Section 6.2: Calculating Large Sample Confidence Intervals for the Mean

Thursday, November 28, 2019 9:32 AM

when we talk about a confidence interval:

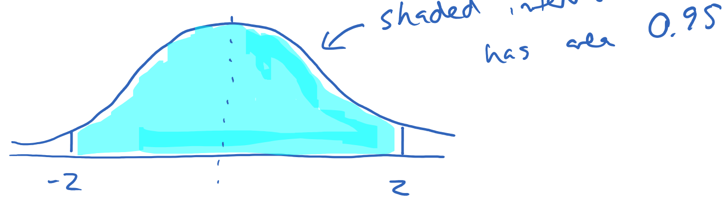
"The average temperature in downtown Victoria tomorrow will be between 12°C and 16°C with 90% confidence."

- what we mean is that we calculated the interval 12°C to 16°C using a method that gives correct results 90% of the time

in other words, 9 times out of 10, the true value will fall within that interval **and** 1 time out of 10, the prediction will be incorrect

---

the most common confidence interval is a 95% confidence interval



by symmetry, this area from 0 to z is $\frac{0.95}{2} = 0.475$

what z is that? $z = 1.96$

note: not exactly 2, but very close (Empirical Rule)

other common confidence intervals are

90%, 98%, and 99%, and they also have
z-scores associated with them

confidence interval		z-score
90%	(0.90)	1.645
95%	(0.95)	1.96 *
98%	(0.98)	2.33
99%	(0.99)	2.575

* don't use "2" from Empirical Rule,
since 1.96 is more accurate

any other confidence intervals: (92%, 96%)
- use the method above
  - divide the area by 2 and look up in standard normal table

to construct a confidence interval, you need the
following pieces of information:

- confidence level (example: 95%), which gives you the associated z-score (1.96)
- sample mean $\bar{x}$
- standard deviation σ (for large samples, can use sample std dev s instead)
- sample size n (and for the technique we're using, n must be ≥ 30)

then you plug into the following formula:

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

if σ not known, can use sample std dev s here

here

population mean we are estimating

sample mean we have measured

example: A sample of 75 five-year-old Douglas fir trees randomly chosen in BC produced a sample mean of 85 cm with a standard deviation of 12 cm for the diameter.

Estimate with 95% confidence the average diameter of five-year-old Douglas fir trees.

95% confidence $\rightarrow$ $z = 1.96$

we use s here

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

$$= 85 \pm \frac{1.96(12)}{\sqrt{75}}$$

$$= 85 \pm \underbrace{2.71586}_{\text{ridiculous number of decimal places}}$$

$$= 85 \pm 3 \quad \underline{\text{or}} \quad 85 \pm 2.7$$

how much should you round? either round to the same precision as the mean (same number of decimal places), <u>or</u> one extra decimal place confidence interval is then:

from 82 to 88 cm with 95% confidence

---

example: Forty students were asked how many hours they studied the weekend before final exams. The mean number of hours studied was 15.1 and the standard deviation was 6.5 hours.

Construct confidence intervals for the mean number of hours studied by students with

a) 90% confidence
b) 95% confidence
c) 99% confidence

What happens to the size of the interval as the confidence level increases?

a) 90% confidence : $z = 1.645$

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

$$= 15.1 \pm \frac{1.645(6.5)}{\sqrt{40}}$$

$$= 15.1 \pm 1.69063$$

$$= 15.1 \pm 1.7$$

90% CI : 13.4 to 16.8 hours

b) same calculation with 95% $\Rightarrow$ $z = 1.96$

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$
$$= 15.1 \pm 2.01437$$
$$= 15.1 \pm 2.0$$

95% CI: 13.1 to 17.1 hours

c) 99% Confidence $\Rightarrow$ $z = 2.575$

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$
$$= 15.1 \pm 2.64643$$
$$= 15.1 \pm 2.6$$

99% CI: 12.5 to 17.7 hours

as the confidence level increases, the width of the interval also increases

---

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

this entire term is called the margin of error (MOE)

---

we just saw that increasing the confidence level increases the margin of error, and thus makes the interval wider

what does increasing the sample size do?

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

↗

$\sqrt{n}$

because n is in the denominator increasing n decreases the MOE and decreases the width of the interval

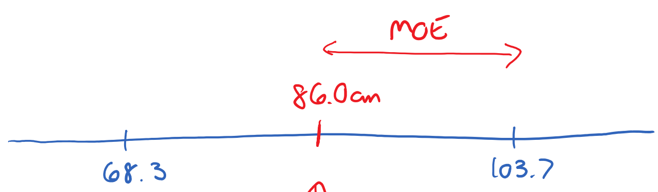
note: if an individual population has large σ (std dev, lots of variation) then the confidence interval will be wider than for a population with less variation

---

trickier question: The 98% confidence interval for a certain measurement is from 68.3 cm to 103.7 cm. If the sample size is 60, what are the sample mean and the sample standard deviation?

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

margin of error



sample mean is midway between these two values

$$\bar{x} = \frac{68.3 + 103.7}{2} = 86.0\text{ cm}$$

so margin of error $= 103.7 - 86.0 = 17.7$

$$\frac{z\sigma}{\sqrt{n}} = 17.7$$

98% confidence $\Rightarrow z = 2.33$

$$\frac{2.33\ \sigma}{\sqrt{60}} = 17.7$$

$$\sigma = \frac{17.7\sqrt{60}}{2.33}$$

$$\sigma = 58.8428$$

$$= 58.8 \text{ cm} \quad \text{or} \quad 59 \text{ cm}$$

$$s \approx \quad '' \qquad\qquad ''$$

$$\boxed{\begin{array}{l} \bar{x} = 86 \text{ cm} \\ s = 59 \text{ cm} \end{array}}$$

---

drawing conclusions based on confidence intervals

example: A study conducted by the doctors at a particular hospital involved monitoring a random sample of 75 surgery patients. The results showed that it took on average an amount of 3.2 cc of anesthetic A to put a patient to sleep, with a standard deviation of 0.4 cc.

However, the latest medical research indicates that the average amount of anesthetic A needed is 3.0 cc.

a) Calculate a 99% confidence interval for the amount of anesthetic A needed in this hospital.

$$\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

99% confidence
$z = 2.575$

$$= 3.2 \pm \frac{2.575(0.4)}{\sqrt{75}}$$

$$= 3.2 \pm 0.1189$$

$$= 3.2 \pm 0.1 \qquad (\text{or } 3.2 \pm 0.12)$$

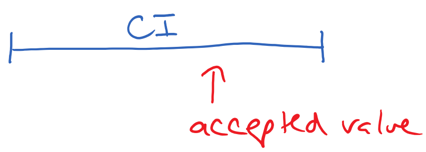
99% CI = 3.1 to 3.3 cc

b) Is there reason to believe that the hospital's amount of anesthetic A differs from the research value? Explain briefly.

Yes, there is reason to believe that the hospital's value is different because the research value lies <u>outside</u> the confidence interval

rule: confidence interval vs. accepted value:

if value is within the interval:



∴ agreement

if value is not in interval:

∴ disagreement

c) A similar study shows that the confidence interval for the amount of anesthetic B required at that hospital is 2.9 to 3.2 cc.

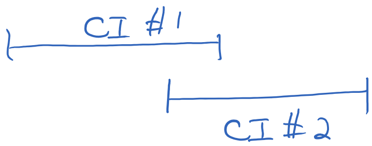
Is there reason to believe that the amounts need for A and B are different?

A : 3.1 to 3.3 cc
B : 2.9 to 3.2 cc

answer: No, the confidence intervals overlap.

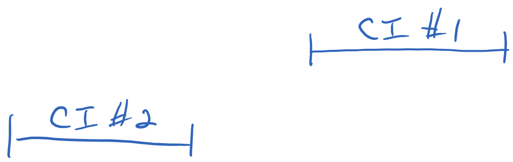
rule: comparing two confidence intervals

if they overlap:



∴ agreement

if they don't overlap:



∴ disagreement

confidence intervals: choosing the sample size

- precision of your estimate (width of the confidence interval) comes from the margin of error
- so, when designing your sampling plan, choose the sample size to ensure that you get the precision you want / need

note: sample size is an <u>integer</u>

you cannot measure the length of 0.85 fish

$\Rightarrow$ ROUND UP

example: suppose you wish to estimate the mean time between failures for a certain brand of disk drive. From previous experience, you know that σ is in the neighbourhood of 200 hours. If you want your estimate of the mean to be precise, with 99% confidence, to within $\pm$ 50 hours of the true value, how many disk drives will you need to test?

$$\mu = \bar{x} \pm \underbrace{\frac{z\sigma}{\sqrt{n}}}_{\text{MOE}}$$

we want MOE $\le$ 50 hours

let's call this the bound B

$$\frac{z\sigma}{\sqrt{n}} \le 50$$

$$\sqrt{n}\left(\frac{z\sigma}{\sqrt{n}}\right) \le (B)\sqrt{n}$$

$$z\sigma \le B\sqrt{n}$$

$$\frac{z\sigma}{B} \le \sqrt{n}$$

$$\left(\frac{z\sigma}{B}\right)^2 \le \left(\sqrt{n}\right)^2$$

$$\left(\frac{z\sigma}{B}\right)^2 \le n$$

$$n \ge \left(\frac{z\sigma}{B}\right)^2$$

99% $\Rightarrow$ $z = 2.575$

$$\ge \left(\frac{2.575 \cdot 200}{50}\right)^2$$

$$\ge 106.09$$

ROUND UP

≥ 107

(I'd say 110, myself)

note: if 90% confidence is sufficient, then can get away with $n \geq 43$ instead

---

how can you decrease the margin of error?

1. increase the sample size

   ↑ good scientific approach

2. decrease the confidence level

   ↑ sadly, the practical ~~weaselly~~ approach (sometimes you are limited by budget/time/resources)