Video

Proportion

Author(s)
David M. Lane

Prerequisites
Introduction to the Normal Distribution, Normal Approximation to the Bin
Sampling Distribution of the Mean, Sampling Distribution of a Proportion,
Confidence Intervals, Confidence Interval on the Mean

Learning Objectives
1. Estimate the population proportion from sample proportions
2. Apply the correction for continuity
3. Compute a confidence interval

A candidate in a two-person election commissions a poll to determine wh
ahead. The pollster randomly chooses 500 registered voters and determir
260 out of the 500 favor the candidate. In other words, 0.52 of the samp
the candidate. Although this point estimate of the proportion is informativ
important to also compute a *confidence interval*. The confidence interval i
computed based on the mean and standard deviation of the *sampling dist*
of a proportion. The formulas for these two parameters are shown below:

$$\mu_p = \pi$$

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{N}}$$

Since we do not know the population *parameter* π, we use the sample pro
p as an estimate. The estimated *standard error* of p is therefore

$$s_p = \sqrt{\frac{p(1-p)}{N}}$$

We start by taking our statistic (p) and creating an interval that ranges (z
in both directions, where $Z_{.95}$ is the number of standard deviations exten
from the mean of a *normal distribution* required to contain 0.95 of the are
the section on the confidence interval for the mean). The value of $Z_{.95}$ is
computed with the normal calculator and is equal to 1.96. We then make
adjustment to correct for the fact that the distribution is discrete rather th
continuous.

Normal Distribution Calculator

s_p is calculated as shown below:

$$s_p = \sqrt{\frac{.52(1-.52)}{500}} = 0.0223$$

To correct for the fact that we are approximating a *discrete distribution* w
continuous distribution (the normal distribution), we subtract 0.5/N from
lower limit and add 0.5/N to the upper limit of the interval. Therefore the
confidence interval is

$$p \pm Z_{.95}\sqrt{\frac{p(1-p)}{N}} \pm \frac{0.5}{N}$$

Lower limit: 0.52 - (1.96)(0.0223) - 0.001 = 0.475
Upper limit: 0.52 + (1.96)(0.0223) + 0.001 = 0.565

0.475 ≤ π ≤ 0.565

Since the interval extends 0.045 in both directions, the *margin of error* is
In terms of percent, between 47.5% and 56.5% of the voters favor the ca
and the margin of error is 4.5%. Keep in mind that the margin of error of

the margin of error for the percent favoring the candidate and not the ma
error for the difference between the percent favoring the candidate and th
percent favoring the opponent. The margin of error for the difference is 9
the margin of error for the individual percent. Keep this in mind when you
reports in the media; the media often get this wrong.

Question 1 out of 3.
Why do we subtract 0.5/N from the lower limit and add 0.5/N to the upp
limit when computing a confidence interval for the population proportion

- We need to correct for the fact that we are approximating a discrete
distribution (the sampling distribution of p) with a continuous distributio
normal distribution).
- The estimate of the population proportion is slightly biased, and we n
correct for it.
- The estimate of the standard error is slightly biased, and we need to
for it.

Check Answer | Previous Question | Next Question

Previous Section | Next Section