As the number of draws increases, the sample averages become more tightly peaked around the expected value.

Inference.

15,125 UCSC students

Population

200 students.

Sample - mean 2.5
SD 0.86

What can I say about the mean GPA of all students in the population, based on my sample?
What's the average GPA at UCSC? = average GPA of the students in the sample.

95% CI for average GPA at UCSC?

GPAs are uniformly distributed between 1 and 4.

For my sample of 200 students

Mean GPA = \boxed{2.5}

SD of sample = 0.86

SE of average = \[
\frac{SD}{\sqrt{\text{# draws}}} = \frac{0.86}{\sqrt{200}} \approx 0.06
\]

95% CI for mean GPA for population

\[
2.5 \pm 2 \times 0.06
\]

\[
2.5 \pm 0.12.
\]
The CI for the mean of the draws is much narrower than the spread in the data.

Width of CI = 0.24

estimate of mean of the population

95% CI

\[
\frac{2.5}{0.12} = 0.24
\]
Hypothesis Testing

Making decisions when uncertainty is present.

Basic question: Is the observed effect due to chance?
- use tests of significance.

Examples:
- A vaccine is known to be 25% effective. A new vaccine is being tested on a random sample of 2000 people. How do we test if the new vaccine is more effective than the old one?

- A machine fills bottles with 333 mL of liquid. Periodically a sample of bottles is taken. How to decide if the average amount is too high / too low?
**Null Hypothesis** vs. **Alternative Hypothesis**

**Ho**
nothing has changed, what we've observed is, due to chance

**Vaccine:**
- **H₀** - effectiveness of new vaccine is 25%
  - the same as the old vaccine (nothing has changed)
- **H₁** - effectiveness > 25%

**bottle filling:**
- **H₀** - average content is 333 ml
- **H₁** - average content is not 333 ml
Revenue neutral changes to tax code

$H_0$ - revenues under new tax code are same as under the old tax code.

$H_1$ - revenues are different

How to test the null ($H_0$) against the alternative ($H_1$)

Build a box model for the null hypothesis ($H_0$)

1 ticket per tax return.

Sample size 100

on each ticket we write the difference in tax paid under the old rules and the new rules.
Sample of 100 tax returns.
Computed old + new tax. => difference in tax paid

Average difference = $219
SD of difference = $725

If Ho is true, expected average difference should be zero.

How far away from the expected value is the observed value?

"How far away" is in terms of standard units.

\[
SE \ of \ average = \frac{SD \ of \ box}{\sqrt{\# \ draws}} = \frac{725}{\sqrt{100}} = 72.5
\]

Compute the difference between observed and expected in standard units.

\[
\frac{-219 - 0}{72.5} = -3.0
\]
If the null hypothesis is true, what's the chance of getting a result as extreme as this or more extreme?

It is very unlikely that chance alone could result in such an extreme value.

**Test statistic.** - measures the difference between data and what's expected under $H_0$

$$ Z = \frac{\text{observed} - \text{expected}}{\text{SE}} \quad \text{assuming } H_0 \text{ is true} $$

- How many SEs away an observed value is from its expected value when expected value is computed using $H_0$
Observed *significance level* is the chance of getting a test statistic as extreme or more than the observed one.

This is usually denoted \( P \), and is called the **P-value**.