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

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

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

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

Exercise

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

Solution

For $\alpha = 0.20 \quad \rightarrow \frac{\alpha}{2} = \frac{0.2}{2} = 0.1;$ 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

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

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

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

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

$$4.63 \pm 0.20$$

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

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

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

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

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

$$1522\pm 58$$

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

$$1464<\mu<1580$$

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

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

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

$$1522\pm 77$$

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

$$1445<\mu<1599$$

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

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

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

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

$$133.93 \pm 5.24$$

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

$$\underline{128.7 < \mu < 139.2}$$

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; $n = 37$, $\bar{x} = 10.4\ yrs$, $\sigma = 2.2\ yrs$

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

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

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

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

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

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

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

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

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

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

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

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

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

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*; $n = 20$, $\bar{x} = \$9004$, $s = \$569$. 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) $E = \dfrac{t_{\alpha/2}\, s}{\sqrt{n}}$

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

$\boxed{= 0.06 \ \ grams/mile}$

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

0.12 ± 0.06

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

$\boxed{0.06 < \mu < 0.18 \quad grams/mile}$

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 $df = 46\,[45]$

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

		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	$\boxed{0.025}$	0.05	0.10
$\boxed{100}$	2.626	2.364	$\boxed{1.984}$	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	$\boxed{0.025}$	0.05	0.10
$\boxed{80}$	2.639	2.374	$\boxed{1.990}$	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

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

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

$71.70 < \mu < 86.30$