

Chapter 10: Confidence Intervals

Friday, December 01, 2023 12:04 PM

Section 10.1: Estimating with Confidence

statistical inference: making predictions (estimates) about populations based on samples

example: Globe and Mail newspaper polls 1000 Canadians and based on this poll, says

"68% of Canadians think ..."



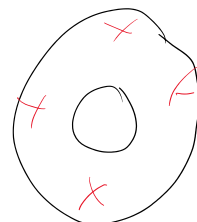
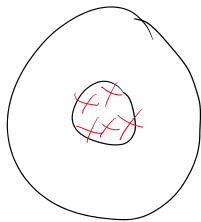
did they measure all Canadians? No!

this is an estimate based on their sample

note: statistical inference also makes decisions about populations based on samples, but that's beyond the scope of this course

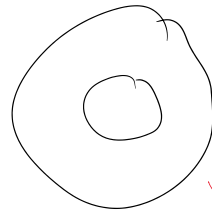
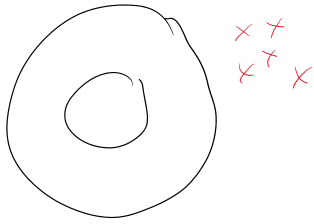
precision vs. accuracy

precise
and
accurate



accurate but
not precise

precise
but
not
accurate



neither precise
nor accurate

Section 10.1: Cent'd

how do you get good accuracy?

- you make sure that as much as possible
your sample is representative

how do you get good precision?

- good measuring instruments and a large
sample size

example:

you measure a sample of adult chinook salmon
and find that the mean length is 75 cm.

what can you say about the entire
population of adult chinook salmon?

answer: if your sample is representative
(which we're going to assume
from now on), you can estimate

that the mean length of all adult
chinook salmon is about 75 cm.

but what does "about 75 cm" mean?

$75 \text{ cm} \pm 1 \text{ cm}$? from 74 to 76 cm
 $75 \text{ cm} \pm 10 \text{ cm}$? 65 to 85 cm

⏟
this is
an interval

and we're going to learn in the next section how to calculate these intervals

but is it not also true that occasionally the sample may contain an outlier? or have more individuals above the mean rather than below it?

occasionally, your sample will, due to random fluctuation, not predict the population mean accurately

how do we handle this? we say something like

"Based on our sample, we have determined that the mean length of adult chinook salmon is between 65 and 85 cm with 95% confidence."

↑
this means that the method we have chosen will give an interval that contains the true value 95% of the time

and 5% of the time, our interval will not contain the true value