population → Simple Random Sample → sample

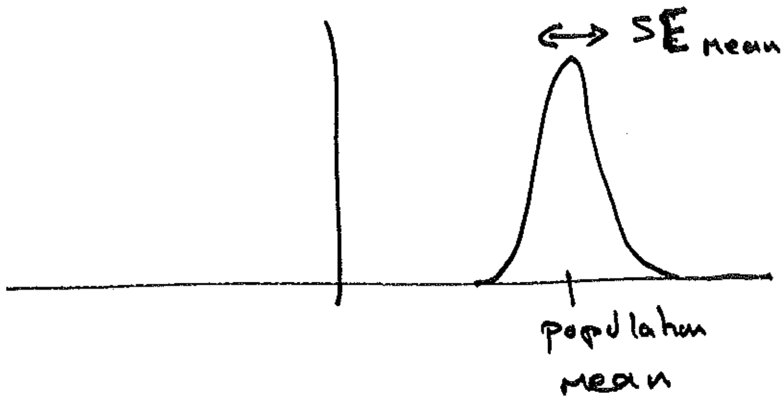
parameter

mean height of all students

↳ data: heights of the students in our sample.

If know population mean, SD

then know what to expect in terms of the variability in the sample means.



← distribution of sample means knowing population mean + SD

Now: based just on the data from one sample, what is our best estimate of the population parameter?

1) Our best estimate of the mean height of all students is the mean of the sample.

2) Derive CI

95% CI is $\underset{\text{mean}}{\text{sample}}$ $\pm$ 2 SE_{mean}.

for 95% of the CIs that we generate from different samples, the population parameter will lie inside the CI.

(the CI "covers" the population parameter value)

<u>sample</u> : mean 66.1 inches
std. 5.1 inches
n 142.

Estimate of the mean height of all students is 66.1"

$$SE_{mean} = \frac{SD.}{\sqrt{\#\text{samples}}} = \frac{5.1}{\sqrt{142}} = \frac{5.1}{11.9} = 0.43 \text{ inches}$$

95% CI for mean height of all students is

<u>66.1 $\pm$ 0.86 inches.</u>

Hypothesis Tests.

- decisions under uncertainty.

H_0 - null hypothesis

H_1 - alternative hypothesis

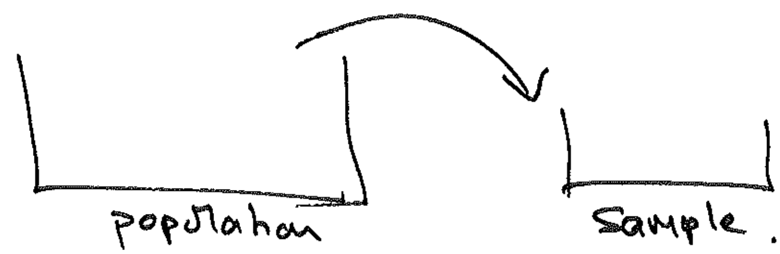
Can we explain the data in the sample ~~by~~ just by sampling variability assuming that H_0 is true, or is the data sufficiently extreme (different from what H_0 would predict) that it is unreasonable to believe that it came from H_0?

- given a box model for H_0, and considering the data as draws from the box, can chance variation explain the data?

Tax code example.

H_0 : revenue under the revised tax code will be the same as under the existing tax code

H_1 : revenue under the revised tax code will be different.



Use data from the sample to decide if the population mean has changed.

100 returns drawn at random, compute the difference in tax between the old rules and the new rules.

If the null hypothesis is true, then the average difference should be zero.

For the 100 tax returns in the sample.

Average difference was $-\$219$, SD of the differences was $\$725$

Need to measure how far the observed value (-219) is from the predicted value (0), <u>on the scale</u> of the amount of variability we expect.

- the amount of variability we expect is given by the standard error.

- SE for average of the tickets in a box $= \dfrac{\text{SD of box}}{\sqrt{\#\text{ samples}}}$.

→ use SD of sample in place of SD box

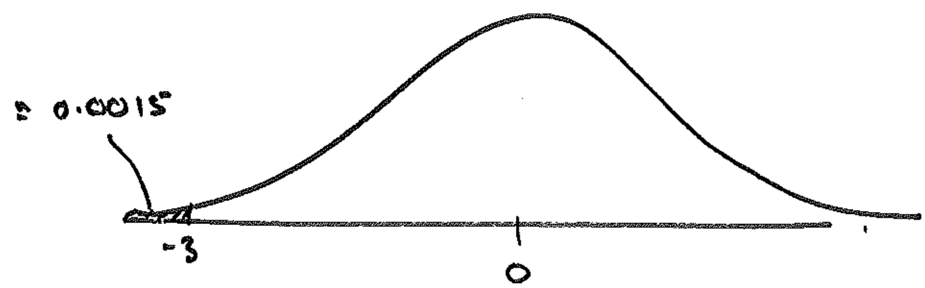
$$SE = \frac{725}{\sqrt{100}} = \frac{725}{10} = 72.5$$

→ Convert the observed value to standard units to get a measure of how large it is on the scale of the expected variability.

$$\frac{-219 - 0}{72.5} \simeq \underline{\underline{-3.0}}.$$

↑ test statistic.

$$z = \frac{\text{observed} - \text{expected}}{SE}$$



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.

Reject the null hypothesis that the revenue under the two tax codes will be the same and conclude instead that they will be different.

Quantify "very unlikely"

the observed *significance level is the chance of getting a test statistic as extreme or more than the observed one.

It is usually denoted "p" and called the p-value.

Note: the p-value is not the chance of H_0 being correct / incorrect.

It is the chance that, if the null hypothesis is true, that your observed data will be so extreme due to only chance variability.

Argument by contradiction.

1/ Assume the null hypothesis is true

2/ Compute the chance of the data, assuming the null hypothesis.

3/ Reject the null hypothesis if the chance is too small.

Significance levels.

$p < 5\%$ "statistically significant"

$p < 1\%$ "highly statistically significant"

- these are arbitrary, but conventional.

[p-values are tail areas under the normal curve]

Summary

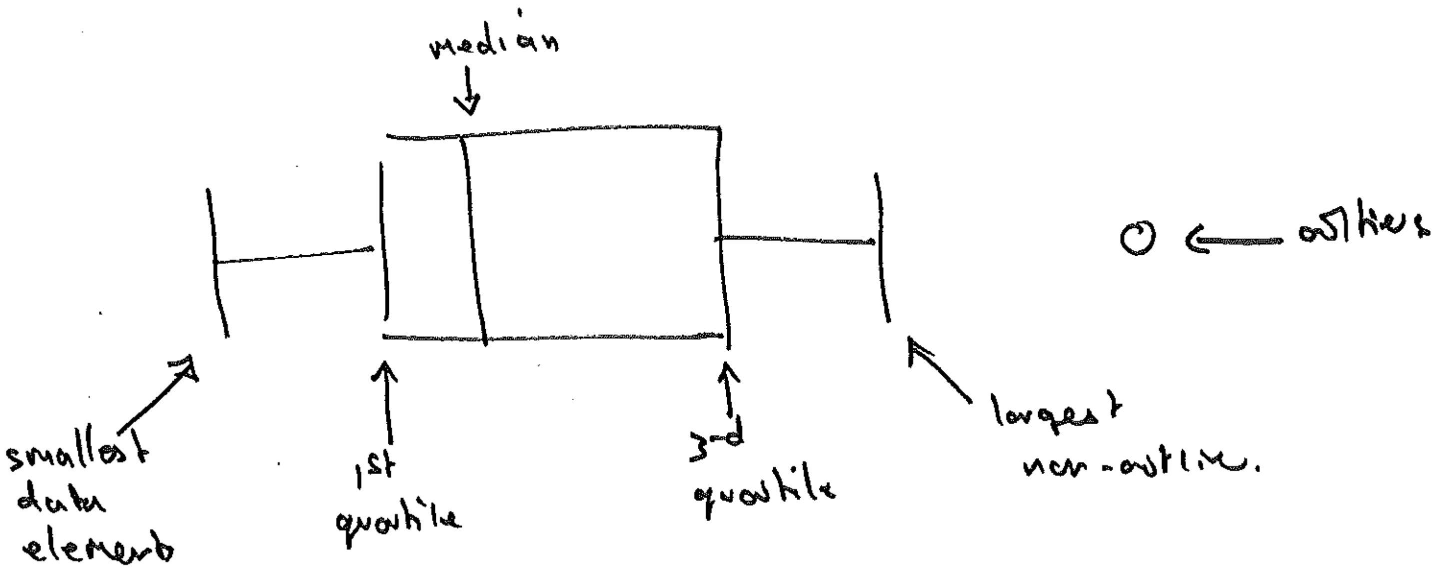
1/ Set up H_0 and H_1

2/ Pick a test statistic to measure the difference between ~~what is~~ ~~exp.~~ the data and what is expected under H_0

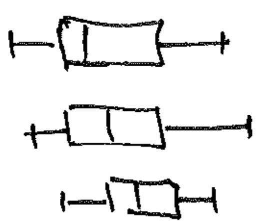
3/ Compute the test statistic + corresponding significance level.

4/ Reject / do not reject H_0.

Box Plot. - shows the distribution of data.



(outlier is > 1.5 IQR above 3rd quartile).



Another example.

An administrator claims that the mean score of 8^{th} graders on a national test is above 260.

A random sample of 85 8^{th} graders has

mean score 265

SD 55

Does this data support the administrator's claim?

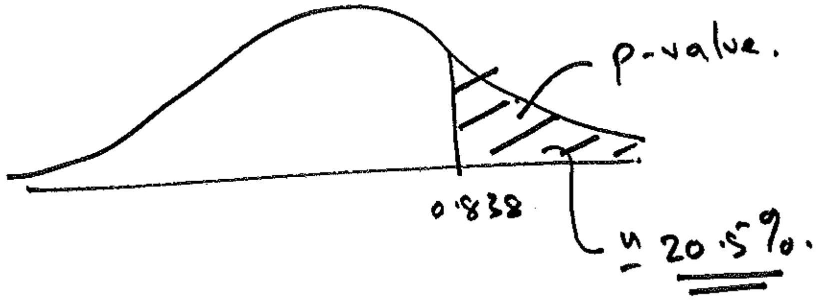
H_0 : mean ≤ 260

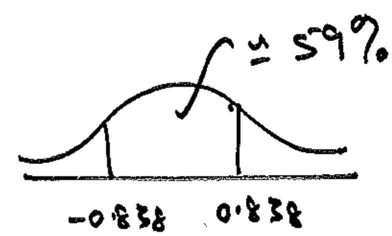
H_1 : mean > 260

if we can reject the null hypothesis, then we support the administrator's claim.

test statistic $z = \dfrac{\text{observed} - \text{expected}}{SE}$

$$= \frac{265 - 260}{55/\sqrt{85}} = \underline{0.838.}$$





p-value is $\simeq 20\%$

ie ~~Ther~~ One time in 5, if the mean is ≤ 260,

we will see a mean score in our sample

of ≥ 265.

This is not enough evidence to reject Ho.

Cannot conclude that mean score > 260.

---

light bulbs are claimed to have a mean

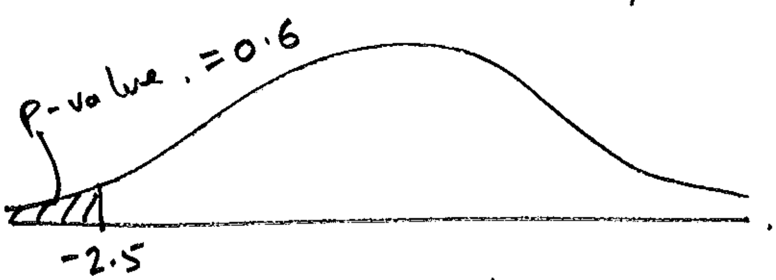
lifetime of 750 hours.

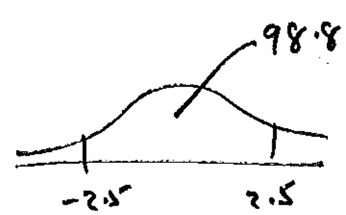
Random sample of 36 has mean lifetime 725

SD 60

Is there evidence to reject the claim?

test statistic $z = \dfrac{725 - 750}{60/\sqrt{36}} = -2.5$

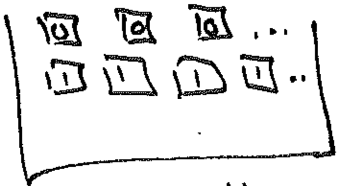




significant evidence against Ho.

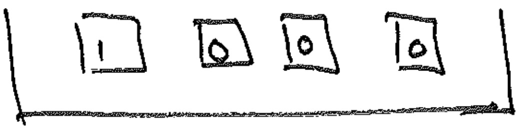
Significance tests for Classifying + Counting.

Vaccine example :



population

vaccine is either
effective
not effective.

H_0 

vaccine is 25% effective

H_1 : vaccine is more than 25% effective.

In sample of size 2000 , new vaccine was effective for 534 people.

Does this give evidence to reject H_0?

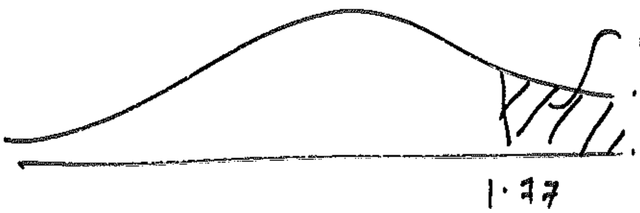
$$Z = \frac{\text{obs} - \text{expected}}{SE}$$

$$SE_{\text{of sum of draws}} = \sqrt{\#\text{draws}} \times SD_{\text{box}}$$

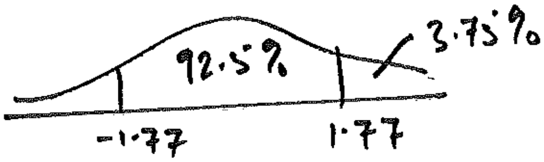
$$= \sqrt{2000} \times (1-0) \times \sqrt{1/4 \times 3/4} = \underline{\underline{19.2}}.$$

$$Z = \frac{534 - 500}{19.2} = \underline{\underline{1.77}}.$$

(How often would we expect a test statistic this large if the box does contain 25% $\boxed{1}$'s ?)



p-value = 3.75%.



p-value is < 5% so reject Ho that the vaccine is 25% effective, and conclude that it is in fact > 25% effective, at 5% significance level.

---

p-value is the chance of rejecting Ho when it is actually true.
