22S:105
Statistical Methods and Computing

Confidence Intervals

Lecture 12
Mar. 6, 2013

Kate Cowles
374 SH, 335-0727
kcowles@stat.uiowa.edu

Confidence in estimation

Example: Studying the quantitative skills of young Americans of working age

We might use the quantitative scores from the national Assessment of Educational Progress (NAEP) Young Adult Literacy Assessment Survey

- possible scores from 0 to 500
- in a recent year, 840 men aged 21 to 25 years were in NAEP sample
  - can be considered a simple random sample from the population of 9.5 million young men in this age range
- mean quantitative score: $\bar{x} = 272$

What can we conclude about the population mean score μ of all 9.5 million young men?

Point estimation

If we had to guess a single number for the population mean μ, our best "educated guess" is $\bar{x}$, the sample mean.

$\bar{x}$ is our *point estimate* of μ.

How great is the uncertainty in this estimate?

Interval estimation: the prelude

Recall essentials about sampling distribution of $\bar{x}$:

- The mean $\bar{x}$ of 840 scores has a distribution that is close to normal (by the Central Limit Theorem)
- The mean of this normal sampling distribution is the same as the unknown mean μ of the entire population
- The standard deviation of $\bar{x}$ for a simple random sample of 840 men is $\frac{\sigma}{\sqrt{840}}$
  - where σ is the standard deviation of individual NAEP scores among all young men

If we knew σ...

Imagine that we know that the true population standard deviation of quantitative scores among all young men is $\sigma = 60$.

Then the standard deviation of $\bar{x}$ is

$$\frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{840}} = 2.1$$

Imagine also that we could choose *many* samples of size 840 and find the mean NEAP quantitative score from each one.

If we collect all these different $\bar{x}$s and display their distribution, we get the normal distribution with

- mean equal to the unknown μ
- standard deviation 2.1

Statistical confidence

- The 68-95-99.7 rule says that in 95% of all samples, the mean score $\bar{x}$ for the sample will be within two standard deviations of the population mean score μ.
  - So the $\bar{x}$ will be within 4.2 points of μ in 95% of samples of 840 NAEP scores
- But if $\bar{x}$ is within 4.2 points of the unknown μ, then μ also has to be within 4.2 points of the observed $\bar{x}$!
  - This will happen in 95% of all samples.
- That is, in 95% of all possible samples of size 840 from this population
  - the unknown μ lies between $\bar{x} - 4.2$ and $\bar{x} + 4.2$

95% confidence

Our sample of 840 young men gave $\bar{x} = 272$.

We say that we are *95% confident* that the unknown mean NAEP quantitative score for all young men lies between

$$\bar{x} - 4.2 \;=\; 272 - 4.2 \;=\; 267.8$$

and

$$\bar{x} + 4.2 \;=\; 272 + 4.2 \;=\; 276.2$$

Every sample would give slightly different values for this interval.

Why are we so confident that μ lies in the interval we happened to get?

There are only two things that could have happened with our particular sample:

- We got a sample such that the true μ does lie in our resulting interval. That is, μ really is between 267.8 and 276.2.
- We were unlucky, and our simple random sample was one of the 5% of all possible samples where $\bar{x}$ is not within 4.2 points of the true μ.

We cannot know for sure which thing happened with our particular sample.

Saying "We are 95% confident that the unknown μ lies in the interval (267.8, 276.2)" means

- "We got these numbers by a method that gives correct results 95% of the time."

What a 95% confidence interval does not mean

Saying "We are 95% confident that the unknown μ lies in the interval (267.8, 276.2)" doesn *not* mean

- μ is a random variable that has a value within the interval 95% of the time
- 95% of the population values lie in the interval

What if we wanted to be *more* confident that our interval contained μ?

We would use a *confidence level* other than 95%.

Example: we will compute a 99% confidence interval for the mean of NAEP quantitative scores in young men

We need the values for a standard normal distribution that cut off the top 0.005 and the bottom 0.005 of values.

- Table A.1 gives several possibilities (due to rounding).
- The most accurate choice is 2.58 for the upper cutoff.

So a 99% confidence interval for μ would be

$$(\bar{x} - 2.58(2.1), \bar{x} + 2.58(2.1)$$

If we didn't need to be all that confident, how would we compute an 80% confidence interval for μ?

Two-sided confidence intervals for a population mean

- Draw a simple random sample of size n from a population having
  - unknown mean μ
  - known standard deviation σ
- A level C confidence interval for μ is

$$\bar{x} \ \pm \ z^* \frac{\sigma}{\sqrt{n}}$$

  where z^* is the value that cuts off the upper $\frac{1}{2}$ of (1 - C) of the area of a standard normal distribution

z^* is called the *critical value*

Critical values for the most commonly used confidence levels

Confidence level	Tail area	z^*
90%	0.05	1.645
95%	0.025	1.960
99%	0.005	2.576

What affects the width of a confidence interval

The width of a confidence interval gets *smaller* if

- The confidence coefficient gets smaller (equivalently, if the level of confidence gets smaller)
- σ gets smaller
- n gets larger

One-sided confidence intervals

What if we only need to be confident that μ is below some upper bound (or above some lower bound), but we don't care how far it might be in the opposite direction?

Example: We are concerned that μ for the NAEP scores might be very low, so we want to find a lower bound. That is, we want to find a value m such that we are 95% confident that $\mu \geq m$.

Begin by drawing the picture!

Now we will use Table A to find the value that cuts off the lower 5% of the area under a standard normal curve.

This is -1.645.

Therefore, we are 95% confident that $\mu \geq \bar{x} - 1.645 \frac{\sigma}{\sqrt{n}}$.

In other words, our one-sided confidence interval for μ is

$$
\begin{aligned}
\mu &\geq \bar{x} - 1.645 \frac{\sigma}{\sqrt{n}} \\
&\geq 272 - 1.645\,(2.1) \\
&\geq 268.55
\end{aligned}
$$