = 1600 \quad \rightarrow \quad z = \frac{\overline{x} - \mu}{\sigma} = \frac{1600 - 1518}{40.625} = 2.02$$

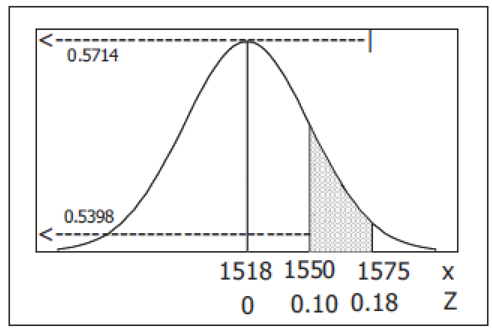


$$P(\overline{x} > 1600) = P(z > 2.02)$$
$$= 1 - P(z < 2.02)$$
$$= 1 - 0.9783$$
$$= 0.0217$$

e) Normal distribution with: $\mu = 1518, \quad \sigma = 325$

$$x = 1550 \quad \rightarrow \quad z = \frac{x - \mu}{\sigma} = \frac{1550 - 1518}{325} = 0.10$$

$$x = 1575 \quad \rightarrow \quad z = \frac{x - \mu}{\sigma} = \frac{1575 - 1518}{325} = 0.18$$



$$P(1550 < x < 1575) = P(0.10 < z < 0.18)$$
$$= P(z < 0.18) - P(z < 0.10)$$
$$= 0.5714 - 0.5398$$
$$= 0.0316$$

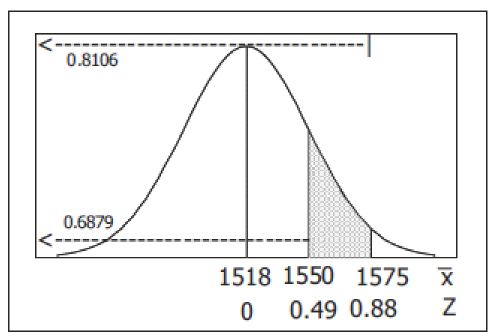
f) Normal distribution, since the original distribution is so

$$\mu_{\overline{x}} = \mu = 1518$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{325}{\sqrt{6425}} = 65$$

$$\overline{x} = 1550 \quad \rightarrow \quad z = \frac{\overline{x} - \mu}{\sigma} = \frac{1550 - 1518}{65} = 0.49$$

$$\overline{x} = 1575 \quad \rightarrow \quad z = \frac{\overline{x} - \mu}{\sigma} = \frac{1575 - 1518}{65} = 0.88$$



$$P(1550 < \overline{x} < 1575) = P(0.49 < z < 0.88)$$
$$= P(z < 0.88) - P(z < 0.49)$$
$$= 0.8106 - 0.6879$$
$$= 0.1227$$

Exercise

Assume that weights of mean are normally distributed with a mean of 172 lb. and a standard deviation of 29 lb.

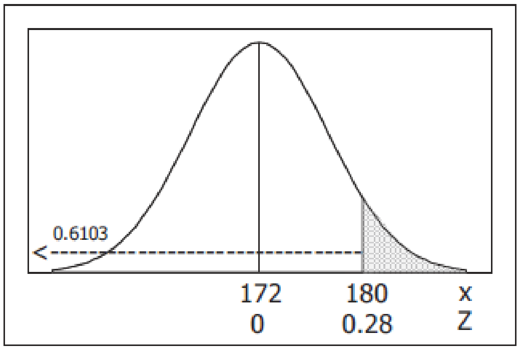
a) Find the probability that if an individual man is randomly selected, his weight will be greater than 180 lb.

b) Find the probability that 20 randomly selected men will have a mean weight that is greater than 180 lb.

c) If 20 mean have a mean weight greater than 180 lb., the total weight exceeds the 3500 lb. safe capacity of a particular water taxi. Based on the proceeding results, is this safety concern? Why or why not?

Solution

a) Normal distribution with: $\mu = 172, \quad \sigma = 29$



$$x = 180 \quad \to \quad z = \frac{x-\mu}{\sigma} = \frac{180-172}{29} \underline{= 0.28}$$

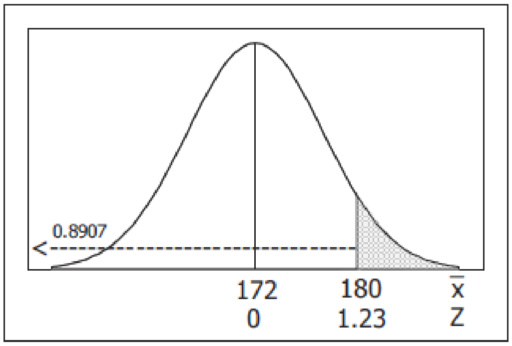
$$\begin{aligned} P(x > 180) &= P(z > 0.28) \\ &= 1 - P(z < 0.28) \\ &= 1 - 0.6103 \\ &\underline{= 0.3897} \end{aligned}$$

b) Normal distribution, since the original distribution is so

$$\mu_{\overline{x}} = \mu = 172$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{29}{\sqrt{20}} = 6.48$$

$$\overline{x} = 180 \quad \to \quad z = \frac{\overline{x}-\mu}{\sigma} = \frac{180-172}{6.48} \underline{= 1.23}$$



$$\begin{aligned} P(\overline{x} > 180) &= P(z > 1.23) \\ &= 1 - P(z < 1.23) \\ &= 1 - 0.8907 \\ &\underline{= 0.1093} \end{aligned}$$

c) Yes. A capacity of 20 is not appropriate when the passengers are all adult men, since a 10.93% probability of overloading is too much of a risk.

Exercise

Membership requires an IQ score above 131.5. Nine candidates take IQ tests, and their summary results indicated that their mean IQ score is 133, (IQ scores are normally distributed with mean of 100 and a standard deviation of 15.)

a) If 1 person is randomly selected from the general population, find the probability of getting someone with an IQ score of at least 133.

b) If 9 people are randomly selected, find the probability that their mean IQ score is at least 133.

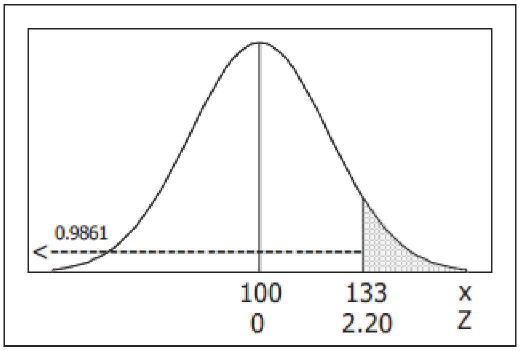
c) Although the summary results are available, the individual IQ test scores have been lost. Can it be concluded that all 9 candidates have IQ scores above 131.5 so that they are all eligible for membership?

Solution

a) Normal distribution with: $\mu = 100, \quad \sigma = 15$

$$x = 133 \quad \to \quad z = \frac{x-\mu}{\sigma} = \frac{133-100}{15} = 2.20$$

$$\begin{aligned} P(x > 133) &= P(z > 2.20) \\ &= 1 - P(z < 2.20) \\ &= 1 - 0.9861 \\ &= 0.0139 \end{aligned}$$



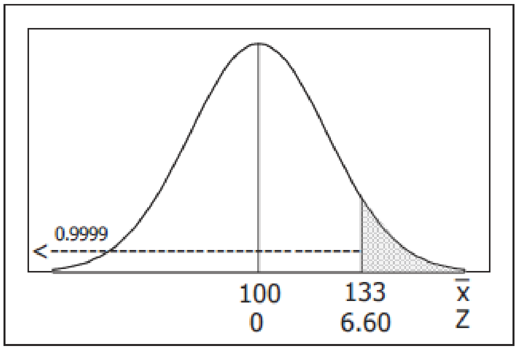
b) Normal distribution, since the original distribution is so

$$\mu_{\bar{x}} = \mu = 100$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{9}} = 5$$

$$\bar{x} = 133 \quad \to \quad z = \frac{\bar{x}-\mu}{\sigma} = \frac{133-100}{5} = 6.60$$

$$\begin{aligned} P(\bar{x} > 133) &= P(z > 6.60) \\ &= 1 - P(z < 6.60) \\ &= 1 - 0.9999 \\ &= 0.0001 \end{aligned}$$



c) No. Even though the mean score is 133, some of the individual scores may be below 131.5.

Exercise

For women aged 18-24, systolic blood pressures (in mm Hg) are normally distributed with a mean of 114.8 and a standard deviation of 13.1. Hypertension is commonly defined as a systolic blood pressure above 140.

a) If a woman between the ages of 18 and 24 is randomly selected, find the probability that her systolic blood pressure is greater than 140.

b) If 4 women in that age bracket are randomly selected, find the probability that their mean systolic blood pressure is greater than 140.

c) Given that part (*b*) involves a sample size that is not larger than 30, why can the central limit theorem be used?

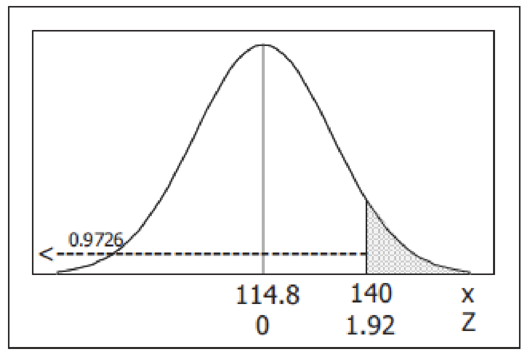
d) If a physician is given a report stating that 4 women have a mean systolic, blood pressure below 140, can she conclude that none of the women have hypertension (with a blood pressure greater than 140)?

Solution

a) Normal distribution with: $\mu = 114.8, \quad \sigma = 13.1$

$$x = 140 \quad \to \quad z = \frac{x-\mu}{\sigma} = \frac{140-114.8}{13.1} = 1.92$$

$$\begin{aligned} P(x>140) &= P(z>1.92) \\ &= 1-P(z<1.92) \\ &= 1-0.9726 \\ &= 0.0274 \end{aligned}$$



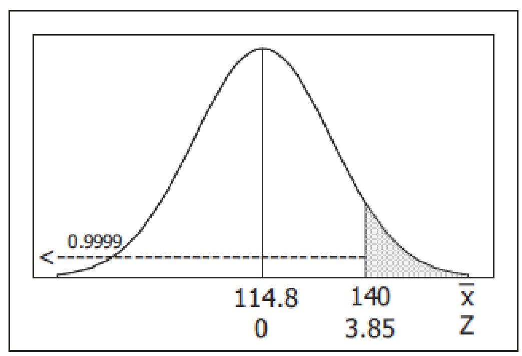
b) Normal distribution, since the original distribution is so

$$\mu_{\overline{x}} = \mu = 114.8$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{13.1}{\sqrt{4}} = 6.55$$

$$\overline{x} = 140 \quad \to \quad z = \frac{\overline{x}-\mu}{\sigma} = \frac{140-100}{6.55} = 3.85$$

$$\begin{aligned} P(\overline{x}>140) &= P(z>3.85) \\ &= 1-P(z<3.85) \\ &= 1-0.9999 \\ &= 0.0001 \end{aligned}$$



c) Since the original distribution is normal, the Central Limit Theorem can be used in part (b) even though the sample size does not exceed 30.

d) The mean can be less than 140 when one or more of the values is greater than 140.

Exercise

Engineers must consider the breadths of male heads when designing motorcycle helmets. Men have head breaths that are normally distributed with a mean of 6.0 in. and a standard deviation of 1.0 in.

a) If one male is randomly selected, find the probability that his head breadth is less than 6.2 in.

b) The Safeguard Helmet Company plans an initial production run of 100 helmets. Find the probability that 100 randomly selected men have a mean head breadth less than 6.2 in.

c) The production manager sees the result from part (b) and reasons that all helmets should be made for men with head breadths less than 6.2 in., because they would fit all but a few men. What is wrong with that reasoning?

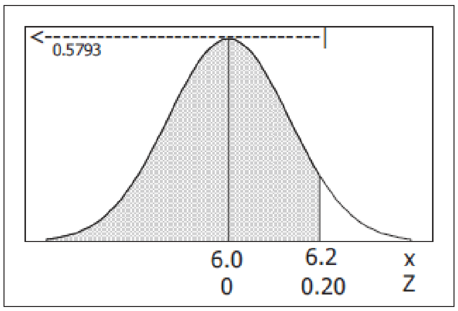
Solution

a) Normal distribution with: $\mu = 6.0, \quad \sigma = 1.0$

$$x = 6.2 \quad \rightarrow \quad z = \frac{x-\mu}{\sigma} = \frac{6.2-6.0}{1.0} \underline{= 0.20}$$

$$P(x < 6.2) = P(z < 0.20)$$

$$\underline{= 0.5793}$$



b) Normal distribution, since the original distribution is so

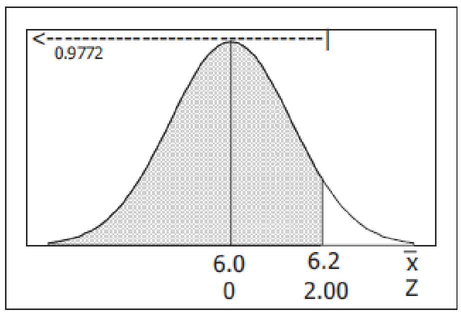
$$\mu_{\overline{x}} = \mu = 6.0$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.0}{\sqrt{100}} = 0.10$$

$$\overline{x} = 6.2 \quad \rightarrow \quad z = \frac{\overline{x}-\mu}{\sigma} = \frac{6.2-6.0}{0.10} \underline{= 2.0}$$

$$P(\overline{x} < 6.2) = P(z < 2.00)$$

$$\underline{= 0.9772}$$



c) Probabilities concerning means do not apply to individuals, It is the information from part (a) that is relevant, since the helmets will be worn by one man at a time – and that indicates that the proportion of men with head breadth greater than 6.2 inches is 1 – 0.5793 = 0.4207 = 42.07%.

Exercise

Currently, quarters have weighs that are normally distributed with a mean 5,670 g and a standard deviation of 0.062 g. A vending machine is configured to accept only those quarters with weights between 5.550 g and 5.790 g.

a) If 280 different quarters are inserted into the vending machine, what is the expected number of rejected quarter?

b) If 280 different quarters are inserted into the vending machine, what is the probability that the mean falls between the limits of 5.550 g and 5.790 g?

c) If you own the vending machine, which result would concern you more? The result from part (*a*) or the result from part (*b*)? Why?

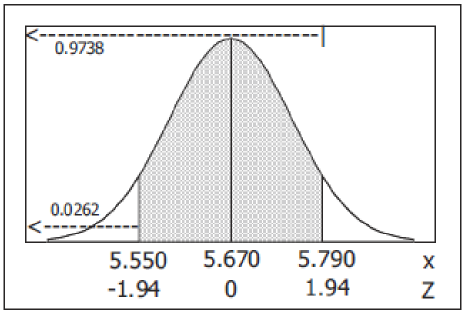
Solution

a) Normal distribution with: $\mu = 5.67, \quad \sigma = 0.062$

$$x = 5.55 \quad \to \quad z = \frac{x-\mu}{\sigma} = \frac{5.55-5.67}{0.062} = -1.94$$

$$x = 5.79 \quad \to \quad z = \frac{x-\mu}{\sigma} = \frac{5.79-5.67}{0.062} = 1.94$$

$$\begin{aligned} P(5.55 < x < 5.79) &= P(-1.94 < z < 1.94) \\ &= P(z < 1.94) - P(z < -1.94) \\ &= 0.9738 - 0.0262 \\ &= 0.9476 \end{aligned}$$



If 0.9476 of the quarters are accepted, then 1 – 0.9476 = 0.0524 of the quarters are rejected.

For 280 quarters, we expect $(0.0524)(280) = 14.7$ of them rejected.

b) Normal distribution, since the original distribution is so

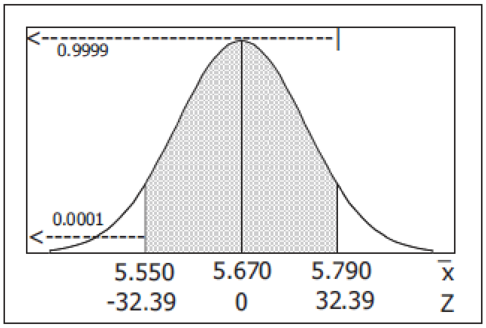
$$\mu_{\overline{x}} = \mu = 5.67$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.062}{\sqrt{280}} = 0.00371$$

$$\overline{x} = 5.55 \quad \to \quad z = \frac{\overline{x}-\mu}{\sigma} = \frac{5.55-5.67}{.00371} = -32.39$$

$$\overline{x} = 5.79 \quad \to \quad z = \frac{\overline{x}-\mu}{\sigma} = \frac{5.79-5.67}{.00371} = 32.39$$

$$\begin{aligned} P(5.55 < \overline{x} < 5.79) &= P(-32.39 < z < 32.39) \\ &= P(z < 32.39) - P(z < -32.39) \\ &= 0.9999 - .0001 \\ &= 0..9998 \end{aligned}$$



c) Probabilities concerning means do not apply to individuals, It is the information from part (a) that is relevant, since the vending machine deals with quarters one at a time.

Exercise

The annual precipitation amounts in a certain mountain range are normally distributed with a mean of 101 inches, and a standard deviation of 12 inches. What is the probability that the mean annual precipitation during 36 randomly picked years will be less than 103.8 inches?

Solution

Given: $x = 103.8; \quad \mu = 101$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = 2$$

$$z = \frac{x - \mu}{\sigma} = \frac{103.8 - 101}{2} = 1.4$$

$$P(x < 103.8) = P(z < 1.4)$$

$$= 0.9192$$

Exercise

The annual precipitation amounts in a certain mountain range are normally distributed with a mean of 72 inches, and a standard deviation of 14 inches. What is the probability that the mean annual precipitation during 49 randomly picked years will be less than 74.8 inches?

Solution

Given: $x = 74.8; \quad \mu = 72$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{14}{\sqrt{49}} = 2$$

$$z = \frac{x - \mu}{\sigma} = \frac{74.8 - 72}{2} = 1.4$$

$$P(x < 74.8) = P(z < 1.4)$$

$$= 0.9192$$

Exercise

The weights of the fish in a certain lake are normally distributed with a mean of 13 lb. and a standard deviation of 6. If 4 fish are randomly selected, what is the probability that the mean weight will be between 10.6 and 16.6 lb.?

Solution

Given: $x_1 = 10.6; \quad x_2 = 16.6; \quad \mu = 13$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{4}} = 3$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{10.6 - 13}{3} = -0.8 \qquad z_2 = \frac{x_2 - \mu}{\sigma} = \frac{16.6 - 13}{3} = 1.2$$

$$P(10.6 < x < 16.6) = P(-0.8 < z < 1.2)$$
$$= P(z = 1.2) - P(z = -0.8)$$
$$= .8849 - .2119$$
$$\underline{= 0.6730}$$

Exercise

For women aged 18-24, systolic blood pressures (in mm Hg) are normally distributed with a mean of 114.8 and a standard deviation of 13.1. If 23 women aged 18-24 are randomly selected, find the probability that their mean systolic blood pressure is between 119 and 122.

Solution

Given: $x_1 = 119$; $x_2 = 122$; $\mu = 114.8$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{114.8}{\sqrt{23}} = 23.937$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{119 - 114.8}{23.937} = 0.18 \qquad z_2 = \frac{x_2 - \mu}{\sigma} = \frac{122 - 114.8}{23.937} = 0.3$$

$$P(119 < x < 122) = P(0.18 < z < 0.3)$$
$$= P(z = 0.3) - P(z = 0.18)$$
$$= 0.6179 - 0.5714$$
$$\underline{= 0.0465}$$

Exercise

A study of the amount of time it takes a mechanic to rebuild the transmission for 2005 Chevy shows that the mean is 8.4 hours and the standard deviation is 1.8 hours. If 40 mechanics are randomly selected, find the probability that their mean rebuild time

a) Exceeds 8.7 hours.

b) Exceeds 8.1 hours.

Solution

a) Given: $x = 8.7$; $\mu = 8.4$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.8}{\sqrt{40}} = 0.285$$

$$z = \frac{x - \mu}{\sigma} = \frac{8.7 - 8.4}{0.285} = 1.05$$

$$P(x > 8.7) = P(z > 1.05)$$
$$= 1 - P(z < 1.05)$$
$$= 1 - 0.8531$$
$$\underline{= 0.1469}$$

b) ***Given***: $x = 8.1; \quad \mu = 8.4$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.8}{\sqrt{40}} = 0.285$$

$$z = \frac{x-\mu}{\sigma} = \frac{8.1-8.4}{0.285} = -1.05$$

$$\begin{aligned} P(x > 8.1) &= P(z > -1.05) \\ &= 1 - P(z < -1.05) \\ &= 1 - 0.1469 \\ &= 0.8531 \end{aligned}$$

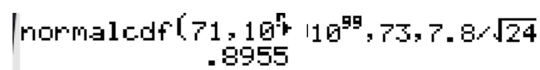
Exercise

A final exam in Math 160 has a mean of 73 with standard deviation 7.8. If 24 students are randomly selected, find the probability that the mean of their test scores is greater than 71.

Solution

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{7.8}{\sqrt{24}} = 0.159$$

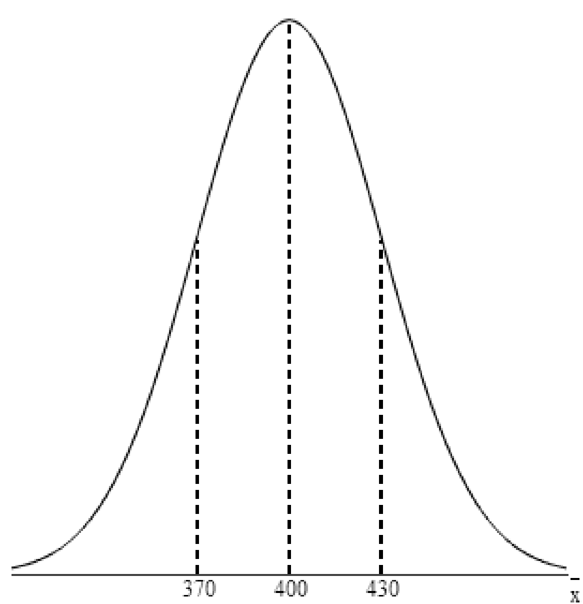
$$z = \frac{x-\mu}{\sigma} = \frac{71-73}{0.159} = -1.25$$



$$\begin{aligned} P(x > 71) &= P(z > -1.25) \\ &= 1 - P(z < -1.25) \\ &= 0.8955 \end{aligned}$$

Exercise

The sampling distribution of the sample mean shown in the graph.



a) What is the value of $\mu_{\overline{x}}$?

b) What is the value of $\sigma_{\overline{x}}$?

c) If the sample size is $n = 9$, what is likely true about the shape of the population?

d) If the sample size is $n = 9$, what is the standard deviation of the population for which the sample was drawn?

Solution

a) $\mu_{\overline{x}} = 400$

b) $\sigma_{\overline{x}} = 430 - 400 = 30$

c) The shape of the population is normal

d) $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} \quad \Rightarrow \quad 30 = \frac{\sigma}{\sqrt{9}}$

$\underline{|\sigma} = 3(30) \ \underline{=90|}$

Exercise

A sample random of size $n = 81$ is obtained from a population with $\mu = 77 \quad and \quad \sigma = 18$.

a) Describe the sampling distribution of $\overline{x}$

b) What is $P(\overline{x} > 79.6)$?

c) What is $P(\overline{x} \le 72.5)$?

d) What is $P(75.1 < \overline{x} < 80.9)$?

Solution

a) The distribution is approximately normal

$\mu_{\overline{x}} = \mu \ \underline{=77|}$

$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{18}{\sqrt{81}} \ \underline{=2|}$

b) $z = \frac{\overline{x} - \mu}{\sigma_{\overline{x}}} = \frac{79.6 - 77}{2} \ \underline{=1.30|}$

$P(\overline{x} > 79.6) = 1 - P(z = 1.30) \ \underline{=0.0968|}$

c) $z = \frac{\overline{x} - \mu}{\sigma_{\overline{x}}} = \frac{72.5 - 77}{2} \ \underline{=-2.25|}$

$P(\overline{x} > 72.5) = P(z = -2.25) \ \underline{=0.0122|}$

d) $z_1 = \frac{\overline{x} - \mu}{\sigma_{\overline{x}}} = \frac{75.1 - 77}{2} \ \underline{=-0.95|} \qquad z_2 = \frac{80.9 - 77}{2} \ \underline{=1.95|}$

$$\begin{aligned} P(75.1 < \overline{x} < 80.9) &= P(z_2 = 1.95) - P(z_1 = -0.95) \\ &= .9744 - .1711 \\ &\underline{=0.8033|} \end{aligned}$$

Exercise

The reading speed of second grade students is approximately normal, with a mean of 90 words per minute (*wpm*) and a standard deviation of 10 *wpm*.

a) What is the probability a randomly selected student will read more than 95 wpm?

b) What is the probability that a random sample of 11 second grade students results in a mean reading rate of more than 95 wpm?

c) What is the probability that a random sample of 22 second grade students results in a mean reading rate of more than 95 wpm?

d) What effect does increasing the sample size have on the probability?

Solution

a) $z = \dfrac{\overline{x} - \mu}{\sigma} = \dfrac{95 - 90}{10} = 0.5$

$P(\overline{x} > 95) = 1 - P(z = .5)\ \underline{= 0.3085}$

b) $z = \dfrac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{95 - 90}{10/\sqrt{11}}\ \underline{= 1.66}$

$P(\overline{x} > 95) = 1 - P(z = 1.66)\ \underline{= 0.0486}$

c) $z = \dfrac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{95 - 90}{10/\sqrt{22}}\ \underline{= 2.35}$

$P(\overline{x} > 95) = 1 - P(z = 2.35)\ \underline{= 0.0095}$

d) Increasing the sample size decreases the probability because $\sigma_{\overline{x}}$ increases as *n* increases.

Solution *Section* 3.2 – Estimating a Population Proportion

Exercise

Find the critical value $z_{\alpha/2}$ that corresponds to a 99% confidence level.

Solution

For 99% confidence, $\alpha = 1 - 0.99 = 0.01 \quad \Rightarrow \quad \frac{\alpha}{2} = \frac{0.01}{2} = 0.005$

For the upper 0.005: $A = 0.995 \quad \Rightarrow \quad z = 2.575$

$z_{\alpha/2} = z_{0.005} = 2.575$

z score	Area
1.645	0.9500
2.575	0.9950

Exercise

Find the critical value $z_{\alpha/2}$ that corresponds to a 99.5% confidence level.

Solution

For 95% confidence, $\alpha = 1 - 0.995 = 0.005 \quad \Rightarrow \quad \frac{\alpha}{2} = \frac{0.005}{2} = 0.0025$

For the upper 0.0025: $A = 0.9975 \quad \Rightarrow \quad z = 2.81$

$z_{\alpha/2} = z_{0.0025} = 2.81$

z	.00	.01	.02
2.8	.9974	.9975	.9976

Exercise

Find the critical value $z_{\alpha/2}$ that corresponds to a 98% confidence level.

Solution

For 95% confidence, $\alpha = 1 - 0.98 = 0.02 \quad \Rightarrow \quad \frac{\alpha}{2} = \frac{0.02}{2} = 0.01$

For the upper 0.01: $A = 0.99 \quad \Rightarrow \quad z = 2.33$

$z_{\alpha/2} = z_{0.01} = 2.33$

Exercise

Find $z_{\alpha/2}$ for $\alpha = 0.10$.

Solution

For $\alpha = 0.1 \quad \Rightarrow \quad \frac{\alpha}{2} = \frac{0.1}{2} = 0.05$

For the upper 0.05: $A = 0.95 \quad \Rightarrow \quad z = 1.645$

$z_{\alpha/2} = z_{0.05} = 1.645$

z score	Area
1.645	0.9500
2.575	0.9950

Exercise

Find $z_{\alpha/2}$ for $\alpha = 0.02$.

Solution

For $\alpha = 0.02 \quad \Rightarrow \quad \frac{\alpha}{2} = \frac{0.02}{2} = 0.01$

For the upper 0.01: $A = 0.99 \quad \Rightarrow \quad z = 2.33$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916

$z_{\alpha/2} = z_{0.01} = 2.33$

Exercise

Express the confidence interval $0.200 < p < 0.500$ in the form $\hat{p} \pm E$

Solution

Let L = the lower confidence limit $\qquad U$ = the upper confidence limit

$\hat{p} = \frac{L+U}{2} = \frac{0.2+0.5}{2} = 0.35$

$E = \frac{U-L}{2} = \frac{0.5-0.2}{2} = 0.15$

$\therefore$ The interval can be expressed as 0.35 ± 0.15

Exercise

Express the confidence interval $0.42 < p < 0.54$ in the form $\hat{p} \pm E$

Solution

$\hat{p} = \frac{L+U}{2} = \frac{0.42+0.54}{2} = 0.48$

$E = \frac{U-L}{2} = \frac{0.54-0.42}{2} = 0.06$

$\therefore$ The interval can be expressed as 0.48 ± 0.06

Exercise

Express the confidence interval 0.222 ± 0.044 in the form $\hat{p} - E < p < \hat{p} + E$

Solution

Given: $\quad \hat{p} = 0.222 \quad$ *and* $\quad E = 0.044$

$L = \hat{p} - E = 0.222 - 0.044 = 0.178 \qquad\qquad L = \hat{p} + E = 0.222 + 0.044 = 0.266$

$\therefore$ The interval can be expressed as $0.178 < p < 0.266$

Exercise

Find the point estimate $\hat{p}$ and the margin of error E of $(0.320,\ 0.420)$

Solution

$$\hat{p} = \frac{L+U}{2} = \frac{0.32+0.42}{2} \underline{=0.370|}$$

$$E = \frac{U-L}{2} = \frac{0.42-0.32}{2} \underline{=0.050|}$$

Exercise

Find the margin of error E of $0.542 < p < 0.576$

Solution

$$E = \frac{U-L}{2} = \frac{0.576-0.542}{2} \underline{=0.017|}$$

Exercise

Find the point estimate $\hat{p}$ of $0.824 < p < 0.868$

Solution

$$\hat{p} = \frac{L+U}{2} = \frac{0.824+0.868}{2} \underline{=0.846|}$$

Exercise

Find the point estimate $\hat{p}$ and the margin of error E of $0.772 < p < 0.776$

Solution

$$\hat{p} = \frac{L+U}{2} = \frac{0.772+0.776}{2} \underline{=0.774|}$$

$$E = \frac{U-L}{2} = \frac{0.776-0.772}{2} \underline{=0.002|}$$

Exercise

Find the point estimate $\hat{p}$ and the margin of error E of $0.433 < p < 0.527$

Solution

$$\hat{p} = \frac{L+U}{2} = \frac{0.433+0.527}{2} \underline{=0.480|}$$

$$E = \frac{U-L}{2} = \frac{0.527-0.433}{2} \underline{=0.047|}$$

Exercise

Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given $n = 1000, \quad x = 400, \quad 95\%$ *confidence*

Solution

$$\frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \quad \Rightarrow \quad A = 1 - 0.025 = 0.975$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

$$A = 0.95 \quad \Rightarrow \quad z_{\alpha/2} = z_{0.025} = 1.96$$

$$\hat{p} = \frac{x}{n} = \frac{400}{1000} \underline{= 0.40} \qquad \hat{q} = 1 - \hat{p} = 1 - 0.4 \underline{= 0.6}$$

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96\sqrt{\frac{(0.4)(0.6)}{1000}} \underline{= 0.0304}$$

Exercise

Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given $n = 500, \quad x = 220, \quad 99\%$ *confidence*

Solution

$$\frac{\alpha}{2} = \frac{1-0.99}{2} = 0.005 \quad \Rightarrow \quad A = 1 - 0.005 = 0.995$$

$$A = 0.995 \quad \Rightarrow \quad z_{\alpha/2} = z_{0.005} = 2.575$$

z score	Area
1.645	0.9500
2.575	0.9950

$$\hat{p} = \frac{x}{n} = \frac{220}{500} \underline{= 0.44} \qquad \hat{q} = 1 - \hat{p} = 1 - 0.44 \underline{= 0.56}$$

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 2.575\sqrt{\frac{(0.44)(0.56)}{500}} \underline{= 0.0572}$$

Exercise

Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given $n = 390, \quad x = 130, \quad 90\%$ *confidence*

Solution

z score	Area
1.645	0.9500
2.575	0.9950

$$\frac{\alpha}{2} = \frac{1-0.90}{2} = 0.05 \quad \Rightarrow \quad A = 1 - 0.05 = 0.95$$

$$A = 0.95 \quad \Rightarrow \quad z_{\alpha/2} = z_{0.05} = 1.645$$

$$\hat{p} = \frac{x}{n} = \frac{130}{390} \underline{= 0.33} \qquad \hat{q} = 1 - 0.33 \underline{= 0.67}$$

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.645\sqrt{\frac{(0.33)(0.67)}{390}} \underline{= 0.0392}$$

Exercise

Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given 98% confidence; the sample size is 1230, of which 40% are successes.

Solution

$$\frac{\alpha}{2}=\frac{1-0.98}{2}=0.01 \quad \Rightarrow \quad A=1-0.01=0.99$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916

$$A=0.99 \quad \Rightarrow \quad z_{\alpha/2}=z_{0.01}=2.33$$

$$\hat{p}=\frac{x}{n}=\frac{492}{1230}\ \underline{=0.4}$$

$$\hat{q}=1-\hat{p}=1-0.4\ \underline{=0.6}$$

$$E=z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}=2.33\sqrt{\frac{(0.4)(0.6)}{1230}}\ \underline{=0.0325}$$

Exercise

Construct the confidence interval estimate of the population proportion p that corresponds to the given $n=200, \quad x=40, \quad 95\%$ *confidence*

Solution

$$\frac{\alpha}{2}=\frac{1-0.95}{2}=0.025 \quad \Rightarrow A=1-0.025=0.975 \quad \to z_{\alpha/2}=z_{0.025}=1.96$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

$$\hat{p}=\frac{x}{n}=\frac{40}{200}\ \underline{=0.2}$$

$$\hat{q}=1-\hat{p}=1-0.2\ \underline{=0.8}$$

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

$$0.20\pm 1.96\sqrt{\frac{(0.2)(0.8)}{200}}$$

$$0.20\pm 0.0554$$

$$0.20-0.0554<p<0.20+0.0554$$

$$\underline{0.145<p<0.255}$$

Exercise

Construct the confidence interval estimate of the population proportion p that corresponds to the given $n=1236, \quad x=109, \quad 99\%$ *confidence*

Solution

$$\frac{\alpha}{2}=\frac{1-0.99}{2}=0.005 \quad \Rightarrow A=1-0.005=0.995 \quad \rightarrow z_{\alpha/2}=z_{0.005}=2.575$$

z score	Area
1.645	0.9500
2.575	0.9950

$$\hat{p}=\frac{x}{n}=\frac{109}{1236}=0.0882$$

$$\hat{q}=1-\hat{p}=1-0.0882=0.9118$$

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

$$0.0882 \pm 2.575\sqrt{\frac{(0.0882)(0.9118)}{1236}}$$

$$0.0882 \pm 0.0207$$

$$0.0882-0.0207<p<0.0882+0.0207$$

$$0.0674<p<0.109$$

Exercise

Construct the confidence interval estimate of the population proportion p that corresponds to the given $n=5200, \quad x=4821, \quad 99\%$ *confidence*

Solution

$$\frac{\alpha}{2}=\frac{1-0.99}{2}=0.005 \quad \Rightarrow A=1-0.005=0.995 \quad \rightarrow z_{\alpha/2}=z_{0.005}=2.575$$

z score	Area
1.645	0.9500
2.575	0.9950

$$\hat{p}=\frac{x}{n}=\frac{4821}{5200}=0.9271$$

$$\hat{q}=1-\hat{p}=1-0.9271=0.0729$$

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

$$0.9271 \pm 2.575\sqrt{\frac{(0.9271)(0.0729)}{5200}}$$

$$0.9271 \pm 0.0093$$

$$0.9271-0.00937<p<0.9271+0.0093$$

$$0.918<p<0.936$$

Exercise

Find the minimum sample size requires to estimate a population proportion or percentage:
Margin of error: 0.045; confidence level: 95%: $\hat{p}$ *and* $\hat{q}$ *unknown*

Solution

$$\frac{\alpha}{2}=\frac{1-0.95}{2}=0.025 \quad \Rightarrow A=1-0.025=0.975 \quad \rightarrow z_{\alpha/2}=z_{0.025}=1.96$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

$E=0.045; \quad \hat{p}$ *unknown* *use* $\hat{p}=0.5$

$$n=\frac{\left(z_{\alpha/2}\right)^2 \hat{p}\hat{q}}{E^2}$$

$$=\frac{(1.96)^2(0.5)(.5)}{(.045)^2}$$

$$\approx 475$$

Exercise

Find the minimum sample size requires to estimate a population proportion or percentage:
Margin of error: 2% points; confidence level: 99%: from prior study, $\hat{p}$ is estimate by the decimal equivalent of 14%

Solution

$$\frac{\alpha}{2}=\frac{1-0.99}{2}=0.005 \quad \Rightarrow \quad A=1-0.005=0.995 \quad \rightarrow z_{\alpha/2}=z_{0.005}=2.575$$

$E=0.02; \quad \hat{p}\approx 0.14$

z score	Area
1.645	0.9500
2.575	0.9950

$$n=\frac{\left(z_{\alpha/2}\right)^2 \hat{p}\hat{q}}{E^2}$$

$$=\frac{(2.575)^2(0.14)(.86)}{(.02)^2}$$

$$\approx 1996$$

Exercise

The Genetics and IVF Institute conducted a clinical trial of the XSORT method designed to increase the probability of conceiving a girl. As of this writing, 574 babies were born to parents using the XSORT method, and 525 of them were girls.

a) What is the best point estimate of the population proportion of girls born to parents using the XSORT method?

b) Use the sample data to construct a 95% confidence interval estimate of the percentage of girls born to parents using the XSORT method.

c) Based on the results, does the XSORT method appear to be effective? Why or why not?

Solution

Let x = the number of girls born using the method

a) $\hat{p} = \frac{x}{n} = \frac{525}{574} \underline{\approx 0.9146}$

b) $\frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \quad \Rightarrow A = 0.975 \quad \rightarrow z_{\alpha/2} = z_{0.025} = 1.96$

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

$$0.9146 \pm 1.96\sqrt{\frac{(0.9146)(.0854)}{5200}}$$

$$0.9146 \pm 0.0229$$

$$0.9146 - 0.0229 < p < 0.9146 + 0.0229$$

$$\underline{0.892 < p < 0.937}$$

c) Yes. Since 0.5 is not within the confidence interval, and below the interval, we can be 95% certain that the method is effective.

Exercise

An important issue facing Americans is the large number of medical malpractice lawsuits and the expenses that they generate. In a study of 1228 randomly selected medical malpractice lawsuits, it is found that 856 of them were later dropped or dismissed.

a) What is the best point estimate of the proportion of medical malpractice lawsuits that are dropped or dismissed?

b) Construct a 99% confidence interval estimate of the proportion of medical malpractice lawsuits that are dropped or dismissed.

c) Does it appear that the majority of such suits are dropped or dismissed?

Solution

Let x = the number of suits dropped or dismissed

a) $\hat{p} = \frac{x}{n} = \frac{856}{1228} \underline{= 0.6971}$

b) $\frac{\alpha}{2}=\frac{1-0.99}{2}=0.005 \quad \Rightarrow A=0.995 \quad \rightarrow z_{\alpha/2}=z_{0.005}=2.575$

z score	Area
1.645	0.9500
2.575	0.9950

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

$$0.6971 \pm 2.575\sqrt{\frac{(0.6971)(0.3029)}{1228}}$$

$$0.6971 \pm 0.0338$$

$$0.6971-0.0338 < p < 0.6971+0.0338$$

$$\boxed{0.663 < p < 0.731}$$

c) Yes. Since 0.5 is not within the confidence interval, and below the interval, we can be 99% certain that more than half the suits are dropped or dismissed.

Exercise

A study of 420,095 Danish cell phone users found that 135 of them developed cancer was found to be 0.0340% for those not using cell phones.

a) Use the sample data to construct a 95% confidence interval estimate of the percentage of cell phone users who develop cancer of the brain or nervous system.

b) Do cell phone users appear to have a rate of cancer of the brain or nervous system that is different from the rate of such cancer among those not using cells phones? Why or why not?

Solution

Let *x* = the number that develop those types of cancer.

a) $\hat{p}=\frac{x}{n}=\frac{135}{420,095}$ $\boxed{0.0003214}$

b) $\frac{\alpha}{2}=\frac{1-0.95}{2}=0.025 \quad \Rightarrow A=0.975 \quad \rightarrow z_{\alpha/2}=z_{0.025}=1.96$

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

$$0.0003214 \pm 1.96\sqrt{\frac{(0.0003214)(0.9996786)}{420095}}$$

$$0.0003214 \pm 0.0000542$$

$$0.0003214-0.0000542 < p < 0.0003214+0.0000542$$

$$\boxed{0.0267\% < p < 0.0376\%}$$

c) No. Since 0.034% is within the confidence interval, it is a reasonable possibility for the true population value. The results do not provide evidence that cell phone users have a different cancer rate than the general population.

Exercise

In an Account survey of 150 senior executives, 47% said that the most common job interview mistake is to have little or no knowledge of the company. Construct a 99% confidence interval estimate of the proportion of all senior executives who have that same opinion. Is it possible that exactly half of all senior executives believe that the most common job interview mistake is to have little or no knowledge of the company? Why or why not?

Solution

Let x = the number who display little or no knowledge of the company.

$$\frac{\alpha}{2} = \frac{1-0.99}{2} = 0.005 \quad \Rightarrow A = 0.995 \quad \to z_{\alpha/2} = z_{0.005} = 2.575$$

$$\hat{p} = \frac{x}{n} = \frac{x}{150} = 0.47 \qquad \Rightarrow x \approx 71$$

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

$$0.47 \pm 2.575\sqrt{\frac{(.47)(.53)}{150}}$$

$$0.47 \pm 0.1049$$

$$0.47 - 0.1049 < p < 0.47 + 0.1049$$

$$0.365 < p < 0.575$$

Yes. Since 0.50 is within the confidence interval, it is a likely value for the true population proportion.

Solution *Section* 3.3 – Estimating a Population Mean

Exercise

A design engineer of the Ford Motor Company must estimate the mean leg length of all adults. She obtains a list of the 1275 employees at her facility; then obtains a simple random sample of 50 employees. If she uses this sample to construct a 95% confidence interval to estimate the mean leg length for the population of all adults, will her estimate be good? Why or why not?

Solution

No. The list of the employees at her facility from which she obtained her simple random sample is itself a convenience sample. Those employees are likely not representative of the population age, gender, ethnicity, or other factors that may affect leg length.

Exercise

Find the critical value $z_{\alpha/2}$ that corresponds to a 90% confidence level.

Solution

For 90% confidence, $\frac{\alpha}{2} = \frac{1-0.90}{2} = 0.05 \quad \Rightarrow A = 0.95$

z score	Area
1.645	0.9500
2.575	0.9950

$\to z_{\alpha/2} = z_{0.05} \ \underline{= 1.645}$

Exercise

Find the critical value $z_{\alpha/2}$ that corresponds to a 98% confidence level.

Solution

For 98% confidence, $\frac{\alpha}{2} = \frac{1-0.98}{2} = 0.01 \quad \Rightarrow A = 0.99$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916

$\to z_{\alpha/2} = z_{0.01} \ \underline{= 2.33}$

Exercise

Find $z_{\alpha/2}$ for $\alpha = 0.20$

Solution

For $\alpha = 0.20 \quad \to \frac{\alpha}{2} = \frac{0.2}{2} = 0.1;$ For upper 0.10, $A = 0.90$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015

$\to z_{\alpha/2} = z_{0.1} \ \underline{= 1.28}$

Exercise

Find $z_{\alpha/2}$ for $\alpha = 0.07$

Solution

For $\alpha = 0.07 \quad \to \frac{\alpha}{2} = \frac{0.07}{2} = 0.035$

For upper 0.035, $A = 1 - 0.035 = 0.965$

z	.00	.01	.02	.03	.04
1.8	.9641	.9649	.9656	.9664	.9671

$\to z_{\alpha/2} = z_{0.035} \underline{= 1.81}$

Exercise

How many adults must be randomly selected to estimate the mean FICO (credit rating) score of working adults in U.S.? We want 95% confidence that the sample mean is within 3 points of the population mean, and the population standard deviation is 68.

Solution

$\frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \quad \Rightarrow A = 0.975$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

$z_{\alpha/2} = z_{0.025} = 1.96$

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

$$= \left(\frac{(1.96)(68)}{3}\right)^2$$

$\underline{\approx 1974}$

Exercise

A simple random sample of 40 salaries of NCAA football coaches has a mean of \$415,953. Assume that σ = \$463,364.

a) Find the best estimate of the mean salary of all NCAA football coaches.
b) Construct a 95% confidence interval estimate of the mean salary of an NCAA football coach.
c) Does the confidence interval contain the actual population mean of \$474,477?

Solution

a) $\overline{x} = \$415{,}953$

b) $\frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \quad \Rightarrow A = 0.975 \quad \to z_{\alpha/2} = z_{0.025} = 1.96$

$$\overline{x} \pm \frac{z_{\alpha/2} \cdot \sigma}{n}$$

$$415953 \pm \frac{(1.96)(463{,}364)}{\sqrt{40}}$$

$$415953 \pm 143{,}598$$

$$415953 - 143{,}598 < \mu < 415953 + 143{,}598$$

$$\boxed{\$272{,}355 < \mu < \$559{,}551}$$

c) Yes. In this case the confidence interval includes the true population mean.

Exercise

A simple random sample of 50 adults (including males and females) is obtained, and each person's red blood cell count (in cells per microliter) is measured. The sample mean is 4.63. The population standard deviation for red blood cell counts is 0.54.

a) Find the best point estimate of the mean red blood cell count of adults.
b) Construct a 99% confidence interval estimate of the mean red blood cell count of adults.
c) The normal range of red blood cell counts for adults is 4.7 to 6.1 for males and 4.3 to 5.4 for females. What does the confidence interval suggest about these normal ranges?

Solution

a) $\overline{x} = 4.63$ *cells / microliter*

b) $\frac{\alpha}{2} = \frac{1-0.99}{2} = 0.005 \quad \Rightarrow A = 0.995 \quad \to z_{\alpha/2} = z_{0.005} = 2.575$

$$\overline{x} \pm \frac{z_{\alpha/2} \cdot \sigma}{n}$$

$$4.63 \pm \frac{(2.575)(0.54)}{\sqrt{50}}$$

$$4.63 \pm 0.20$$

$$4.63 - 0.20 < \mu < 4.63 + 0.20$$

$$\boxed{4.43 < \mu < 4.83} \ \textit{cells / microliter}$$

c) The intervals are not comparable, since the 2 given in part (*c*) are normal ranges for individual counts and the one calculated in part (*b*) is a confidence interval for mean counts. One would expect the confidence interval for mean counts to be well within the normal ranges for individual counts. The fact that the point estimate and the lower confidence interval limit for the mean are so close to the lower limit of the normal ranges for individuals suggests that the sample may consist of persons with lower red blood cell counts.

Exercise

A simple random sample of 125 SAT scores has a mean of 1522. Assume that SAT scores have a standard deviation of 333.

a) Construct a 95% confidence interval estimate of the mean SAT score.
b) Construct a 99% confidence interval estimate of the mean SAT score.
c) Which of the preceding confidence intervals is wider? Why?

Solution

a) $\frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \quad \Rightarrow A = 0.975 \quad \to z_{\alpha/2} = z_{0.025} = 1.96$

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

$$1522 \pm \frac{(1.96)(333)}{\sqrt{125}}$$

$$1522 \pm 58$$

$$1522 - 58 < \mu < 1522 + 58$$

$$\boxed{1464 < \mu < 1580}$$

b) $\frac{\alpha}{2} = \frac{1-0.99}{2} = 0.005 \quad \Rightarrow A = 0.995 \quad \to z_{\alpha/2} = z_{0.005} = 2.575$

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

$$1522 \pm \frac{(2.575)(333)}{\sqrt{125}}$$

$$1522 \pm 77$$

$$1522 - 77 < \mu < 1522 + 77$$

$$\boxed{1445 < \mu < 1599}$$

c) The 99% confidence interval in part (b) is wider than the 95% confidence interval in part (a). For an interval to have more confidence associated with it, it must be wider to allow for more possibilities.

Exercise

When 14 different second-year medical students measured the blood pressure of the same person, they obtained the results listed below. Assuming that the population standard deviation is known to be 10 mmHg, construct a 95% confidence interval estimate of the population mean. Ideally, what should the confidence interval be in this situation?

138 130 135 140 120 125 120 130 130 144 143 140 130 150

Solution

$n = 14 \quad \sum x = 1875 \quad \bar{x} = \frac{1875}{14} = 133.93$

$\frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \quad \Rightarrow A = 0.975$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

$z_{\alpha/2} = z_{0.025} = 1.96$

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

$133.93 \pm \frac{1.96(10)}{\sqrt{14}}$

133.93 ± 5.24

$133.93 - 5.24 < \mu < 133.93 + 5.24$

$128.7 < \mu < 139.2$

There is sense in which all the measurements should be the same – and in that case there would be no need for a confidence interval. It is unclear what the given $\sigma = 10$ represents in this situation. Is it the true standard deviation in the values of all people in the population (in which case it would not be appropriate in this context where only a single person is involved)?

Is it the true standard deviation in readings from evaluator to evaluator (when they are supposedly evaluating the same thing)? Using the methods of this section and assuming $\sigma = 10$, the confidence interval would be $128.7 < \mu < 139.2$ as given above even if all the readings were the same.

Exercise

Do the given conditions justify using the margin of error $E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$ when finding a confidence interval estimate of the population mean μ?

a) The sample size is $n = 4$, $\sigma = 12.5$, and the original population is normally distributed

b) The sample size is $n = 5$ and σ is not known

Solution

a) Yes

b) No, σ is not given

Exercise

Use the confidence level and sample data to find the margin of error E.

a) Replacement times for washing machines: 90% confidence; $n = 37, \quad \overline{x} = 10.4\ yrs, \quad \sigma = 2.2\ yrs$

b) College students' annual earnings: 99% confidence; $n = 76, \quad \overline{x} = \$4196, \quad \sigma = \$848$

Solution

a) $\frac{\alpha}{2} = \frac{1-0.90}{2} = 0.05 \quad \Rightarrow A = 0.95$

$$z_{\alpha/2} = z_{0.05} = 1.645$$

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$= 1.645\left(\frac{2.2}{\sqrt{37}}\right)$$

$$\approx 0.6\ yr$$

b) $\frac{\alpha}{2} = \frac{1-0.99}{2} = 0.005 \quad \Rightarrow A = 0.995$

$$z_{\alpha/2} = z_{0.005} = 2.575$$

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$= 2.575\left(\frac{848}{\sqrt{76}}\right)$$

$$\approx \$250$$

Exercise

Use the confidence level and sample data to find a confidence interval for estimating the population μ. A laboratory tested 89 chicken eggs and found that the mean amount of cholesterol was 203 milligrams with $\sigma = 11.4\ mg$. Construct a 95% confidence interval for the true mean cholesterol content μ, of all such eggs.

Solution

$n = 89 \qquad \overline{x} = 203$

$\frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \quad \Rightarrow A = 0.975 \quad \Rightarrow z_{\alpha/2} = z_{0.025} = 1.96$

$$\overline{x} \pm \frac{z_{\alpha/2} \cdot \sigma}{n} = 203 \pm \frac{1.96(11.4)}{\sqrt{89}} = 203 \pm 2.37$$

$$203 - 2.37 < \mu < 203 + 2.37$$

$$201\ mg < \mu < 205\ mg$$

Exercise

Use the confidence level and sample data to find a confidence interval for estimating the population μ. A group of 66 randomly selected students have a mean score of 34.3 on a placement test. The population standard deviation $\sigma = 3$. What is the 90% confidence interval for the mean score, μ, of all students taking the test?

Solution

$$n = 66 \qquad \bar{x} = 34.3$$

$$\frac{\alpha}{2} = \frac{1-0.9}{2} = 0.05 \quad \Rightarrow A = 0.95 \ \Rightarrow \ z_{\alpha/2} = z_{0.05} = 1.645$$

$$\bar{x} \pm \frac{z_{\alpha/2} \cdot \sigma}{n} = 34.3 \pm \frac{1.645(3)}{\sqrt{66}} = 34.3 \pm 0.6$$

$$34.3 - 0.6 < \mu < 34.3 + 0.6$$

$$\underline{33.7 < \mu < 34.9}$$

Exercise

Use the given information to find the minimum sample size required to estimate an unknown population mean μ. Margin error: \$139, confidence level: 99%, σ = \$522

Solution

$$\frac{\alpha}{2} = \frac{1-0.99}{2} = 0.005 \quad \Rightarrow A = 0.995 \ \Rightarrow \ z_{\alpha/2} = z_{0.005} = 2.575$$

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 = \left(\frac{2.575 \times 522}{139}\right)^2 \ \underline{\approx 94}$$

Exercise

What does it mean when we say that the methods for constructing confidence intervals in this section are robust against departures from normality? Are the methods for constructing confidence intervals in this section robust against poor sampling methods?

Solution

Robust against departure from normality mans that the requirement that the original population be approximately normal is not a strong requirement, and that the methods of this section still give good results if the departure from normality is not too extreme. The methods of this section are not robust against poor sampling methods, as poor sampling methods can yield data that are entirely useless.

Exercise

Assume that we want tic instruct a confidence interval using the given confidence level

95%; $n = 23$; σ is unknown; population appears to be normally distributed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

Solution

σ unknown, normal population, $n = 23$; It is *t*-distribution with $df = 22$

$$\alpha = 0.05; \quad t_{df,\alpha/2} = t_{22,\,0.025} = 2.074$$

***t* Distribution: Critical *t* Values**

Degrees of Freedom	Area in One Tail 0.005	0.01	0.025	0.05	0.10
22	**2.819**	2.508	**2.074**	1.717	1.321

Exercise

Assume that we want tic instruct a confidence interval using the given confidence level

99%; $n = 25$; σ is known; population appears to be normally distributed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

Solution

σ unknown, normal population, *Use* ***z***

$$\alpha = 0.01; \quad z_{\alpha/2} = z_{0.005} = 2.575$$

z score	Area
1.645	0.9500
2.575	0.9950

Exercise

Assume that we want tic instruct a confidence interval using the given confidence level

99%; $n = 6$; σ is unknown; population appears to be very skewed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

Solution

σ unknown, population not normal, $n = 6$: Neither normal nor *t* applies.

Exercise

Assume that we want tic instruct a confidence interval using the given confidence level 90%; $n = 200$; σ = 15.0; population appears to be skewed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

Solution

σ known, population not normal, *n* = 200, *Use* **z**

$\alpha = 0.1; \quad z_{\alpha/2} = z_{0.05} = 1.645$

z score	Area
1.645	0.9500
2.575	0.9950

Exercise

Assume that we want tic instruct a confidence interval using the given confidence level 95%; $n = 9$; σ is unknown; population appears to be very skewed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

Solution

σ unknown, population not normal, *n* = 9: Neither normal nor *t* applies.

Exercise

Given 95% *confidence*; $n = 20, \quad \bar{x} = \$9004, \quad s = \$569$. Assume that the sample is a simple random and the population has a normal distribution.

a) Find the margin error

b) Find the confidence interval for the population mean μ.

Solution

σ unknown, normal population, *n* = 20; It is *t*-distribution with $df = 19$

$\alpha = 0.05; \quad t_{df,\alpha/2} = t_{19,\,0.025} = 2.093$

t Distribution: Critical *t* Values

Degrees of Freedom	Area in One Tail: 0.005	0.01	0.025	0.05	0.10
19	2.861	2.539	2.093	1.729	1.328

a) $E = \dfrac{t_{\alpha/2}\, s}{\sqrt{n}}$

$= \dfrac{(2.093)(569)}{\sqrt{20}}$

$= \$266$

b) $\bar{x} \pm E$

9004 ± 266

$9004 - 266 < \mu < 9004 + 266$

$\$8738 < \mu < \$9270$

Exercise

Given 99% *confidence*; $n = 7$, $\bar{x} = 0.12$, $s = 0.04$. Assume that the sample is a simple random and the population has a normal distribution.

a) Find the margin error

b) Find the confidence interval for the population mean μ.

Solution

σ unknown, normal population, *n* = 7; It is *t*-distribution with $df = 6$

$\alpha = 0.01;\quad t_{df,\alpha/2} = t_{6,\,0.005} = 3.707$

t Distribution: Critical *t* Values

Degrees of Freedom	Area in One Tail: 0.005	0.01	0.025	0.05	0.10
6	3.707	3.143	2.447	1.943	1.440

a) $E = \dfrac{t_{\alpha/2}\, s}{\sqrt{n}}$

$= \dfrac{(3.707)(0.04)}{\sqrt{7}}$

$= 0.06 \ \ grams/mile$

b) $\bar{x} \pm E$

0.12 ± 0.06

$0.12 - 0.06 < \mu < 0.12 + 0.06$

$0.06 < \mu < 0.18 \quad grams/mile$

Exercise

In a test of the effectiveness of garlic for lowering cholesterol, 47 subjects were treated with Garlicin, which is garlic in a processed tablet form. Cholesterol levels were measured before and after the treatment. The changes in their levels of LDL cholesterol (in mg/dL) have a mean of 3.2 and standard deviation of 18.6.

a) What is the best point estimate of the population mean net change LDL cholesterol after Garlicin treatment?

b) Construct a 95% confidence interval estimate of the mean net change in LDL cholesterol after the Garlicin treatment. What does the confidence interval suggest about the effectiveness of Garlicin in reducing LDL cholesterol?

Solution

a) $\bar{x} = 3.2\ \ mg/dl$

b) σ unknown, $n > 30$; use t-distribution with $df = 46\,[45]$

$\alpha = 0.05;\ \ t_{df,\,\alpha/2} = t_{46,\,0.025} = 2.014$

t Distribution: Critical *t* Values

Degrees of Freedom	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
45	2.690	2.412	2.014	1.679	1.301

$$\bar{x} \pm \frac{t_{\alpha/2}\, s}{\sqrt{n}}$$

$$3.2 \pm \frac{(2.014)(18.6)}{\sqrt{47}}$$

$$3.2 \pm 5.5$$

$$3.2 - 5.5 < \mu < 3.2 + 5.5$$

$$-2.3 < \mu < 8.7\ \ (mg/dl)$$

Since the confidence interval includes 0, there is a reasonable possibility that the true value is zero – i.e., that the Garlicin treatment has no effect on LDL cholesterol levels.

Exercise

A random sample of the birth weights of 186 babies has a mean of 3103 g and a standard deviation of 696 g. These babies were born to mothers who did not use cocaine during their pregnancies.

a) What is the best point estimate of the mean weight of babies born to mothers who did not use cocaine during their pregnancies?

b) Construct a 95% confidence interval estimate of the mean birth for all such babies.

c) Compare the confidence interval from part (*b*) to this confidence interval obtained from birth weights of babies born to mothers who used cocaine during pregnancy: $2608\ g < \mu < 2792\ g$. Does cocaine use appear to affect the birth weight of a baby?

Solution

a) $\bar{x} = 3103\ \ grams$

b) σ unknown, $n > 30$; use t-distribution with $df = 185\ [200]$

$\alpha = 0.05;\quad t_{df,\ \alpha/2} = t_{185,\ 0.025} = 1.972$

t Distribution: Critical t Values

Degrees of Freedom	Area in One Tail 0.005	0.01	0.025	0.05	0.10
200	2.601	2.345	1.972	1.653	1.286

$$\bar{x} \pm \frac{t_{\alpha/2}\, s}{\sqrt{n}}$$

$$3103 \pm \frac{(1.972)(696)}{\sqrt{186}}$$

$$3103 \pm 101$$

$$3103 - 101 < \mu < 3103 + 101$$

$$3002 < \mu < 3204\quad (grams)$$

c) Yes. Since the confidence interval for the mean birth weight for mothers who used cocaine is entirely below the confidence interval in part (b), it appears that cocaine use is associated with lower birth rates.

Exercise

In a study designed to test the effectiveness of acupuncture for treating migraine, 142 subjects were treated with acupuncture and 80 subjects were given a sham treatment. The numbers of migraine attacks for the acupuncture treatment group had a mean of 1.8 and a standard deviation of 1.4. The numbers of migraine attacks for the sham treatment group had a mean of 1.6 and standard deviation of 1.2

a) Construct a 95% confidence interval estimate of the mean number of migraine attacks for those treated with acupuncture.

b) Construct a 95% confidence interval estimate of the mean number of migraine attacks for those given a sham treatment.

c) Compare the two confidence intervals. What do the results about the effectiveness of acupuncture?

Solution

a) σ unknown, $n > 30$; use *t*-distribution with $df = 141$ $[100]$

$\alpha = 0.05;\quad t_{df,\ \alpha/2} = t_{141,\ 0.025} = 1.984$

t Distribution: Critical *t* Values					
Degrees of Freedom	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
100	2.626	2.364	1.984	1.660	1.290

$$\bar{x} \pm \frac{t_{\alpha/2}\, s}{\sqrt{n}}$$

$$1.8 \pm \frac{(1.984)(1.4)}{\sqrt{142}}$$

$$1.8 \pm 0.2$$

$$1.8 - 0.2 < \mu < 1.8 + 0.2$$

$$1.6 < \mu < 2.0 \quad (headaches)$$

b) σ unknown, $n > 30$; use *t*-distribution with $df = 79$ $[80]$

$\alpha = 0.05;\quad t_{df,\ \alpha/2} = t_{79,\ 0.025} = 1.990$

t Distribution: Critical *t* Values					
Degrees of Freedom	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
80	2.639	2.374	1.990	1.664	1.292

$$\bar{x} \pm \frac{t_{\alpha/2}\, s}{\sqrt{n}}$$

$$1.6 \pm \frac{(1.990)(1.2)}{\sqrt{80}}$$

$$1.6 \pm 0.3$$

$$1.6 - 0.3 < \mu < 1.6 + 0.3$$

$$1.3 < \mu < 1.9 \quad (headaches)$$

c) The 2 confidence intervals are very similar and overlap considerably. There is no evidence that the acupuncture treatment is effective.

Exercise

30 randomly selected students took the statistics final. If the sample mean was 79 and the standard deviation was 14.5, construct a 99% confidence interval for the mean score of all students.
Use the given degree of confidence and sample data to construct a confidence level interval for the population mean μ. Assume that the population has a normal distribution.

Solution

$s = 14.5,\ \bar{x} = 79,\ n = 30$

$df = 29\ [30]$

$\alpha = 1 - .99 = 0.01;\quad t_{df,\ \alpha/2} = t_{29,\ 0.005} = 2.756$

t Distribution: Critical *t* Values

Degrees of Freedom	Area in One Tail 0.005	0.01	0.025	0.05	0.10
29	2.756	2.462	2.045	1.699	1.311

$$\bar{x} \pm E = \bar{x} \pm \frac{t_{\alpha/2}s}{\sqrt{n}} = 79 \pm \frac{(2.756)(14.5)}{\sqrt{30}} = 79 \pm 7.296$$

$79 - 7.296 < \mu < 79 + 7.296$

$71.70 < \mu < 86.30$

Solution Section 3.4 – Estimating a Population Standard Deviation

Exercise

Using the weights of the M&M candies. We use the standard deviation of the sample $(s = 0.05179\ g)$ to obtain the following 95% confidence interval estimate of the standard deviation of the weights of all M&Ms: $0.0455\ g < \sigma < 0.0602\ g$. Write a statement that correctly interprets that confidence interval.

Solution

We can be 95% confident that the interval from 0.0455 grams to 0.0602 grams includes the true value of the standard deviation in the weights for the population of all M&M's

Exercise

Find χ^2_L and χ^2_R that corresponds to: 95%; $n = 9$

Solution

$$\alpha = 0.05 \to \frac{\alpha}{2} = 0.025 \quad and \quad df = 8$$

TABLE Chi-Square (χ^2) Distribution

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955

$$\chi^2_L = \chi^2_{8,\ 0.0975} \underline{= 2.180} \qquad \chi^2_R = \chi^2_{8,\ 0.025} \underline{= 17.535}$$

Exercise

Find χ^2_L and χ^2_R that corresponds to: 99%; $n = 81$

Solution

$$\alpha = 1 - .99 = 0.01 \to \frac{\alpha}{2} = 0.005 \quad and \quad df = 80$$

TABLE Chi-Square (χ^2) Distribution

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.32

$$\chi^2_L = \chi^2_{80,\ 0.0995} \underline{= 51.172} \qquad \chi^2_R = \chi^2_{80,\ 0.005} \underline{= 116.321}$$

Exercise

Find χ_L^2 and χ_R^2 that corresponds to: 90%; $n = 51$

Solution

$$\alpha = 1 - .90 = 0.1 \to \frac{\alpha}{2} = 0.05 \quad and \quad df = 50$$

TABLE	Chi-Square (χ^2) Distribution									
Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490

$$\chi_L^2 = \chi_{50,\ 1-.05}^2 = \chi_{50,\ .95}^2 \underline{= 34.764}$$

$$\chi_R^2 = \chi_{50,\ .05}^2 \underline{= 67.505}$$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005

Exercise

Find a confidence interval for the population standard deviation σ

95% *confidence*; $n = 30$, $\overline{x} = 1533$, $s = 333$ (*Assume has a normal distribution*)

Solution

$$\alpha = 1 - .95 = 0.05 \to \frac{\alpha}{2} = 0.025 \quad and \quad df = 29$$

TABLE	Chi-Square (χ^2) Distribution									
Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336

$$\chi_L^2 = \chi_{29,\ 1-.025}^2 = \chi_{29,\ .975}^2 \underline{= 16.047}$$

$$\chi_R^2 = \chi_{29,\ .025}^2 \underline{= 45.722}$$

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

$$\sqrt{\frac{(29)(333)^2}{45.722}} < \sigma < \sqrt{\frac{(29)(333)^2}{16.047}}$$

$\underline{265 < \sigma < 448}$

Exercise

Find a confidence interval for the population standard deviation σ

95% *confidence*; $n = 25$, $\bar{x} = 81.0\ mi/h$, $s = 2.3\ mi/h$ (*Assume has a normal distribution*)

Solution

$$\alpha = 1 - .95 = 0.05 \to \frac{\alpha}{2} = 0.025 \quad and \quad df = n - 1 = 24$$

TABLE Degrees of Freedom	Chi-Square (χ^2) Distribution									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559

$$\chi_L^2 = \chi_{24,\ 1-.025}^2 = \chi_{24,\ .975}^2 \ \underline{= 12.401} \qquad \chi_R^2 = \chi_{24,\ .025}^2 \ \underline{= 39.364}$$

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

$$\sqrt{\frac{(24)(2.3)^2}{39.364}} < \sigma < \sqrt{\frac{(24)(2.3)^2}{12.401}}$$

$$\underline{1.8 < \sigma < 3.2} \quad (mph)$$

Exercise

Find a confidence interval for the population standard deviation σ

99% *confidence*; $n = 7$, $\bar{x} = 7.106$, $s = 2.019$ (*Assume has a normal distribution*)

Solution

$$\alpha = 1 - .99 = 0.01 \to \frac{\alpha}{2} = 0.005 \quad and \quad df = n - 1 = 6$$

TABLE Degrees of Freedom	Chi-Square (χ^2) Distribution									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548

$$\chi_L^2 = \chi_{6,\ 1-.005}^2 = \chi_{6,\ .995}^2 \ \underline{= 0.676} \qquad \chi_R^2 = \chi_{6,\ .025}^2 \ \underline{= 18.548}$$

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

$$\sqrt{\frac{(6)(2.019)^2}{18.548}} < \sigma < \sqrt{\frac{(6)(2.019)^2}{0.676}}$$

$$\underline{1.148 < \sigma < 6.015} \quad (cells / microliter)$$

Exercise

In a study of the effects of prenatal cocaine use on infants, the following sample data were obtained for weights at birth: $n = 190, \quad \bar{x} = 2700\ g, \quad s = 645\ g$. Use the sample data to construct a 95% confidence interval estimate of the standard deviation of all birth weights of infants born to mothers who used cocaine during pregnancy. Because from the Table, a maximum of 100 degrees of freedom while we require 189 degrees of freedom, use these critical values to obtained $\chi_L^2 = 152.8222$ and $\chi_R^2 = 228.9638$.

Based on the result, does the standard deviation appear to be different from the standard deviation of 696g for birth weights of babies born to mothers who did not use cocaine during pregnancy?

Solution

Given: $\chi_L^2 = 152.8222 \qquad \chi_R^2 = 228.9638$

$\alpha = 1 - .95 = 0.05 \rightarrow \frac{\alpha}{2} = 0.025 \quad$ *and* $\quad df = n - 1 = 189$

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

$$\sqrt{\frac{(189)(645)^2}{228.9638}} < \sigma < \sqrt{\frac{(189)(645)^2}{152.8222}}$$

$$\underline{586 < \sigma < 717} \quad (g)$$

No. Since the confidence interval includes 696, it is a reasonable possibility for σ.

Exercise

In the course of designing theather seats, the sitting heights (in *mm*) of a simple random sample of adults women is obtained, and the results are

849 807 821 856 864 877 772 848 802 807 887 815

Use the sample data to construct a 95% confidence interval estimate of σ, the standard deviation of sitting heights of all women. Does the confidence contain the value of 35 *mm*, which is believed to be the standard deviation of sitting heights of women?

Solution

```
1-Var Stats
 x̄=833.75000
 Σx=10005.00000
 Σx²=8354987.00
 Sx=34.79583
 σx=33.31447
↓n=12.00000
```

Using the calculator: $n = 12, \quad \bar{x} = 833.75 \quad s = 34.796$

$\alpha = 1 - .95 = 0.05 \rightarrow \frac{\alpha}{2} = 0.025 \quad$ *and* $\quad df = n - 1 = 11$

TABLE Chi-Square (χ^2) Distribution

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757

$$\chi_L^2 = \chi^2_{11,\ 1-.025} = \chi^2_{11,\ .975} \underline{= 3.816|} \qquad \chi_R^2 = \chi^2_{11,\ .025} \underline{= 21.920|}$$

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

$$\sqrt{\frac{(11)(34.796)^2}{21.920}} < \sigma < \sqrt{\frac{(11)(34.796)^2}{3.816}}$$

$$\underline{24.7 < \sigma < 59|} \quad (mm)$$

Yes. The interval contains the traditionally believed value of 35 *mm*.

Exercise

One way to measure the risk of a stock is through the standard deviation rate of return of the stock. The following data represent the weekly rate of return (in percent) of Microsoft for 15 randomly selected weeks. Compute the 90% confidence interval for the risk of Microsoft stock.

5.34	9.63	–2.38	3.54	–8.76	2.12	–1.95	0.27
0.15	5.84	–3.90	–3.80	2.85	–1.61	–3.31	

Solution

A normal probability plot and boxplot indicate the data is approximately normal with no outliers.

$s = 4.6974 \quad s^2 = 22.0659$

$df = 15 - 1 \underline{= 14|}$

Degrees of Freedom	Chi-Square (χ^2) Distribution — Area to the Right of Critical Value									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319

$\chi^2_{0.95} = 6.571 \qquad \chi^2_{0.05} = 23.685$

Lower bound: $\dfrac{(n-1)s^2}{\chi_R^2} = \dfrac{14(22.0659)}{23.685} \underline{= 13.04|}$

Upper bound: $\dfrac{(n-1)s^2}{\chi_L^2} = \dfrac{14(22.0659)}{6.571} \underline{= 47.01|}$

We are 90% confident that the population standard deviation rate of return of the stock is between 13.04 and 47.01.

Solution *Section* 3.5 – Language of Hypothesis Testing

Exercise

Bottles of Bayer aspirin are labeled with a statement that the tablets each contain 325 mg of aspirin. A quality control manager claims that a large sample of data can be used to support the claim that the mean amount of aspirin in the tablets is equal to 325 mg, as the label indicates. Can a hypothesis test be used to support that claim? Why or Why not?

Solution

No. Since the claim that the mean is equal to a specific value must be the null hypothesis, the only possible conclusions are to reject that claim or to fail to reject that claim, Hypothesis testing cannot be used to support a claim that a parameter is equal to a particular value.

Exercise

In the preliminary results from couples using the Gender Choice method of gender selection to increase the likelihood of having a baby girl, 20 couples used the Gender Choice method with the result that 8 of them had baby girls and 12 had baby boys. Given that the sample proportion of girls is $\frac{8}{20}$ *or* 0.4, can the sample data support the claim that the proportion of girls is greater than 0.5? Can any sample proportion less than 0.5 be used to support a claim that the population proportion is greater than 0.5?

Solution

No. Sample data that is not consistent with a claim can't be used to support that claim. In particular, no sample proportion less than 0.5 can ever be used to support a claim that the population proportion is greater than 0.5.

Exercise

Express the null hypothesis H_0 and alternative hypothesis H_1 in symbolic form. Be sure to use the correct symbol $(\mu,\ p,\ \sigma)$ for indicated parameter

a) The mean annual income of employees who took a statistics course is greater than \$60,000.
b) The proportion of people aged 18 to 25 who currently use illicit drugs is equal to 0.20 (or 20%).
c) The standard deviation of human body temperatures is equal to 0.62°F.
d) The majority of college students have credit cards.
e) The proportion of homes with fire extinguishers is 0.80.
f) The mean weight of plastic discarded by households in one week is less than 1 kg.

Solution

a) Original claim: $\mu > \$60,000$

$H_0: \mu = \$60,000 \qquad H_1: \mu > \$60,000$

b) Original claim: $p = 0.20$

$H_0 : p = 0.20 \qquad H_1 : p \neq 0.20$

c) Original claim: $p = 0.20$

$H_0 : \sigma = 0.62°F \qquad H_1 : p \neq 0.62°F$

d) Original claim: $p > 0.5$

$H_0 : p = 0.5 \qquad H_1 : p > 0.5$

e) Original claim: $p = 0.80$

$H_0 : p = 0.80 \qquad H_1 : p \neq 0.80$

f) Original claim: $\mu < 1\ kg$

$H_0 : \mu = 1\ kg \qquad H_1 : \mu < 1\ kg$

Exercise

Assume that the normal distribution applies and find the critical z values.

a) Two-tailed test: $\alpha = 0.01$.

b) Right-tailed test: $\alpha = 0.02$.

c) Left-tailed test: $\alpha = 0.10$.

d) $\alpha = 0.05;\quad H_1$ *is* $p \neq 0.4$

e) $\alpha = 0.01;\quad H_1$ *is* $p > 0.5$

f) $\alpha = 0.005;\quad H_1$ *is* $p < 0.8$

g) $\alpha = 0.05$ for two-tailed test

h) $\alpha = 0.05$ for left-tailed test

i) $\alpha = 0.08;\quad H_1$ *is* $\mu \neq 3.25$

Solution

a) Two-tailed test; place $\frac{\alpha}{2} = \frac{0.01}{2} = 0.005$ in each tail.

z score	Area
1.645	0.9500
2.575	0.9950

$A = 1 - \frac{\alpha}{2} = 0.995$

Critical value: $\pm z_{\alpha/2} = \pm z_{0.005}\ \underline{= \pm 2.575}$

b) Right-tailed test; place $\alpha = 0.02$ in the upper tail. $\Rightarrow A = 1 - \alpha = 0.98$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817

Critical value: $z_{\alpha/2} = z_{0.02}\ \underline{= 2.05}$

c) Left-tailed test; place $\alpha = 0.10$ in the lower tail. $\Rightarrow A = \alpha = 0.1$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985

Critical value: $z_{\alpha} = z_{0.1}\ \underline{= -1.28}$

d) Two-tailed test; place $\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$ in each tail. $\rightarrow A = 1 - \frac{\alpha}{2} = 0.975$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
$\boxed{1.9}$	.9713	.9719	.9726	.9732	.9738	.9744	$\boxed{.9750}$	.9756	.9761	.9767

Critical value: $\pm z_{\alpha/2} = \pm z_{0.025}$ $\underline{= \pm 1.96|}$

e) Right-tailed test; place $\alpha = 0.01$ in the upper tail. $A = 1 - \alpha = 0.99$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
$\boxed{2.3}$	.9893	.9896	.9898	$\boxed{.9901}$	.9904	.9906	.9909	.9911	.9913	.9916

Critical value: $z_{\alpha} = z_{0.02}$ $\underline{= 2.33|}$

f) Left-tailed test; place $\alpha = 0.005$ in the lower tail. $\Rightarrow A = 1 - \alpha = 0.995$

Critical value: $-z_{\alpha} = -z_{0.005}$ $\underline{= -2.575|}$

z score	Area
1.645	0.9500
2.575	0.9950

g) Two-tailed test; place $\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$ in each tail.

$A = 1 - \frac{\alpha}{2} = 0.975$

Critical value: $\pm z_{\alpha/2} = \pm z_{0.025}$ $\underline{= \pm 1.96|}$

h) Left-tailed test; $\alpha = 0.05$ $\Rightarrow A = 1 - \alpha = 0.95$

Critical value: $z_{\alpha} = z_{0.05}$ $\underline{= -1.645|}$

i) Two-tailed test; place $\frac{\alpha}{2} = \frac{0.08}{2} = 0.04$ in each tail. $A = \alpha = 0.04$

Critical value: $\pm z_{\alpha/2} = \pm z_{0.04}$ $\underline{= \pm 1.75|}$

Exercise

The claim is that the proportion of peas with yellow pods is equal to 0.25 (or 25%). The sample statistics from one of Mendel's experiments include 580 peas with 152 of them having yellow pods. Find the value of the test statistic z using $z = \dfrac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$

Solution

$$\hat{p} = \frac{x}{n} = \frac{152}{580} = 0.262$$

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.262 - 0.250}{\sqrt{\frac{(0.25)(.75)}{580}}} \underline{= 0.67|}$$

Exercise

The claim is that less than $\frac{1}{2}$ of adults in U.S. have carbon monoxide detectors. A KRC Research survey of 1005 adults resulted in 462 who have carbon monoxide detectors. Find the value of the test statistic z using $z=\dfrac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$

Solution

$$\hat{p}=\frac{x}{n}=\frac{462}{1005}=0.460$$

$$z=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} \quad =\frac{0.46-0.5}{\sqrt{\frac{(0.5)(.5)}{1005}}} \quad \underline{=-2.56|}$$

Exercise

The claim is that more than 25% of adults prefer Italian food as their favorite ethnic food. A Harris Interactive survey of 1122 adults resulted in 314 who say that Italian food is their favorite ethnic food. Find the value of the test statistic z using $z=\dfrac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$

Solution

$$\hat{p}=\frac{x}{n}=\frac{314}{1122}=0.28$$

$$z=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} \quad =\frac{0.28-0.25}{\sqrt{\frac{(0.25)(.75)}{1122}}} \quad \underline{=2.31|}$$

Exercise

Find *P*-value by using a 0.05 significance level and state the conclusion about the null hypothesis. (Reject the null hypothesis or fail to reject the null hypothesis)

a) The test statistic in a left-tailed test is $z = -1.25$

b) The test statistic in a right-tailed test is $z = 2.50$

c) The test statistic in a two-tailed test is $z = 1.75$

d) With $H_1: p \neq 0.707$, the test statistic is $z = -2.75$

e) With $H_1: p > \frac{1}{4}$, the test statistic is $z = 2.30$

f) With $H_1: p < 0.777$, the test statistic is $z = -2.95$

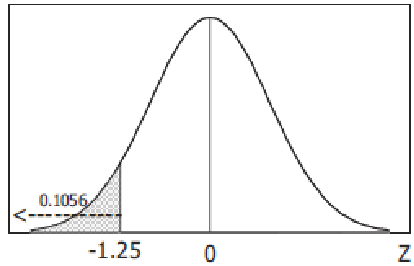
Solution

a) $P\text{-value} = P(z < -1.25)$

$= 0.1056$

z	.00	.01	.02	.03	.04	.05
-1.2	.1151	.1131	.1112	.1093	.1075	.1056

Since 0.1056 > 0.05, fail to reject H_0



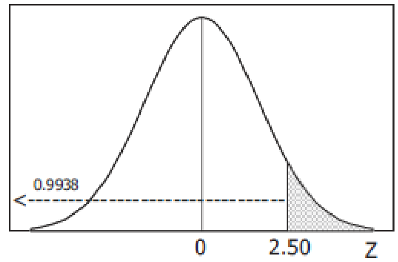
b) $P\text{-value} = P(z > 2.5)$

$= 1 - 0.9938$

$= 0.0062$

z	.00	.01
2.5	.9938	

Since 0.0062 < 0.05, reject H_0

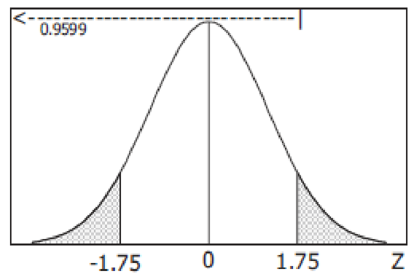


c) $P\text{-value} = 2 \cdot P(z > 1.75)$

$= 2(1 - 0.9599)$

$= 0.0802$

Since 0.0802 > 0.05, fail to reject H_0

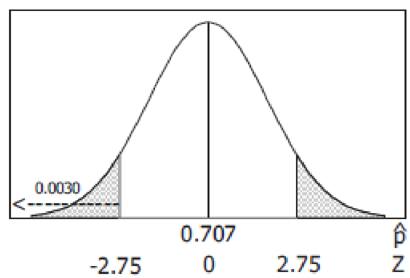


d) $P\text{-value} = 2 \cdot P(z < -2.75)$

$= 2(0.003)$

$= 0.006$

Since 0.006 > 0.05, reject H_0

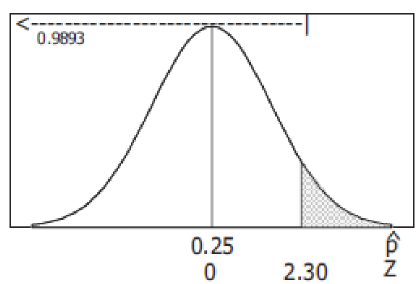


e) $P\text{-value} = P(z > 2.3)$

$= 1 - 0.9893$

$= 0.0107$

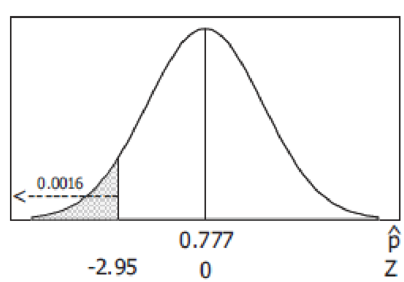
Since 0.0107 < 0.05, reject H_0



f) $P\text{-value} = P(z < -2.95)$

$= 0.0016$

Since 0.0016 < 0.05, reject H_0



Exercise

The percentage of nonsmokers exposed to secondhand smoke is equal to 41%. Identify the type I error and type II error.

Solution

Original claim: $p = 0.41$

$H_0 : p = 0.41$

Type I error: rejecting H_0 when H_0 is actually true rejecting the claim that the percentage of non-smokers exposed to secondhand smoke is 41% when that percentage actually is 41%

Type II error: failing to reject H_0 when H_1 is actually true failing to reject the claim that the percentage is actually different from 41%

Exercise

The percentage of Americans who believe that life exists only on earth is equal to 20%. Identify the type I error and type II error.

Solution

Original claim: $p = 0.20$ $\qquad H_0 : p = 0.20$

Type I error: rejecting H_0 when H_0 is actually true rejecting the claim that the percentage of Americans who believe that life exists only on earth is 20% when that percentage actually is 20%

Type II error: failing to reject H_0 when H_1 is actually true failing to reject the claim that the percentage of Americans who believe that life exists only on earth is 20% when that percentage is actually different from 20%

Exercise

The percentage of college students who consume alcohol is greater than 70%. Identify the type I error and type II error.

Solution

Original claim: $p > 0.70$ $\qquad H_0 : p = 0.70$

Type I error: rejecting H_0 when H_0 is actually true rejecting the claim that the percentage of college students who use alcohol is 70% when that percentage actually is 70%.

Type II error: failing to reject H_0 when H_1 is actually true failing to reject the claim that the percentage of college students who use alcohol is 70% when that percentage actually is actually greater than 70%

Exercise

An entomologist writes an article in a scientific journal which claims that fewer than 13 in 10,000 male fireflies are unable to produce light due to a genetic mutation. Use the parameter p, the true proportion of fireflies unable to produce light. Express the null hypothesis and the alternative hypothesis in symbolic form. $(\mu,\ p,\ \sigma)$

Solution

$$p = \frac{13}{10,000} = 0.0013$$

Since the claims are fewer than it will be "<"

$H_0 : p = 0.0013$

$H_1 : p < 0.0013$

Solution Section 3.6 – Hypothesis Tests for a Population Proportion

Exercise

In a Harris poll, adults were asked if they are in favor of abolishing the penny. Among the responses, 1261 answered "no", and 491 answered "yes", and 384 had no opinion. What is the sample proportion of *yes* responses, and what notation is used to represent it?

Solution

Total Responses: $1261+491+384 = 2136$

The sample proportion of yes responses is $\hat{p}=\frac{x}{n}=\frac{491}{2136} = 0.230$

$\hat{p}$: is used to represent a sample proportion.

Exercise

A recent study showed that 53% of college applications were submitted online. Assume that this result is based on a simple random sample of 1000 college applications, with 530 submitted online. Use a 0.01 significance level to test the claim that among all college applications the percentage submitted online is equal to 50%

a) What is the test statistic?
b) What are the critical values?
c) What is the *P*-Value?
d) What is the conclusion?
e) Can a hypothesis test be used to "prove" that the percentage of college applications submitted online is equal to 50% as claimed?

Solution

a) $\hat{p}=\frac{x}{n}=\frac{530}{1000} = 0.530$

$$z_{\hat{p}}=\frac{\hat{p}-\mu_{\hat{p}}}{\sigma_{\hat{p}}}=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$$

$$=\frac{0.53-0.5}{\sqrt{\frac{(.5)(.5)}{1000}}}$$

$$= 1.90$$

z score	Area
1.645	0.9500
2.575	0.9950

b) $\frac{\alpha}{2}=\frac{0.01}{2}=0.005 \quad \Rightarrow \quad A=1-0.005=0.995$

$$z=\pm z_{\alpha/2}=\pm z_{0.01/2}=\pm z_{0.005} = \pm 2.575$$

c) $P-value=2\cdot P(z>1.90)$

$$=2\cdot(1-0.9713)$$

z	.00	.01

$\underline{= 0.0574}$

1.9	.9713	.9719

d) Do not reject H_0; there is not sufficient evidence to reject the claim that the percentage of all college applications that are submitted online is 50%.

e) No. A hypothesis test will either "reject" or "fail to reject" a claim that a population parameter is equal to a specified value.

Exercise

In a survey, 1864 out of 2246 randomly selected adults in the U.S. said that texting while driving should be illegal. Consider a hypothesis test that uses a 0.05 significance level to test the claim that more than 80% of adults believe that testing while driving should be illegal

a) What is the test statistic?
b) What are the critical values?
c) What is the *P*-Value?
d) What is the conclusion?

Solution

a) $\hat{p} = \frac{x}{n} = \frac{1864}{2246} \underline{= 0.830}$

$$z_{\hat{p}} = \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.83 - 0.80}{\sqrt{\frac{(0.8)(0.2)}{2246}}}$$

$$\underline{= 3.54}$$

b) $\alpha = 0.05 \quad \Rightarrow \quad A = 1 - 0.05 = 0.95$

$z = z_\alpha = z_{0.05} \underline{= 1.645}$

z score	Area
$\boxed{1.645}$	$\boxed{0.9500}$
2.575	0.9950

c) $P - value = P(z > 3.54)$

$= 1 - 0.9999$

$\underline{= 0.0001}$

3.50 and up	.9999

d) Reject H_0; there is sufficient evidence to conclude that the proportion of adults who believe that texting while driving should be illegal is greater than 80%.

Exercise

In a Pew Research Center poll of 745 randomly selected adults, 589 said that it is morally wrong to not report all income on tax returns. Use a 0.01 significance level to test the claim that 75% of adults say that it is morally wrong to not report all income on tax returns.
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

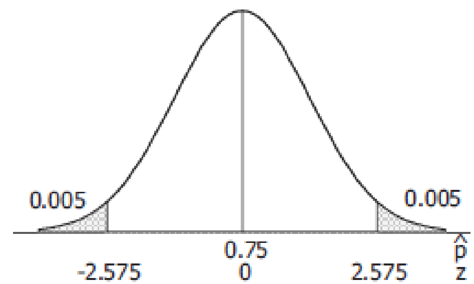
Solution

Original claim: $p = 0.75$

$\hat{p} = \dfrac{x}{n} = \dfrac{589}{745} \underline{= 0.791}$

$H_0 : p = 0.75$

$H_1 : p \neq 0.75$



Assume: $\alpha = 0.01$

$\dfrac{\alpha}{2} = \dfrac{0.01}{2} = 0.005 \quad \Rightarrow \quad A = 1 - 0.005 = 0.995$

Critical value: $z = \pm z_{\alpha/2} = \pm z_{0.005} = \pm 2.575$

z score	Area
1.645	0.9500
2.575	0.9950

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.791 - 0.75}{\sqrt{\frac{(0.75)(0.25)}{745}}}$$

$$\underline{= 2.56}$$

$P - value = 2 \cdot P(z > 2.56)$

$= 2 \cdot (1 - 0.9948)$

$\underline{= 0.0104}$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952

Conclusion

Reject H_0; there is sufficient evidence to conclude that the proportion of adults who believe that texting while driving should be illegal is greater than 80%.

Exercise

308 out of 611 voters surveyed said that they voted for the candidate who won. Use a 0.01 significance level to test the claim that among all voters, the percentage who believe that they voted for the winning candidate is equal to 43%, which is the actual percentage of votes for the winning candidate. What does the result suggest about voter perceptions?
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

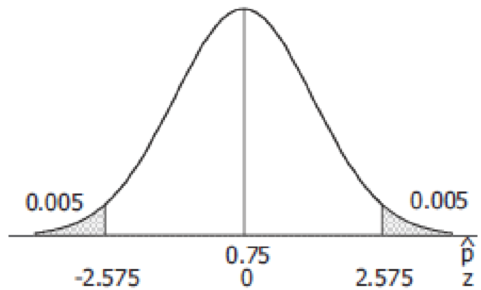
Solution

Original claim: $p = 0.43$

$$\hat{p} = \frac{x}{n} = \frac{308}{611} \underline{= 0.504}$$

$H_0 : p = 0.43$

$H_1 : p \neq 0.43$



Assume: $\alpha = 0.01$

$$\frac{\alpha}{2} = \frac{0.01}{2} = 0.005 \quad \Rightarrow \quad A = 1 - 0.005 = 0.995$$

Critical value: $z = \pm z_{\alpha/2} = \pm z_{0.005} = \pm 2.575$

z score	Area
1.645	0.9500
2.575	0.9950

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.504 - 0.43}{\sqrt{\frac{(0.43)(0.57)}{611}}}$$

$$\underline{= 3.70}$$

$$P - value = 2 \cdot P(z > 3.70)$$

$$= 2 \cdot (1 - 0.9999)$$

$$\underline{= 0.0002}$$

3.50 and up	.9999

Conclusion

Reject H_0; there is sufficient evidence to reject the claim that $p = 0.43$ and conclude that $p \neq 0.43$ (in fact, that $p > 0.43$). There is sufficient evidence to reject the claim that the proportion of adults who believe they voted for the winning candidate is 43%. Either the voters are deliberately not telling the truth, or they have faulty memories about how they actually voted.

Exercise

The company Drug Test Success provides a “1-Panel-THC” test for marijuana usage. Among 300 tested subjects, results from 27 subjects were wrong (either a false positive or a false negative). Use a 0.05 significance level to test the claim that less than 10% of the test results are wrong. Does the test appear to be good for most purposes?

Solution

Original claim: $p < 0.10$

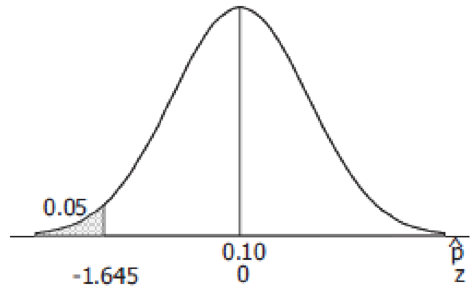
$$\hat{p} = \frac{x}{n} = \frac{27}{300} = 0.090$$

$H_0 : p = 0.10$

$H_1 : p < 0.10$

$\alpha = 0.05$

Critical Value: $z = -z_\alpha = -z_{0.05} \underline{= -1.645}$



z score	Area
1.645	0.9500
2.575	0.9950

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.09 - 0.1}{\sqrt{\frac{(0.1)(0.9)}{300}}}$$

$$\underline{= -0.58}$$

$P - value = P(z < -0.58)$

$\underline{= 0.2810}$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776

Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $p < 0.10$. There is not sufficient evidence to support the claim that the proportion of test results that are incorrect is less than 10%.

No; the test appears to have too high of an error rate to be considered reliable for most purposes.

Exercise

When testing gas pumps in Michigan for accuracy, fuel-quality enforcement specialists tested pumps and found that 1299 of them were not pumping accurately (within 3.3 oz. when 5 gal. is pumped), and 5686 pumps were accurate. Use a 0.01 significance level to test the claim of an industry representative that less than 20% of Michigan gas pumps are inaccurate. From the perspective of the consumer, does that rate appear to be low enough?

Solution

Original claim: $p < 0.20$

$$\hat{p} = \frac{x}{n} = \frac{1299}{1299+5686} = 0.186$$

$H_0 : p = 0.20$

$H_1 : p < 0.20$

$\alpha = 0.01$

Critical Value:

$z = -z_\alpha = -z_{0.01}$

$= -2.326$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.186 - 0.2}{\sqrt{\frac{(0.2)(0.8)}{6985}}}$$

$= -2.93$

$P - value = P(z < -2.93)$

$= 0.0017$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014

Conclusion

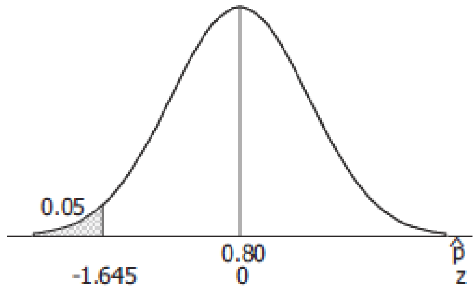
Reject H_0; there is sufficient evidence to conclude that $p < 0.20$. There is sufficient evidence to support the claim that less than 20% of Michigan gas pumps are inaccurate.

No; from the perspective of the consumer, the rate does not appear to be low enough. While the point estimate of 0.186 indicates the rate is lower than 20%. It should probably be about $\frac{1}{10}$ of that.

Exercise

Trials in an experiment with a polygraph include 98 results that include 24 cases of wrong results and 74 cases of correct results. Use a 0.05 significance level to test the claim that such polygraph results are correct less than 80% of the time. Based on the results, should polygraph test results be prohibited as evidence in trials?

Solution



Original claim: $p < 0.80$

$\hat{p} = \frac{x}{n} = \frac{74}{98} = 0.755$

$H_0 : p = 0.80$

$H_1 : p < 0.80$

$\alpha = 0.05$

Critical Value: $z = -z_{\alpha} = -z_{0.05} = -1.645$

z score	Area
1.645	0.9500
2.575	0.9950

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.755 - 0.8}{\sqrt{\frac{(0.8)(0.2)}{98}}}$$

$$= -1.11$$

$P - value = P(z < -1.11)$

$= 0.1335$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170

Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $p < 0.80$. There is not sufficient evidence to support the claim polygraph tests are correct less than 80% of the time.

Yes; based on these results, polygraph test results should probably be prohibited as evidence in trials. Even though the point estimate of 75.5% accuracy does not support the less than 80% claim, the accuracy rate is still far too small to make conclusions beyond a reasonable doubt.

Exercise

In recent years, the Town of Newport experienced an arrest rate of 25% for robberies. The new sheriff compiles records showing that among 30 recent robberies, the arrest rate is 30%, so she claims that her arrest rate is greater than the 25% rate in the past. Is there sufficient evidence to support her claim that the arrest rate is greater than 25%?

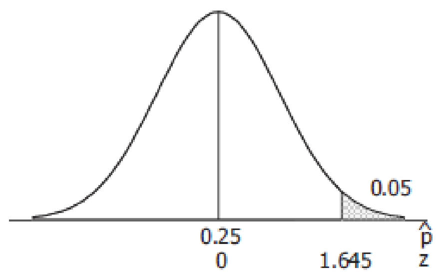
Solution

Original claim: $p > 0.25$

$$\hat{p} = \frac{x}{n} = \frac{(0.3)(30)}{30} = 0.300$$

$$H_0 : p = 0.25$$

$$H_1 : p > 0.25$$



z score	Area
1.645	0.9500
2.575	0.9950

Assumed: $\alpha = 0.05$

Critical Value: $z = z_{\alpha} = z_{0.05} = 1.645$

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.3 - 0.25}{\sqrt{\frac{(0.25)(0.75)}{30}}}$$

$$= 0.63$$

$$P - value = P(z > 0.63)$$

$$= 1 - .7357$$

$$= 0.2643$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549

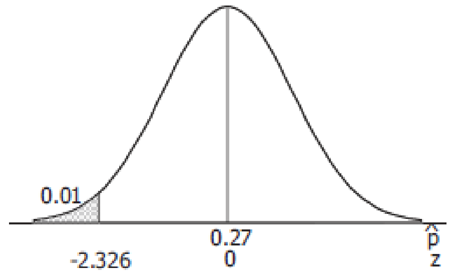
Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $p > 0.30$. There is not sufficient evidence to support the new sheriff's claim that the new arrest rate is greater than 25%.

Exercise

A survey showed that among 785 randomly selected subjects who completed 4 years of college, 18.3 % smoke and 81.7% do not smoke. Use a 0.01 significance level to test the claim that the rate of smoking among those with 4 years of college is less than the 27% rate for the general population. Why would college graduates smoke at a lower rate than others?

Solution



Original claim: $p < 0.27$

$$\hat{p} = \frac{x}{n} = \frac{(0.183)(785)}{785} = 0.183$$

$$H_0 : p = 0.27$$

$$H_1 : p < 0.27$$

$$\alpha = 0.01$$

Critical Value:

$$z = -z_{\alpha} = -z_{0.01}$$
$$= -2.326$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.183 - 0.27}{\sqrt{\frac{(0.27)(0.73)}{785}}}$$

$$\underline{= -5.46}$$

−3.50 and lower	.0001

$$P - value = P\left(z < -5.46\right) \underline{= 0.0001}$$

Conclusion

Reject H_0; there is sufficient evidence to conclude that $p < 0.27$. There is sufficient evidence to support the claim that the proportion of smokers among those with four years of college is less than the general rate of 27%. College graduates would smoke at a lower rate others because those with better education tend to make wiser decisions and are more likely to recognize the various disadvantages of smoking.

Exercise

When 3011 adults were surveyed, 73% said that they use the Internet. Is it okay for a newspaper reporter to write that "3/4 of all adults use the internet"? Why or Why not?

Solution

The value for x is not given,

$$(72.5\%)(3011) < x < (73.5\%)(3011)$$

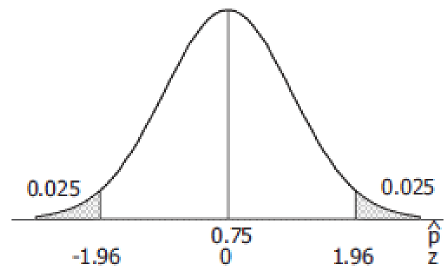
$$2183 < x < 2213$$

Original claim: $p = 0.75$

$$\hat{p} = \frac{x}{n} = \frac{x}{3011} = 0.73$$

$$H_0 : p = 0.75$$

$$H_1 : p \neq 0.75$$



Assumed: $\alpha = 0.05$

Critical Value:

$$z = \pm z_{\alpha/2} = \pm z_{0.025}$$

$$= \pm 1.96$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.73 - 0.75}{\sqrt{\frac{(0.75)(0.25)}{3011}}}$$

$$= -2.53$$

$$P - value = 2 \cdot P(z < -2.53)$$

$$= 2(0.0057)$$

$$= 0.0114$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	* .0049	.0048

Conclusion

Reject H_0; there is sufficient evidence to conclude that $p = 0.75$ and conclude that $p \neq 0.75$ (in fact, that $p < 0.75$). There is sufficient evidence to reject the claim that the proportion of adults who use the internet is $\frac{3}{4}$. While the difference between 0.73 and 0.75 may be of little practical significance, in the interest of accuracy the reporter should not write that $\frac{3}{4}$ of all adults use the internet.

Exercise

A hypothesis test is performed to test the claim that a population proportion is greater than 0.7. Find the probability of a type II error, β, given that the true value of the population proportion is 0.72. The sample size is 50 and the significance level is 0.05.

Solution

$$z=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} \;\Rightarrow\; \hat{p}-p=z\sqrt{\frac{pq}{n}} \;\Rightarrow\; \hat{p}=p+z\sqrt{\frac{pq}{n}}$$

z score	Area
1.645	0.9500

The rejection region for the test is any sample proportion greater than:

$n=50;\quad p=0.7;\quad q=0.3;\quad z=1.645$

$$\begin{aligned}\hat{p}&=p+z\sqrt{\frac{pq}{n}}\\ &=0.7+1.645\sqrt{\frac{0.7(0.3)}{50}}\\ &\approx 0.81\end{aligned}$$

The probability for the true value of 0.72

$$P\left(Z>\frac{0.81-0.72}{\sqrt{\frac{0.72(0.28)}{50}}}\right)=P(Z>1.42)\ \approx 0.9222$$

Because for a type II error, failure to reject the null when it is false, requires that the true value be in alternative space. Since the in the first set up, the value of 0.72 is in the null space, the null hypothesis is true and β does not exists, because a Type II error is impossible.

Exercise

In a sample of 88 children selected randomly from one town, it is found that 8 of them suffer asthma. Find the *P*-value for a test of the claim that the proportion of all children in the town who suffer from asthma is equal to 11%.

Solution

Original claim: $p=0.11\ \&\ \hat{p}=\frac{x}{n}=\frac{8}{88}\ =0.091$

$$z_{\hat{p}}=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}=\frac{.091-.11}{\sqrt{\frac{(.11)(0.89)}{88}}}\ =-0.57$$

$$\begin{aligned}P-value&=2\cdot P(z=-0.57)\\ &=2(.2843)\\ &=0.5686\end{aligned}$$

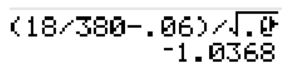
Exercise

An airline claims that the no-show rate for passengers booked on its flights is less than 6%. Of 380 randomly selected reservation, 18 were no-shows. Find the P-value for a test of the airline's claim.

Solution

Original claim: $p = 0.06$

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{18}{380} - .06}{\sqrt{\frac{(.06)(.94)}{380}}} \underline{=-1.04}$$



$$P - value = P(z < -1.04)$$

$$\underline{= 0.1492}$$

Exercise

In 1997, 46% of Americans said they did not trust the media "when it comes to reporting the news fully, accurately and fairly". In a 2007 poll of 1010 adults nationwide, 525 stated they did not trust the media. At the $\alpha = 0.05$ level of significance, is there evidence to support the claim that the percentage of Americans that do not trust the media to report fully and accurately has increased since 1997?

Solution

We want to know if $p > 0.46$. First, we must verify the requirements to perform the hypothesis test:

1. This is a simple random sample.
2. $np_0(1 - p_0) = 1010(0.46)(1 - 0.46) \underline{= 250.9 > 10}$
3. Since the sample size is less than 5% of the population size, the assumption of independence is met.

Step 1: $H_0 : p = 0.46 \quad vs \quad H_1 : p > 0.46$

Step 2: The level of significance is $\alpha = 0.05$.

Step 3: The sample proportion is $\hat{p} = \frac{525}{1010} = 0.52$. The test statistic is then

$$z_0 = \frac{0.52 - 0.46}{\sqrt{\frac{0.46(1-0.46)}{1010}}} \underline{= 3.83}$$

Classical Approach

Step 4: Since this is a right-tailed test, we determine the critical value at the $\alpha = 0.05$ level of significance to be $z_{0.05} = 1.645$

Step 5: Since the test statistic $z_0 = 3.83$, is greater than the critical value 1.645, we reject the null hypothesis.

P-Value Approach

Step 4: Since this is a right-tailed test, the P–value is the area under the standard normal distribution to the right of the test statistic $z_0 = 3.83$. That is, P–value = $P(Z > 3.83) \approx 0$.

Step 5: Since the P-value is less than the level of significance, we reject the null hypothesis.

Step 6: There is sufficient evidence at the $\alpha = 0.05$ level of significance to conclude that the percentage of Americans that do not trust the media to report fully and accurately has increased since 1997.

Exercise

In 2006, 10.5% of all live births in the United States were to mothers under 20 years of age. A sociologist claims that births to mothers under 20 years of age is decreasing. She conducts a simple random sample of 34 births and finds that 3 of them were to mothers under 20 years of age. Test the sociologist's claim at the $\alpha = 0.01$ level of significance.

Solution

Step 1: $H_0: p = 0.105 \quad vs \quad H_1: p > 0.105$

Step 2: From the null hypothesis, we have $p_0 = 0.105$. There were 34 mothers sampled, so

$$np_0\left(1 - p_0\right) = 32(0.105)(1 \; - \; 0.105) \; \boxed{= 3.20 < 10}$$

Thus, the sampling distribution of $\hat{p}$ is not approximately normal.

Step 3: Let X represent the number of live births in the United States to mothers under 20 years of age. We have $x = 3$ successes in $n = 34$ trials so $\hat{p} = \frac{3}{34} = 0.088$. If the population mean is truly 0.105. Thus,

$$\begin{aligned} P - value &= P\left(X \le 3 \text{ assuming } p = 0.105\right) \\ &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0.51 \end{aligned}$$

Step 4: The P-value = 0.51 is greater than the level of significance so we do not reject H_0. There is insufficient evidence to conclude that the percentage of live births in the United States to mothers under the age of 20 has decreased below the 2006 level of 10.5%.

Solution   Section 3.7 – Hypothesis Tests for a Population Mean

Exercise

Because the amounts of nicotine in king size cigarettes listed below

1.1	1.7	1.7	1.1	1.1	1.4	1.1	1.4	1	1.2	1.1	1.1	1.1
1.1	1.1	1.8	1.6	1.1	1.2	1.5	1.3	1.1	1.3	1.1	1.1	

We must satisfy the requirement that the population is normally distributed. How do we verify that a population is normally distributed?

Solution

We consider the normality requirement to be satisfied if there are no outliers and the histogram of the sample data is approximately bell-shaped. More formally, a normal quantile plot could be used to determine whether the sample data are approximately normally distributed.

Exercise

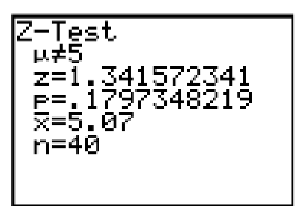
If you want to construct a confidence interval to be used for testing the claim that college students have a mean IQ score that is greater than 100, and you want the test conducted with a 0.01 significance level, what confidence level should be used for the confidence interval?

Solution

A one-tailed test at the 0.01 level of significance rejects the null hypothesis if the sample statistic falls into the extreme 1% of the sampling distribution in the appropriate tail. The corresponding (two-sided) confidence interval test that places 1% each tail would be a 98% confidence interval.

Exercise

A jewelry designer claims that women have wrist breadths with a mean equal to 5 cm. A simple random sample of the wrist breadths of 40 women has a mean of 5.07 cm. Assume that the population standard deviation is 0.33 cm. Use the accompanying TI display to test the designer's claim.



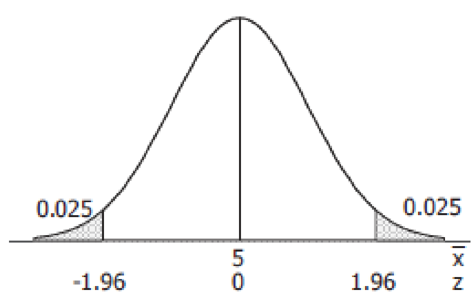
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

Solution

Original claim: $\mu = 5$ cm

$H_0: \ \mu = 5cm$

$H_1: \ \mu \neq 5cm$



Assume: $\alpha = 0.05$

Critical value: $z = \pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$

P-value = 0.1797 [*TI*]

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}} = 1.34 \quad [TI]$$

Conclusion

Do not reject H_0; there is not sufficient evidence to reject the claim that $\mu = 5$. There is not sufficient evidence the claim that women have a mean wrist breadth equal to 5 cm.

Exercise

The U.S. Mint has a specification that pennies have a mean weight of 2.5 g. Assume that weights of pennies have a standard deviation of 0.0165 g and use the accompanying Minitab display to test the claim that the sample is from a population with a mean that is less than 2.5 g. These Minitab results were obtained using the 37 weights of post 1983 pennies.

```
Test of mu = 2.5 vs < 2.5. Assumed s.d. = 0.0165
                              95% Upper
 N      Mean      StDev      Bound          Z         P
37   2.49910   0.01648    2.50356      -0.33     0.370
```

Solution

Original claim: $\mu < 2.5$ g.

$H_0: \mu = 2.5\, g$

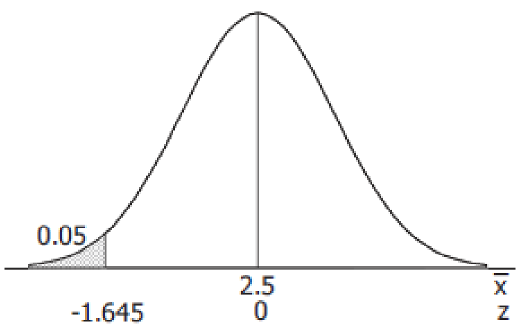
$H_1: \mu < 2.5\, g$

Assume: $\alpha = 0.05$

Critical value: $z = -z_{\alpha} = -z_{0.05} = -1.645$

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}} = -0.33 \quad [Minitab]$$

P-value = 0.370 [***Minitab***]



Conclusion

Do not reject H_0; there is not sufficient evidence to reject the claim that $\mu < 2.5$. There is not sufficient evidence to support the claim the sample is from a population with a mean less than 2.5 g.

Exercise

In the manual "How long to have a Number One the Easy Way," by KLF Publications, it is stated that a song "must be no longer than 3 minutes and 30 seconds" (or 210 seconds). A simple random sample of 40 current hit songs results in a mean length of 252.5 sec. Assume that the standard deviation of song lengths is 54.5 sec. Use a 0.05 significance level to test the claim that the sample is from a population of songs with a mean greater than 210 sec. What do these results suggest about the advice given in the manual?

Solution

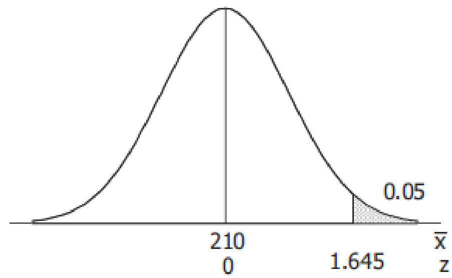
Original claim: $\mu > 210$ sec.

$H_0: \mu = 210$ sec

$H_1: \mu > 210$ sec

Given: $\alpha = 0.05$

Critical value: $z = z_\alpha = z_{0.05} = 1.645$



z score	Area
1.645	0.9500

$$z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{252.5 - 210}{\frac{54.5}{\sqrt{40}}} = 4.93$$

$$P-value = P(z > 4.93) = 1 - 0.9999 = 0.0001$$

Conclusion

Reject H_0; there is sufficient evidence to conclude that $\mu > 210$. There is sufficient evidence to support the claim the sample is from a population of songs with a mean greater than 210 sec. These results suggest that the advice given in the manual is not good advice.

Exercise

A simple random sample of 50 adults is obtained, and each person's red blood cell count (in cells per microliter) is measured. The sample mean is 5.23. The population standard deviation for red blood cell counts is 0.54. Use a 0.01 significance level to test the claim that the sample is from population with a mean less than 5.4, which is value often used for the upper limit of the range of normal values. What do the results suggest about the sample group?

Solution

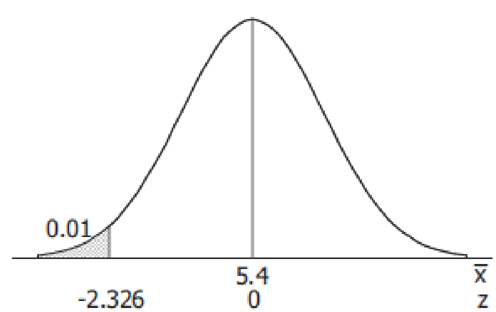
Original claim: $\mu < 5.4$ cells/microliter.

$H_0: \mu = 5.4\ cells / microliter$

$H_1: \mu < 5.4\ cells / microliter$

Given: $\alpha = 0.01$

z	0.00	0.01	0.02	0.03
−2.3	0.0107	0.0104	0.0102	0.0099



Critical value: $z = -z_\alpha = -z_{0.01} = -2.326$

$$z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{5.23 - 5.4}{\frac{0.54}{\sqrt{50}}} = -2.23$$

$(5.23 - 5.4) / (0.54 / \sqrt{50})$

$$P-value = P(z < -2.23) = 0.0129$$

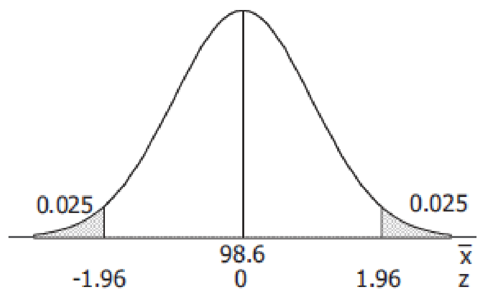
Conclusion

Do not reject H_0; there is not sufficient evidence to reject the claim that $\mu < 5.4$. There is not sufficient evidence to support the claim the sample is from a population with a mean red blood cell count less than 504 cells/microliter. If 5.4 is the upper limit for the range of normal individuals, and if the mean of the sample group is not significantly below the upper limit for an individual and the mean of the sample group may have manually high red cell counts. The $\mu \pm 2\sigma$ guideline for normal values suggests that the population mean is approximately $5.4-2(0.54)=4.32$, and the sample mean of 5.23 is considerably higher than that.

Exercise

A simple random sample of 106 body temperature with a mean of 98.20 °F. Assume that σ is known to be 0.62 °F. Use a 0.05 significance level to test the claim that the mean body temperature of the population is equal to 98.6 °F, as is commonly believed. Is there sufficient evidence to conclude that the common belief is wrong?

Solution



Original claim: μ = 98.6 °F.

$H_0:\ \mu = 98.6°F$

$H_1:\ \mu \neq 98.6°F$

Given: $\alpha = 0.05$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250

Critical value:

$$z = \pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$$

$$z_{\bar{x}} = \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$= \frac{98.20-98.6}{\frac{0.62}{\sqrt{106}}}$$

$(98.2-98.6)/\left(.62/\sqrt{106}\right)$

$$\underline{=-6.64}$$

$$P-value = 2P(z<-6.64)$$

$$= 2(0.0001)$$

$$\underline{=0.0002}$$

Conclusion

Reject H_0; there is sufficient evidence to reject the claim that $\mu = 98.6$ and conclude that $\mu \neq 98.6$ (in fact, that $\mu < 98.6$). There is sufficient evidence to reject the claim that the mean body temperature of the population is 98.6°F. Yes; it appears there is sufficient evidence to conclude that the common belief is wrong.

Exercise

When 40 people used the Weight Watchers diet for one year, their mean weight loss was 3.0 lb. Assume that the standard deviation of all such weight changes is $\sigma = 4.9$ lb. and use a 0.01 significance level to test the claim that the mean weight loss is greater than 0. Based on these results, does the diet appear to be effective? Does the diet appear to have a practical significance?

Solution

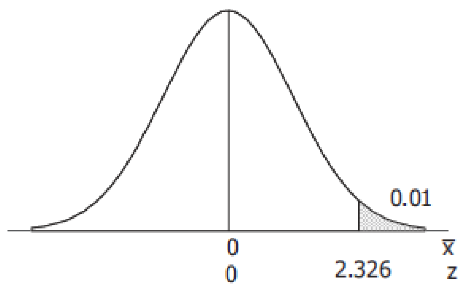
Original claim: $\mu > 0$ lb.

$H_0 : \mu = 0\ lb$

$H_1 : \mu > 0\ lb$

Given: $\alpha = 0.01$

Critical value:



$$z = z_\alpha = z_{0.01} = 2.326$$

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}} = \frac{3.0 - 0}{\frac{4.9}{\sqrt{40}}} = 3.87$$

$30 / \left(4.9 / \sqrt{40}\right)$

$$P - value = P\left(z > 3.87\right) = 1 - 0.9999 = 0.0001$$

Conclusion

Reject H_0; there is sufficient evidence to reject the claim that $\mu > 0$. There is sufficient evidence to support the claim that the mean weight lost is greater than 0. The diet is effective in that the weight loss is statistically significant – but a mere 3.0 lbs. weight loss after following the regimen for an entire year suggests the diet may have no practical significance.

Exercise

The health of the bear population in Yellowstone National Park is monitored by periodic measurements taken from anesthetized bears. A sample of 54 bears has a mean weight of 182.9 lb. Assuming that σ is known to be 121.8 lb. use a 0.05 significance level to test the claim that the population mean of all such bear weights is greater than 150 lb.

Solution

Original claim: $\mu > 150$ lbs.

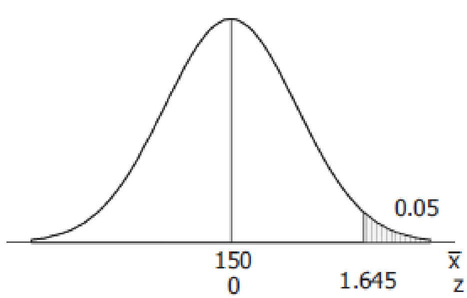
$H_0 : \mu = 150\ lbs$

$H_1 : \mu > 150\ lbs$

Given: $\alpha = 0.05$

Critical value:

z score	Area
1.645	0.9500



$$z = z_\alpha = z_{0.05} = 1.645$$

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}} = \frac{182.9 - 150}{\frac{121.8}{\sqrt{54}}} = 1.98$$

$(182.9 - 150) / \left(121.8 / \sqrt{54}\right)$

$$P - value = P(z > 1.98) = 1 - 0.9761 = 0.0239$$

Conclusion

Reject H_0; there is sufficient evidence to conclude that $\mu > 150$. There is sufficient evidence to support the claim that the mean weight of all such bears is greater than 150 lbs.

Exercise

A simple random sample of 401 salaries of NCAA football coaches in the NCAA has a mean of \$415,953. The standard deviation of all salaries of NCAA football coaches is \$463,364. Use a 0.05 significance level to test the claim that the mean salary of a football coach in the NCAA is less than \$500,000.

Solution

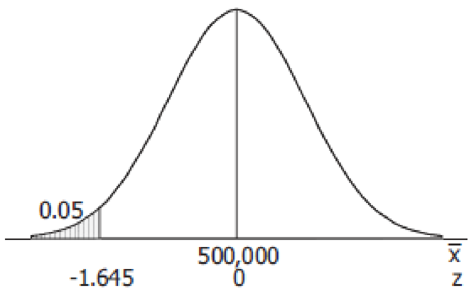
Original claim: μ < \$500,000.

$H_0: \mu = \$500,000$

$H_1: \mu < \$500,000$

Given: $\alpha = 0.05$

Critical value: $z = -z_{\alpha} = -z_{0.05} = -1.645$



$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}} = \frac{415,953 - 500,000}{\frac{463,364}{\sqrt{40}}} = -1.15$$

$(415,953 - 500,000) / \left(463,364 / \sqrt{40}\right)$

$$P - value = P(z < -1.15) = 0.1251$$

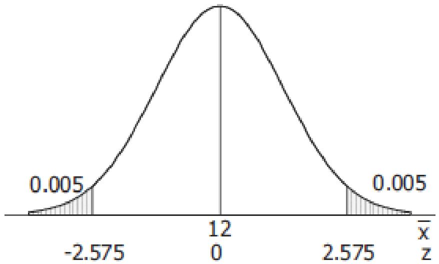
Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $\mu < 500,000$. There is not sufficient evidence to support the claim that the mean salary of an NCAA football coach is less than \$500,000.

Exercise

A simple random sample of 36 cans of regular Coke has a mean volume of 12.19 oz. Assume that the standard deviation of all cans of regular Coke is 0.11 oz. Use a 0.01 significance level to test the claim that cans if regular Coke have volumes with a mean of 12 oz., as stated on the label. If there is a difference, is it substantial?

Solution



Original claim: μ < 12 oz.

$H_0: \ \mu = 12 \ oz$

$H_1: \ \mu \neq 12 \ oz$

Given: $\alpha = 0.01$

Critical value: $z = \pm z_{\alpha/2} = \pm z_{0.025} = \pm 2.575$

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}} = \frac{12.19 - 12}{\frac{0.11}{\sqrt{36}}} = 10.36$$

$$(415{,}953 - 500{,}000) / \left(463{,}364 / \sqrt{40}\right)$$

$$P - value = 2 \cdot P(z > 10.36) = 2(1 - 0.999) = 0.0002$$

Conclusion

Reject H_0; there is sufficient evidence to reject the claim that μ = 12 and conclude that $\mu \neq 12$ (in fact, that $\mu < 12$). There is sufficient evidence to reject the claim that cans of regular Coke have a mean volume of 12 oz. The difference is statistically significant, but the difference is of practical significance only in that it guarantees that virtually 100% of the product meets the volume stated on the label.

Exercise

A simple random sample of FICO credit rating scores is obtained, and the scores are listed below.

714	751	664	789	818	779	698	836	753	834	693	802

As the writing, the mean FICO score was reported to be 678. Assuming the standard deviation of all FICO scores is known to be 58.3, use a 0.05 significance level to test the claim that these sample FICO scores come from a population with a mean equal to 678.
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

Solution

$$n = 12; \quad \sum x = 9131 \quad \sum x^2 = 6{,}985.297 \quad \overline{x} = \frac{\sum x}{n} = \frac{9131}{12} = 760.9$$

Original claim: μ < 678 FICO units.

$H_0: \ \mu = 678 \ \ FICO \ units$

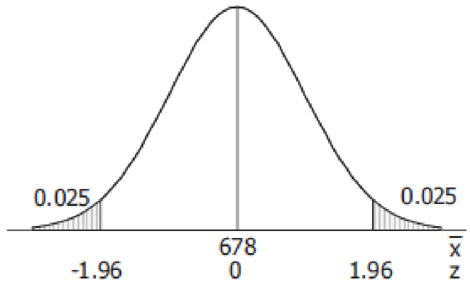
$H_1: \ \mu \neq 678 \ \ FICO \ units$

Given: $\alpha = 0.05$

Critical value: $z = \pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$

$$z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{760.9 - 678}{\frac{58.3}{\sqrt{12}}} \underline{= 4.93|}$$

$$P - value = 2 \cdot P(z > 4.93) = 2(1 - 0.9999) \underline{= 0.0002|}$$



Conclusion

Reject H_0; there is sufficient evidence to reject the claim that $\mu = 678$ and conclude that $\mu \neq 678$ (in fact, that $\mu > 678$). There is sufficient evidence to reject the claim that these FICO scores come from a population with a mean equal to 678.

Exercise

Listed below are recorded speeds (in mi/h) of randomly selected cars traveling on a section of Highway 405 in Los Angeles.

68	68	72	73	65	74	73	72	68	65	65	73	66	71	68	74	66	71	65	73
59	75	70	56	66	75	68	75	62	72	60	73	61	75	58	74	60	73	58	75

That part of the highway has posted speed limit of 65 mi/h. Assume that the standard deviation od speeds is 5.7 mi/h and use a 0.01 significance level to test the claim that the sample is from a population with a mean that is greater than 65 mi/h.

Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

Solution

$n = 40; \quad \sum x = 2735 \quad \sum x^2 = 188,259 \quad \bar{x} = \frac{\sum x}{n} = \frac{2735}{40} = 68.375$

Original claim: $\mu > 65$ mph..

$H_0: \mu = 65\ mph$

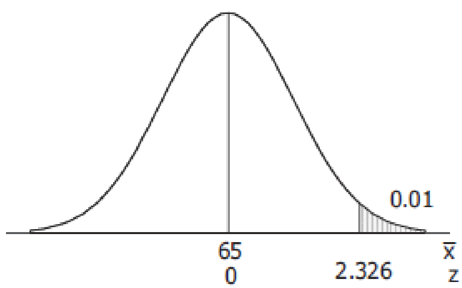
$H_1: \mu > 65\ mph$

Given: $\alpha = 0.01$

Critical value: $z = z_{\alpha} = z_{0.01} = 2.326$

$$z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{68.375 - 65}{\frac{5.7}{\sqrt{40}}} \underline{= 3.74|}$$

$$P - value = P(z > 3.74) = 1 - 0.9999 \underline{= 0.0001|}$$



Conclusion

Reject H_0; there is sufficient evidence to conclude that $\mu > 65$. There is sufficient evidence to support the claim that the sample is from a population with a mean greater than 65 mph.

Exercise

Assume the resting metabolic rate (RMR) of healthy males in complete silence is 5710 kJ/day. Researchers measured the RMR of 45 healthy males who were listening to calm classical music and found their mean RMR to be 5708.07 with a standard deviation of 992.05.

At the $\alpha = 0.05$ level of significance, is there evidence to conclude that the mean RMR of males listening to calm classical music is different than 5710 kJ/day?

Solution

We assume that the RMR of healthy males is 5710 kJ/day. This is a two-tailed test since we are interested in determining whether the RMR differs from 5710 kJ/day.

Since the sample size is large, we follow the steps for testing hypotheses about a population mean for large samples.

Step 1: $H_0 : \mu = 5710 \quad vs. \quad H_1 : \mu \neq 5710$

Step 2: The level of significance is $\alpha = 0.05$.

Step 3: The sample mean is $\bar{x} = 5708.7$ and $s = 992.05$. The test statistic is

$$t_0 = \frac{5708.7 - 5710}{\frac{992.05}{\sqrt{45}}} = -0.013$$

Classical Approach

Step 4: Since this is the two-tailed test, we determine the critical value at the $\alpha = 0.05$ level of significance with $df = 45 - 1 = 44$ to be $-t_{0.025} = -2.021 \quad and \quad t_{0.025} = 2.021$

Step 5: Since the test statistic $t_0 = -0.013$, is between the critical values, we fail to reject the null hypothesis.

P-Value Approach

Step 4: Since this is the two-tailed test, the *P*–value is the area under the *t*–distribution with $df = 44$ to the left of $-t_{0.025} = -2.021$ and to the right of $t_{0.025} = 2.021$.

That is, *P*–value $= P(t < -0.013) + P(t > 0.013)$

$$= 2P(t > 0.013)$$

$$= 0.50 < P - value$$

Step 5: Since the *P*-value is greater than the level of significance (0.05 < 0.5), we fail to reject the null hypothesis.

Step 6: There is insufficient evidence at the $\alpha = 0.05$ level of significance to conclude that the mean RMR of males listening to calm classical music differs from 5710 kJ/day.

Exercise

People have died in boat accidents because an obsolete estimate of the mean weight of men was used. Using the weights of the simple random sample of men from Data Set 1 in Appendix B, we obtain these sample statistics: $n = 40$ and $\bar{x} = 172.55\ lb.$, and $\sigma = 26.33\ lb$. Do not assume that the value of σ is known. Use these results to test the claim that men have a mean weight greater than 166.3 *lb*., which was the weight in the National Transportation and Safety Board's recommendation M-04-04. Use a 0.05 significance level, and the traditional method.

Solution

Requirements are satisfied: simple random sample, population standard deviation is not known, sample size is 40 ($n > 30$)

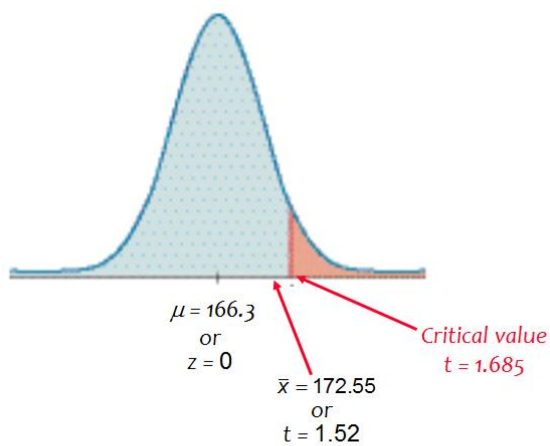
Step 1: Express claim as $\mu > 166.3\ lb.$

Step 2: Alternative to claim is $\mu \le 166.3\ lb.$

Step 3: $\mu > 166.3\ lb.$ does not contain equality, it is the alternative hypothesis:

$H_0: \mu = 166.3\ lb.$ null hypothesis

$H_1: \mu > 166.3\ lb.$ alternative hypothesis and original claim



Step 4: Significance level is $\alpha = 0.05$

Step 5: Claim is about the population mean, so the relevant statistic is the sample mean, 172.55 lb.

Step 6: Calculate t

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{172.55 - 166.3}{\frac{26.33}{\sqrt{40}}} = 1.501$$

df = $n - 1 = 39$, area of 0.05, one-tail yields $t = 1.685$;

Step 7: $t = 1.501$ does not fall in the critical region bounded by $t = 1.685$, we fail to reject the null hypothesis.

✓ Because we fail to reject the null hypothesis, we conclude that there is not sufficient evidence to support a conclusion that the population mean is greater than 166.3 lb, as in the National Transportation and Safety Board's recommendation.

Exercise

Given a simple random sample of speeds of cars on Highway in CA, you want to test the claim that the sample that the sample values are from a population with a mean greater than the posted speed limit of 65 mi/hr. Is it necessary to determine whether the sample is from a normally distributed population? If so, what methods can be used to make that determination?

Solution

Yes, since $n \le 30$, the sample should be from a population that is approximately normally distributed. We consider the normality requirement to be satisfied for such data if there are no outliers and the histogram of the sample data is approximately bell-shaped. More formally, a normal quantile plot could be used to determine whether the sample data are approximately normally distributed.

Exercise

In statistics, what does ***df*** denote. If a simple random sample of 20 speeds of cars is to be used to test the claim that the sample values are from a population with a mean greater than the posted speed limit of 65 mi/h, what is the specific value of ***df***?

Solution

In statistics, ***df*** denotes the degrees of freedom. In general, the degrees of freedom give the number of pieces of information that are free to vary without changing the mathematical constraint of the problem. When using a *t* test with $n = 20$ sample values to test a claim about the mean of a population, $df = 19$

Exercise

Claim about IQ scores of statistics instructors: $\mu > 100$, sample data: $n = 15, \quad \bar{x} = 118, \quad s = 11$. The sample data appear to come from a normally distributed population with unknown μ and σ. Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

Solution

Use *t*. When σ is unknown and the *x*'s approximately normally distributed, use *t*.

Exercise

Claim about FICO credit scores of adults: $\mu = 678$, sample data: $n = 12, \quad \bar{x} = 719, \quad s = 92$. The sample data appear to come from a population with a distribution that is not normal, and σ Is unknown Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

Solution

Neither the *z* nor the *t* applies. When σ is unknown and the *x*'s are not normally distributed sample sizes $n \le 30$ cannot be used with these techniques.

Exercise

Claim about daily rainfall amounts in Boston: $\mu < 0.20\,in$, sample data: $n = 19, \bar{x} = 0.10\,in, s = 0.26\,in.$ The sample data appear to come from a population with a distribution that is very far from normal, and σ is unknown Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

Solution

Neither the *z* nor the *t* applies. When σ is unknown and the *x*'s are not normally distributed sample sizes $n \le 30$ cannot be used with these techniques.

Exercise

Testing a claim about the mean weight of M&M's: Right-tailed test with *n* = 25 and test statistic *t* = 0.430. Find the *P*-value and find a range of values for the *P*-value.

Solution

$$P-value = P\left(t_{24} > 0.430\right)$$

Table for area in one tail: [0.430 < 1.318] $P-value > 0.10$

TI: ***tcdf***(0.430, 99, 24) = 0.3355

Exercise

Test a claim about the mean body temperature of healthy adults: left-tailed test with *n* = 11 and test statistic *t* = –3.158. Find the *P*-value and find a range of values for the *P*-value.

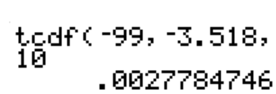
Solution

$$P-value = P\left(t_{10} < -3.518\right)$$

Table for area in one tail: [–3.518 < –3.169]

$P-value < 0.01$

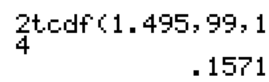
TI: ***tcdf***(–99, –3.518, 10) = 0.0027



Exercise

Two-tailed test with n = 15 and test statistic t = 1.495. Find the P-value and find a range of values for the P-value.

Solution



$$P-value = 2 \cdot P\left(t_{14} < 1.495\right)$$

Table for area in two tails: [1.345 < 1.495 < 1.761]

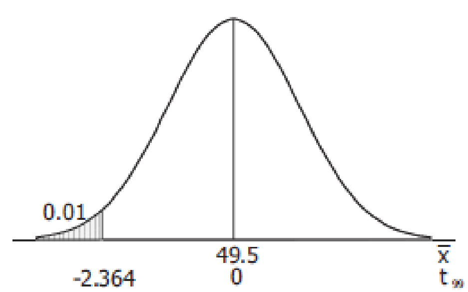
$$P-value < 0.20$$

TI: 2·***tcdf***(1.495, 99, 14) = 0.1571

Exercise

In an analysis investigating the usefulness of pennies, the cents portions of 100 randomly selected checks are recorded. The sample has mean of 23.8 cents and a standard deviation of 32.0 cents. If the amounts from 0 cents to 99 cents are all equally likely, the mean is expected to be 49.5 cents. Use a 0.01 significance level to test the claim that the sample is from a population with a mean less than 49.5 cents. What does the result suggest about the cents portions of the checks?

Solution



Original claim: μ = 49.5 cents.

$H_0 : \ \mu = 49.5 \ cents$

$H_1 : \ \mu < 49.5 \ cents$

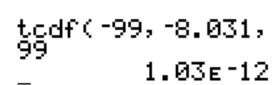
$Given: \ \alpha = 0.01 \quad and \quad df = 99$

Critical value: $t = -t_{\alpha} = -t_{0.01} = -2.364$

$$t_{\overline{x}} = \frac{\overline{x} - \mu}{s_{\overline{x}}} = \frac{23.8 - 49.5}{\frac{32}{\sqrt{100}}} = -8.031$$

$(23.8 - 49.5) / \left(32 / \sqrt{100}\right)$

$$P-value = 2 \cdot P\left(t_{99} < -8.031\right) = 2 \cdot tcdf\ (-99,\ -8.031,\ 99)\ = 2.06\ E-12$$



Conclusion

Reject H_0; there is sufficient evidence to conclude that $\mu < 49.5$. There is sufficient evidence to that the cents portion of all checks has a mean that is less than 49.5 cents. The results suggest that the cents portions of checks are not uniformly distributed from 0 to 99 cents.

Exercise

A simple random sample of 40 recorded speeds (in mi/h) is obtained from cars traveling on a section of Highway 405 in Los Angeles. The sample has a mean of 68.4 mi/h and a standard deviation of 5.7 mi/h.

Use a 0.05 significance level to test the claim that the mean speed of all cars is greater than the posted speed limit of 65 mi/h.

Solution

$n = 40;\quad \sum x = 2735 \quad \sum x^2 = 188{,}259 \quad \bar{x} = \frac{\sum x}{n} = \frac{2735}{40} = 68.375$

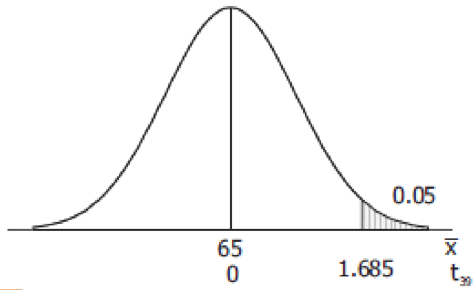
Original claim: $\mu > 65$ mph.

$H_0:\ \mu = 65\ mph$

$H_1:\ \mu > 65\ mph$

Given: $\alpha = 0.05$ *and* $df = 39$

Critical value: $t = t_\alpha = t_{0.05} = 1.685$



Degrees of Freedom	*t* Distribution: Critical *t* Values				
	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
39	**2.708**	2.426	**2.023**	1.685	1.304

$$t_{\bar{x}} = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{68.4 - 65}{\frac{5.7}{\sqrt{40}}} = \underline{3.773}$$

tcdf(3.773,99,39
2.681717486E-4

$$P - value = P\left(t_{39} > 3.7763\right) = tcdf\left(3.773,\ 99,\ 39\right) = \underline{0.0003}$$

Conclusion

Reject H_0; there is sufficient evidence to conclude that $\mu > 65$. There is sufficient evidence to support the claim that the mean speed of all such cars is greater than the posted speed of 65 mph.

Exercise

The heights are measured for the simple random sample of supermodels. They have mean height of 70.0 in. and a standard deviation of 1.5 in. Use a 0.01 significance level to test the claim that supermodels have heights with a mean that is greater than the mean heights of 63.6 in. for women in general population. Given that there are only nine heights represented, can we really conclude that supermodels are taller than the typical woman?

Solution

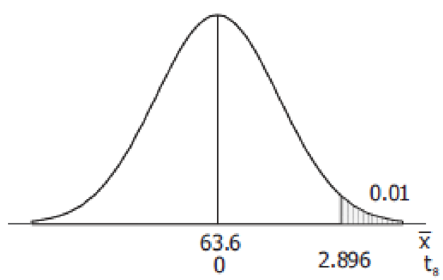
Original claim: $\mu > 63.6$ in.

$H_0:\ \mu = 63.6\ in.$

$H_1:\ \mu > 63.6\ in.$

Given: $\alpha = 0.01$ *and* $df = 9 - 1 = 8$

Critical value: $t = t_\alpha = t_{0.01} = 2.896$



Degrees of Freedom	*t* Distribution: Critical *t* Values				
	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
8	**3.355**	2.896	**2.306**	1.860	1.397

$$t_{\overline{x}} = \frac{\overline{x}-\mu}{s_{\overline{x}}} = \frac{70-63.6}{\frac{1.5}{\sqrt{9}}} = 12.80$$

$$P-value = P\left(t_8 > 12.8\right) = tcdf\,(12.8,\ 99,\ 8) = 6.5465E\text{ - }7 \approx 0.0000007$$

Conclusion

Reject H_0; there is sufficient evidence to conclude that $\mu > 63.6$. There is sufficient evidence to support the claim that supermodels have a mean height that is greater than the 63.6 in. of the general population of women. Yes; assuming that the heights of supermodels are approximately normally distributed around their mean, the test and the conclusion are valid.

Exercise

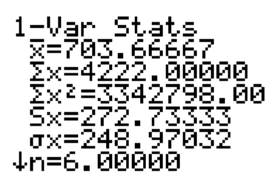
The National Highway Traffic Safety Administration conducted crash tests of child booster seats for cars. Listed below are results from those tests, with measurement given in hic (standard *head injury condition* units). The safety requirement is that the hic measurement should be less than 1000 hic. Use a 0.01 significance level to test the claim that the sample is from a population with a mean less than 1000 hic.

774 649 1210 546 431 612

Do the results suggest that all of the child booster seats meet the specified requirement?

Solution

$n = 6;\quad \sum x = 4222 \quad \sum x^2 = 3{,}342{,}798 \quad \overline{x} = 703.67 \quad s = 272.73$



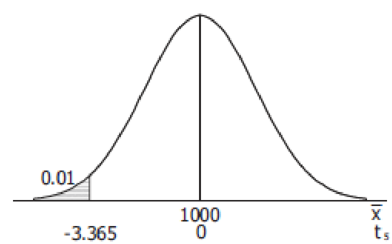
Original claim: $\mu < 1000$ hic.

$H_0:\ \mu = 1000\ hic.$

$H_1:\ \mu < 1000\ hic.$

Given: $\alpha = 0.01$ *and* $df = 6-1 = 5$

Critical value: $t = -t_{\alpha} = -t_{0.01} = -3.365$



***t* Distribution: Critical *t* Values**

Degrees of Freedom	Area in One Tail 0.005	0.01	0.025	0.05	0.10
5	**4.032**	3.365	**2.571**	2.015	1.476

$$t_{\overline{x}} = \frac{\overline{x}-\mu}{s_{\overline{x}}} = \frac{703.67-1000}{\frac{272.73}{\sqrt{6}}} = -2.661$$

$$P-value = P\left(t_5 < -2.661\right) = tcdf\,(-99, -2.661, 5) = 0.0224$$

Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $\mu < 1000$. There is not sufficient evidence to support the claim that the population mean is less than 1000 hic. No; since one of the sample values is 1210, there is proof that not all of the child booster seats meet the specified requirement.

Exercise

The trend of thinner Miss America winners has generated charges that the contest encourages unhealthy diet habits among young women. Listed below are body mass indexes (BMI) for recent Miss America winners. Use a 0.01 significance level to test the claim that recent Miss America winners are from a population with a mean BMI less than 20.16, which was the BMI for winners from the 1920s and 1930s.

19.5 20.3 19.6 20.2 17.8 17.9 19.1 18.8 17.6 16.8

Do recent winners appear to be significantly different from those in the 1920s and 1930s?

Solution

$n=10; \quad \sum x = 187.6 \quad \bar{x} = 18.76 \quad s = 1.1862$

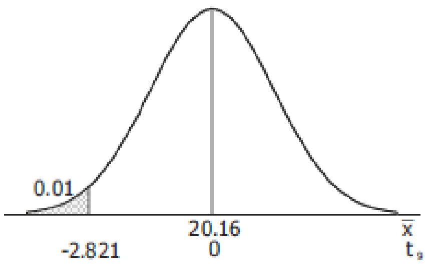
Original claim: $\mu < 20.16$.

$H_0: \ \mu = 20.16$

$H_1: \ \mu < 20.16$

Given: $\alpha = 0.01$ *and* $df = 9$

Critical value: $t = -t_{\alpha} = -t_{0.01} = -2.821$



Degrees of Freedom	t Distribution: Critical t Values				
	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
9	3.250	2.821	2.262	1.833	1.383

$$t_{\bar{x}} = \frac{\bar{x}-\mu}{s_{\bar{x}}} = \frac{18.76-20.16}{\frac{1.1862}{\sqrt{10}}} = -3.732$$

$$P-value = P\left(t_9 < -3.732\right) = tcdf\left(-99, -3.732, 9\right) = 0.0023$$

Conclusion

Reject H_0; there is sufficient evidence to conclude that $\mu < 20.16$. There is sufficient evidence to support the claim that the recent winners are from a population with BMI less than 20.16, which was the BMI for winners in the 1920's and 1930's.

Yes; recent winners appear to be significantly different from those in the earlier years.

Exercise

The list measured voltage amounts supplied directly to the author's home

123.8	123.9	123.9	123.3	123.4	123.3	123.3	123.6	123.5	123.5	123.5	123.7
123.6	123.7	123.9	124.0	124.2	123.9	123.8	123.8	124.0	123.9	123.6	123.5
123.4	123.4	123.4	123.4	123.3	123.3	123.5	123.6	123.8	123.9	123.9	123.8
123.9	123.7	123.8	123.8								

The Central Hudson power supply company states that it has a target power supply of 120 volts. Using those home voltage amounts, test the claim that the mean is 120 volts. Use a 0.01 significance level.

Solution

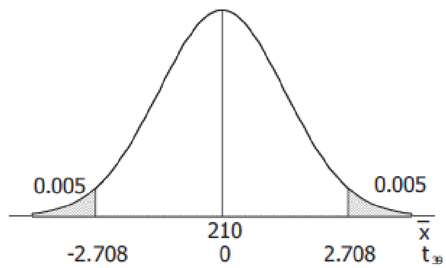
$n = 40; \quad \overline{x} = 123.6625 \quad s = 0.24039$

Original claim: μ = 120 volts.

$H_0 : \ \mu = 120 \ volts$

$H_1 : \ \mu \neq 120 \ volts$

Given: $\alpha = 0.01 \quad and \quad df = 39$



Critical value: $t = \pm t_{\alpha/2} = \pm t_{0.005} = \pm 2.708$

$$t_{\overline{x}} = \frac{\overline{x} - \mu}{s_{\overline{x}}} = \frac{123.6625 - 120}{\frac{120}{\sqrt{40}}} \underline{= 96.359}$$

$$P - value = 2 \cdot P\left(t_{39} > 96.359\right) = 2 \cdot tcdf\,(96.359, 999, 39) = 5.284E - 48 \ \underline{\approx 0}$$

Conclusion

Reject H_0; there is sufficient evidence to conclude that μ = 120 and conclude that $\mu \neq 120$ (in fact the $\mu > 120$). There is sufficient evidence to reject the claim that the mean home voltage amount is 120 volts.

Exercise

When testing a claim about a population mean with a simple random sample selected from a normally distributed population with unknown σ, the student t distribution should be used for finding critical values and/or a P–value. If the standard normal distribution is incorrectly used instead, does that mistake make you more or less likely to reject the null hypothesis, or does it not make a difference? Explain.

Solution

Because the *z* distribution has less spread than a *t* distribution, z_α is less than t_α for any α. This makes the critical *z* value smaller (closer to 0) than the corresponding critical *t* value, which means that rejection is more likely with *z* than with *t*.

Exercise

The list measured human body temperature.

98.6	98.6	98.0	98.0	99.0	98.4	98.4	98.4	98.4	98.6	98.6	98.8	98.6	97.0	97.0	97.0
98.8	97.6	97.7	98.8	98.0	98.0	98.3	98.5	97.3	98.7	97.4	98.9	98.6	99.5	97.5	98.0
97.3	97.6	98.2	99.6	98.7	99.4	98.2	98.0	98.6	98.6	97.2	98.4	98.6	98.2	98.0	97.4
97.8	98.0	98.4	98.6	98.6	97.8	99.0	96.5	97.6	98.0	96.9	97.6	97.1	97.9	98.4	98.4
97.3	98.0	97.5	97.6	98.2	98.5	98.8	98.7	97.8	98.0	97.1	97.4	99.4	98.4	98.6	97.8
98.4	98.5	98.6	98.3	98.7	98.8	99.1	98.6	97.9	98.8	98.0	98.7	98.5	98.9	98.4	98.4
98.6	97.1	97.9	98.8	98.7	97.6	98.2	99.2	97.8	98.0						

Use the temperatures listed for 12 AM on day 2 to test the common belief that the mean body temperature is 98.6 °F. Does that common belief appear to be wrong?

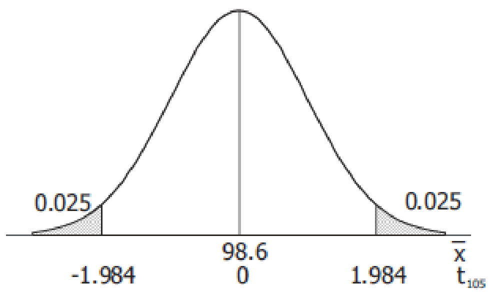
Solution

$n = 106; \quad \overline{x} = 98.2 \quad s = 0.6229$

Original claim: $H_0: \ \mu = 98.6\ °F$ F.

$H_1: \ \mu \ne 98.6\ °F$

Given: $\alpha = 0.05$ (*assume*) *and* $df = 105$



Critical value: $t = \pm t_{\alpha/2} = \pm t_{0.025} = \pm 1.984$

$$t_{\overline{x}} = \frac{\overline{x} - \mu}{s_{\overline{x}}} = \frac{98.2 - 98.6}{\frac{0.6229}{\sqrt{106}}} = -6.611$$

$$P - value = 2 \cdot P\left(t_{105} < -6.611\right) = 2 \cdot tcdf\left(-99, -6.611, 105\right) = 1.627E-9 \approx 0.0000002$$

Conclusion

Reject H_0; there is sufficient evidence to conclude that $\mu = 98.6$ and conclude that $\mu \ne 98.6$ (in fact the $\mu < 98.6$). There is sufficient evidence to reject the claim that the mean body temperature of the population is 98.6 °F. Yes; it appears there is sufficient evidence to conclude that the common belief is wrong.

Exercise

Determine whether the hypothesis test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither

a) Claim $\mu = 981$. Sample data: $n = 20, \quad \overline{x} = 946, \quad s = 27$. The sample data appear to come from a normally distributed population with $\sigma = 30$.

b) Claim $\mu = 105$. Sample data: $n = 16, \quad \overline{x} = 101, \quad s = 15.1$. The sample data appear to come from a normally distributed population with unknown μ *and* σ.

Solution

a) Normal

b) Student *t*. (σ in unknown and the *x*'s approximately normally distributed, use *t*.

Solution Section 3.8 – Hypothesis Tests for a Population Standard Deviation

Exercise

There is a claim that the lengths of men's hands have a standard deviation less than 200 mm. You plan to test that claim with a 0.01 significance level by constructing a confidence interval. What level of confidence should be used for the confidence interval? Will the conclusion based on the confidence interval be the same as the conclusion based on a hypothesis test that uses the traditional method or the *P*-value method?

Solution

A one-tailed test at the 0.01 level of significance rejects the null hypothesis if the sample statistic falls into the extreme 1% of the sampling distribution in the appropriate tail. The corresponding (two-sided) confidence interval test that places 1% each tail would be a 98% confidence interval. When testing claims about a standard deviation, the confidence interval method gives the same results as test using the traditional method or the *P*-value method.

Exercise

There is a claim that daily rainfall amounts in Boston have a standard deviation equal to 0.25 in. Sample data show that daily rainfall amounts are from a population with a distribution that is very far from normal. Can the use of a very large sample compensate for the lack of normality, so that the methods of this section can be used for the hypothesis test?

Solution

No. Unlike tests and confidence intervals involving the mean, which do not require normality when $n > 30$, test and confidence intervals involving standard deviations require approximate normality for all sample sizes.

Exercise

There is a claim that men have foot breaths with a variance equal to 36 mm^2. Is a hypothesis test of the claim that the variance is equal to 36 mm^2 equivalent to a test of the claim that the standard deviation is equal to 6 *mm*.

Solution

Yes. The claim that the variance is equal to 36 mm^2 and the claim that the standard deviation is equal to 6 *mm* are equivalent claims, and their corresponding tests are equivalent.

Exercise

Given: $H_1 : \sigma \neq 696\ g, \quad \alpha = 0.05, \quad n = 25, \quad s = 645\ g$, Find

a) Find the test statistic

b) Find critical value(s)

c) Find *P*-value limits

d) Determine whether there is sufficient evidence to support the given alternative hypothesis.

Solution

a) Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(24)(645)^2}{(696)^2} \underline{= 20.612|}$

b) Critical values for $\alpha = 0.05 \quad and \quad df = 24$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559

$\chi^2 = \chi^2_{1-\alpha/2} = \chi^2_{0.975} \underline{= 12.401|}$

$\chi^2 = \chi^2_{\alpha/2} = \chi^2_{0.025} \underline{= 39.364|}$

c) *P*-value limits: 15.659 < 20.612 (part *a*)

P-value > 0.20

P-value exact: $2 \bullet \chi^2\ cdf\ (0,\ 20.612,\ 24) \underline{= 0.6770|}$

d) Conclusion: Do not reject H_0 ; there is not sufficient to conclude that $\sigma \neq 696$

Exercise

Given: $H_1 : \sigma < 29\ lb, \quad \alpha = 0.05, \quad n = 8, \quad s = 7.5\ lb$, Find

a) Find the test statistic

b) Find critical value(s)

c) Find *P*-value limits

d) Determine whether there is sufficient evidence to support the given alternative hypothesis.

Solution

a) Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(7)(7.5)^2}{(29)^2} \underline{= 0.468|}$

b) Critical values for $\alpha = 0.05 \quad and \quad df = 7$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278

$\chi^2 = \chi^2_{1-\alpha} = \chi^2_{0.95} \underline{= 2.167|}$

c) *P*-value limits: (part *a*) 0.468 < 0.989

P-value < 0.005

P-value exact: $\chi^2\ cdf\,(0,\ 0.468,\ 7) = 4.451E-4\ \underline{=0.0004}$

d) Conclusion: Reject H_0; there is sufficient to conclude that $\sigma < 29$

Exercise

Given: $H_1:\ \sigma > 3.5\ min, \quad \alpha = 0.01, \quad n = 15, \quad s = 4.8\ min$, Find

a) Find the test statistic

b) Find critical value(s)

c) Find *P*-value limits

d) Determine whether there is sufficient evidence to support the given alternative hypothesis.

Solution

a) Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(14)(4.8)^2}{(3.5)^2}\ \underline{=26.331}$

b) Critical values for $\alpha = 0.01 \quad and \quad df = 14$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319

$\chi^2 = \chi^2_\alpha = \chi^2_{0.01}\ \underline{=29.141}$

c) *P*-value limits: 26.119 < 26.331 < 29.141

0.01 < *P*-value < 0.025

P-value exact: $\chi^2\ cdf\,(26,\ 999,\ 14)\ \underline{=0.0235}$

d) Conclusion: Do not reject H_0; there is not sufficient evidence to conclude that $\sigma > 3.5$

Exercise

Given: $H_1:\ \sigma \neq 0.25, \quad \alpha = 0.01, \quad n = 26, \quad s = 0.18$, Find

a) Find the test statistic

b) Find critical value(s)

c) Find *P*-value limits

d) Determine whether there is sufficient evidence to support the given alternative hypothesis.

Solution

a) Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(25)(0.18)^2}{(0.25)^2}\ \underline{=12.960}$

b) Critical values for $\alpha = 0.01 \quad and \quad df = 25$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928

$$\chi^2 = \chi^2_{1-\alpha/2} = \chi^2_{0.995} \underline{=10.520|}$$

$$\chi^2 = \chi^2_{\alpha/2} = \chi^2_{0.005} \underline{=46.928|}$$

c) P-value limits: $12.198 < 12.960 < 13.844$

$0.02 < P\text{-value} < 0.05$

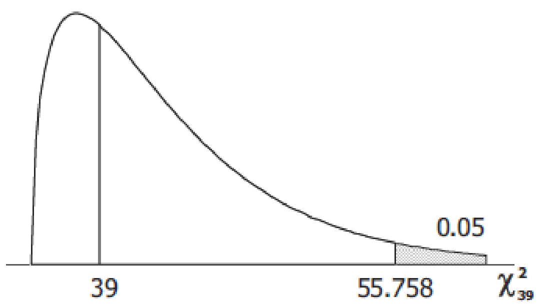
P-value exact: $2 \cdot \chi^2\ cdf\,(0,\ 12.960,\ 25)\ \underline{=0.0460|}$

d) Conclusion: Do not reject H_0; there is not sufficient to conclude that $\sigma \neq 0.25$

Exercise

A simple random sample of 40 men results in a standard deviation of 11.3 beats per minute. The normal range of pulse rates of adults is typically given as 60 to 100 beats per minute. If the range rule of thumb is applied to that normal range, the result is a standard deviation of 10 beats per minute. Use the sample results with a 0.05 significance level to test the claim that rates of men have a standard deviation greater than 10 beats per minute.

Solution



Original claim: $\sigma > 10$ *beats/min*

$H_0: \sigma = 10$ *beats/min*

$H_1: \sigma > 10$ *beats/min*

Given: $\alpha = 0.05$ *and* $df = 39$

Critical value: $\chi^2 = \chi^2_{\alpha} = \chi^2_{0.05}\ \underline{=55.758|}$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(39)(11.3)^2}{(10)^2}\ \underline{=49.799|}$$

P-value exact: $\chi^2\ cdf\,(49.799,\ 999,\ 39)\ \underline{=0.1152|}$

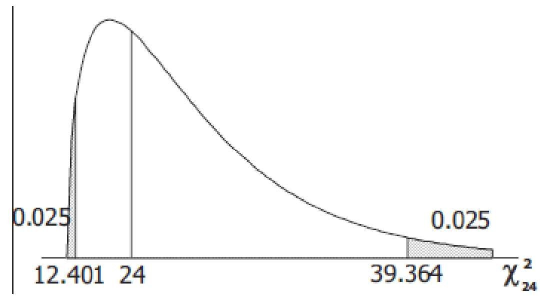
Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $\sigma > 10$. There is not sufficient evidence to support the claim that the pulse rate of men have a standard deviation greater than 10 *beats/min*.

Exercise

A simple random sample of 25 filtered 100 mm cigarettes is obtained, and the tar content of each cigarette is measured. The sample has a standard deviation of 3.7 mg. Use a 0.05 significance level to test the claim that the tar content of filtered 100 mm cigarettes has a standard deviation different from 3.2 mg, which is the standard deviation for unfiltered king size cigarettes.

Solution



Original claim: $\sigma \neq 3.2\ mg$

$H_0:\ \sigma = 3.2\ mg$

$H_1:\ \sigma \neq 3.2\ mg$

Given: $\alpha = 0.05 \quad and \quad df = 24$

Critical value:

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559

$$\chi^2 = \chi^2_{1-\alpha/2} = \chi^2_{0.975} = 12.401$$

$$\chi^2 = \chi^2_{\alpha/2} = \chi^2_{0.025} = 39.364$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(24)(3.7)^2}{(3.2)^2} = 32.086$$

P-value exact: $2 \cdot \chi^2\ cdf\,(32.086,\ 999,\ 24) = 0.2498$

Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $\sigma \neq 3.2$. There is not sufficient evidence to support the claim that the tar content of such cigarettes has a standard deviation different from 3.2 *mg*.

Exercise

When 40 people used the Weight Watchers diet for one year, their weight losses had a standard deviation of 4.9 lb. Use 0.01 significance level to test the claim that the amounts of weight loss have a standard deviation equal to 6.0 lb., which appears to be the standard deviation for the amounts of weight loss with the Zone diet.

Solution

Original claim: $\sigma = 6.2\ lbs$

$H_0:\ \sigma = 6.2\ lbs$

$H_1:\ \sigma \neq 6.2\ lbs$

$Given$: $\alpha = 0.01 \quad and \quad df = 39$

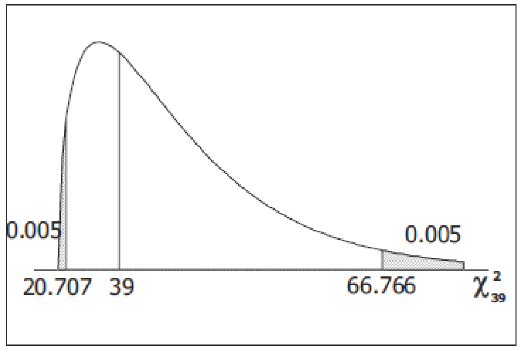
Critical value:

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766

$$\chi^2 = \chi^2_{1-\alpha/2} = \chi^2_{0.995} \underline{= 20.707}$$

$$\chi^2 = \chi^2_{\alpha/2} = \chi^2_{0.005} \underline{= 66.766}$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(39)(4.9)^2}{(6.0)^2} \underline{= 26.011}$$



P-value exact:

$$2 \cdot \chi^2\ cdf\,(0,\ 26.011,\ 39) \underline{= 0.1101}$$

Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $\sigma = 6.0$. There is not sufficient evidence to support the claim that the weight losses from this diet have a standard deviation different of 6.0 *lbs*.

Exercise

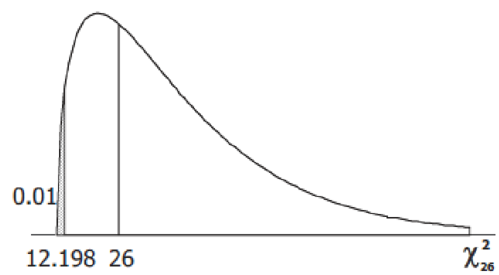
Tests in the statistic classes have scores with a standard deviation equal to 14.1. One of the last classes has 27 test scores with a standard deviation of 9.3. Use a 0.01 significance level to test the claim that this class has less variation than other past classes. Does a lower standard deviation suggest that this last class is doing better?

Solution

Original claim: $\sigma < 14.1$

$H_0 : \sigma = 14.1$

$H_1 : \sigma < 14.1$



$Given$: $\alpha = 0.01 \quad and \quad df = 26$

Critical value:

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290

$$\chi^2 = \chi^2_{1-\alpha} = \chi^2_{0.99} \underline{= 12.198}$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(26)(9.3)^2}{(14.1)^2} \underline{= 11.311}$$

P-value exact: $\chi^2\ cdf\left(0,\ 11.311,\ 26\right) \underline{=0.0056}$

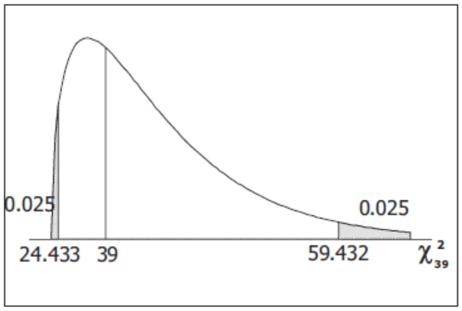
Conclusion

Reject H_0; there is sufficient evidence to conclude that $\sigma < 14.1$. There is sufficient evidence to support the claim that this class has less variation than other past classes.

No; a lower standard deviation means that the scores are closer together, but it says nothing about whether they are higher or lower.

Exercise

A simple random sample of pulse rates of 40 women results in a standard deviation of 12.5 *beats/min*. The normal range of pulse rates of adults is typically given as 60 to 100 *beats/min*. If the range rule of thumb is applied to that normal range, the result is a standard deviation of 10 *beats/min*. Use the sample results with a 0.05 significance level to test the claim that pulse rates of women have a standard deviation equal to 10 *beats/min*.



Solution

Original claim: $\sigma = 10$ *beats/min*

H_0 : $\sigma = 10$ *beats/min*

H_1 : $\sigma \neq 10$ *beats/min*

Given: $\alpha = 0.05$ *and* $df = 39$

Critical value:

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766

$$\chi^2 = \chi^2_{1-\alpha/2} = \chi^2_{0.975} \underline{=24.433}$$

$$\chi^2 = \chi^2_{\alpha/2} = \chi^2_{0.025} \underline{=59.432}$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(39)(12.5)^2}{(10)^2} \underline{=60.9375}$$

P-value exact: $2 \cdot \chi^2\ cdf\left(60.9375,\ 999,\ 39\right) \underline{=0.0277}$

Conclusion

Reject H_0; there is sufficient evidence to reject the claim that $\sigma = 10$ conclude that $\sigma \neq 10$. There is sufficient evidence to reject the claim that the pulse rates of women have a standard deviation equal to 10 *beats/min*.

Exercise

Listed below are the playing times (in seconds) of songs that were popular at the time of this writing. Use a 0.05 significance level to test the claim that the songs are from a population with a standard deviation less than on minute.

448 242 231 246 246 293 280 227 244 213 262 239 213 258 255 257

Solution

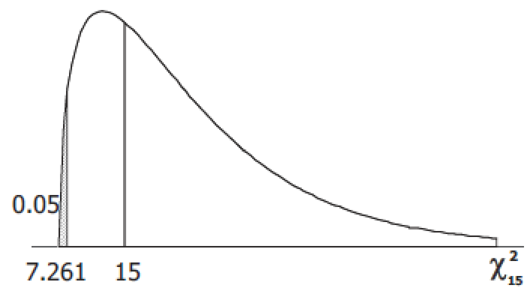
$n = 16, \quad \bar{x} = 259.3, \quad s = 54.549$

Original claim: $\sigma < 60\ sec$

$H_0 : \sigma = 60\ sec$

$H_1 : \sigma < 60\ sec$

Given: $\alpha = 0.05 \quad and \quad df = 15$



Critical value

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801

$$\chi^2 = \chi^2_{1-\alpha} = \chi^2_{0.95} \ \underline{= 7.261}$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(15)(54.549)^2}{(60)^2} \ \underline{= 12.398}$$

P-value exact: $\chi^2\ cdf\,(0,\ 12.398,\ 15)\ \underline{= 0.3513}$

Conclusion

Do not reject H_0; there is not sufficient evidence to conclude that $\sigma < 60$. There is not sufficient evidence to support the claim that the songs are from a population with a standard deviation less than one minute.

Exercise

Find the critical value or values of χ^2 based on the given information

a) $H_0: \sigma = 8.0, \quad \alpha = 0.01, \quad n = 10$

b) $H_1: \sigma > 3.5, \quad \alpha = 0.05, \quad n = 14$

c) $H_1: \sigma < 0.14, \quad \alpha = 0.10, \quad n = 23$

d) $H_1: \sigma \neq 9.3, \quad \alpha = 0.05, \quad n = 28$

Solution

Using Chi-Square $\left(\chi^2\right)$ Distribution Table

a) $\alpha = 0.01 \quad and \quad df = 10 - 1 = 9$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589

Critical value: $\chi^2_{1-\alpha/2} = \chi^2_{0.995} \underline{= 1.735|}$

$\chi^2 = \chi^2_{\alpha/2} = \chi^2_{0.005} \underline{= 23.589|}$

b) $\alpha = 0.05 \quad and \quad df = 14 - 1 = 13$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819

Since $H_1: \sigma > 3.5$, then $\chi^2 = \chi^2_{\alpha} = \chi^2_{0.05} \underline{= 22.362|}$

c) $\alpha = 0.10 \quad and \quad df = 23 - 1 = 22$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796

Since $H_1: \sigma < 0.14$, then $\chi^2 = \chi^2_{1-\alpha} = \chi^2_{0.90} \underline{= 14.042|}$

d) $\alpha = 0.05 \quad and \quad df = 28 - 1 = 27$

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963	49.645

Critical value: $\chi^2_{1-\alpha/2} = \chi^2_{0.975} \underline{= 14.573|}$

$\chi^2 = \chi^2_{\alpha/2} = \chi^2_{0.025} \underline{= 43.194|}$