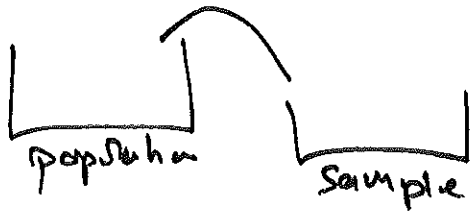
Standard Deviation

Standard Error.



1
2
3
6
4
9
3

* compute the SD for any list of numbers.

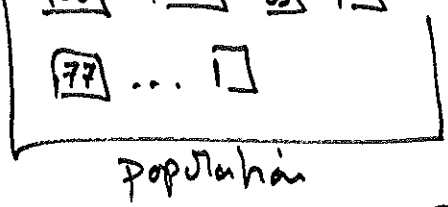
mean $\frac{28}{7} = 4$

SD - compute deviations from mean
square them
form average
take sqrt.

$$\sqrt{\frac{(1-4)^2 + (2-4)^2 + (3-4)^2 + (6-4)^2 + (4-4)^2 + (9-4)^2 + (3-4)^2}{7}}$$

$$\sqrt{\frac{9+4+1+4+0+25+1}{7}} = \sqrt{\frac{44}{7}}$$

SD. - measure of the spread in the data



SD of population
- tells us about the amount of variability in the heights of the people in the population

Standard Error.

Think about taking many samples of size N and for each of the samples, we compute the mean.

There will be variability in the means between samples.

SE measures this variability

$$SE_{mean} = \frac{SD_{box}}{\sqrt{N}}$$

Hypothesis Tests.

Hypothesis – a claim about a population parameter.

Data for a sample

Measure the difference between the observed data and the expected value from the hypothesis, on the scale of the amount of variability expected.

Can we explain the ~~observed data~~ difference between the observed data + the expected value just in terms of sampling variability?

H_0 – null hypothesis

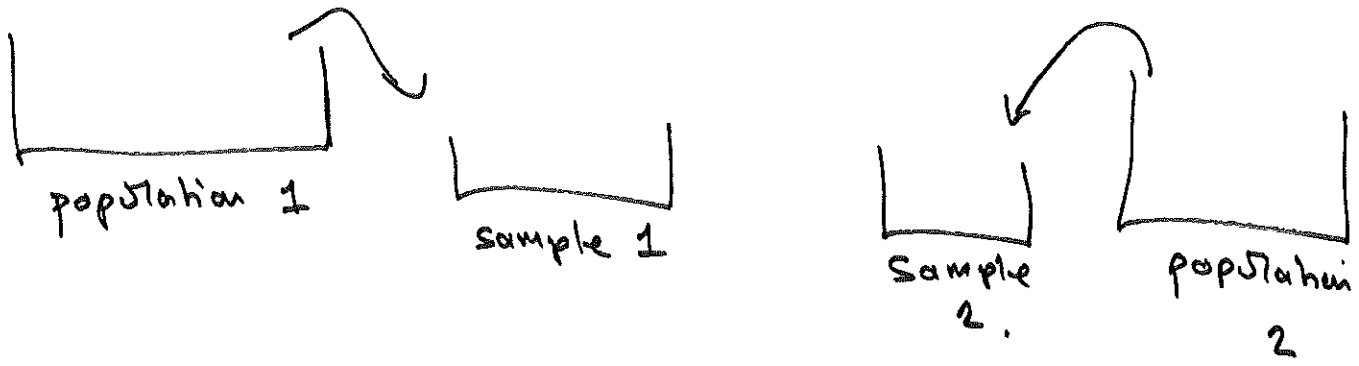
H_1 – alternative hypothesis

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

p-value → probability of getting a test statistic as large or larger than the one we have, assuming H_0 is true.

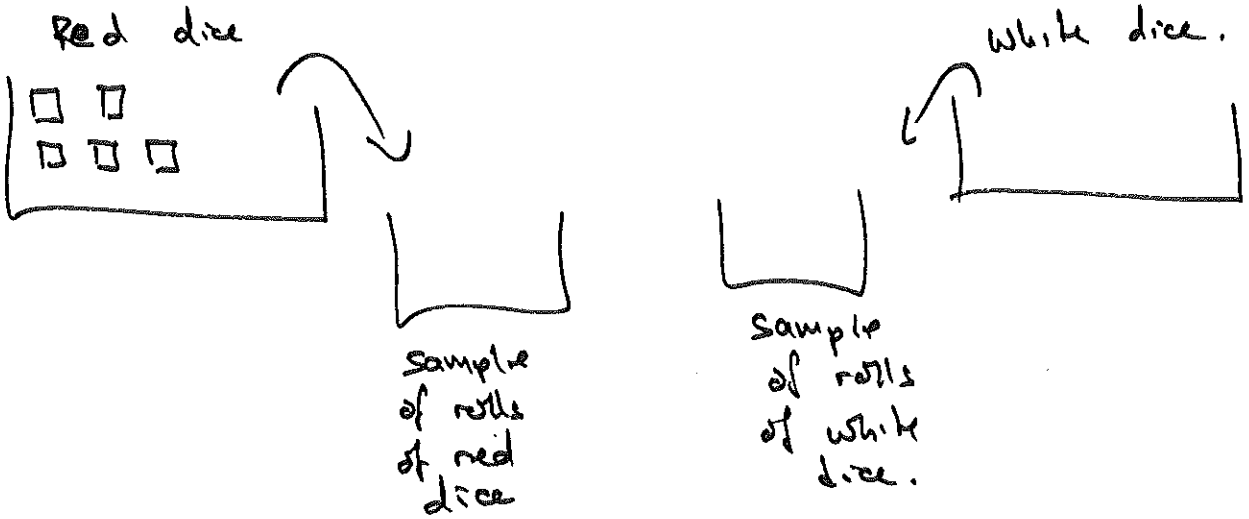
If p-value < 0.05 reject H_0 and say that the result is "statistically significant"

If p-value < 0.01 reject H_0 "highly statistically significant"



Are the two populations the same?

Based on the data in the two samples, can we decide whether parameter 1 is the same as parameter 2?



Does the total score for 10 rolls of the white dice differ from the total score for 10 rolls of the red die?

Each box contains tickets for the total from 10 rolls.

H_0: the averages of the two boxes (ie population means) are the same.

idea: how big is the observed difference in the sample means in terms of the size of the difference we should expect just from sampling variability?

2-sample z-statistic $= \dfrac{\text{observed difference} - \text{expected difference}}{SE_{\text{difference}}}$

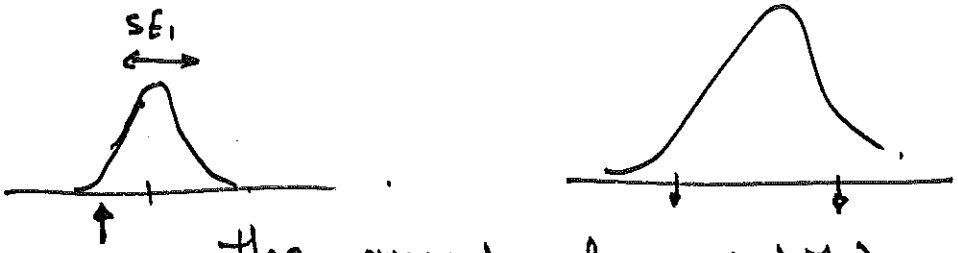
Red :	sample mean	33.7	white :	sample mean	34.9
	sample SD	8.6		sample SD	6.3
	sample size	68		sample size	85

$SE_{mean} = \dfrac{SD}{\sqrt{\text{sample size}}}$

$= \dfrac{8.6}{\sqrt{68}} = 1.04$

$SE_{mean} = \dfrac{6.3}{\sqrt{85}}$

$= 0.68.$



the amount of variability we expect in the difference of the sample means is <u>larger</u> than the variability in either of them.

SE for the difference of means of

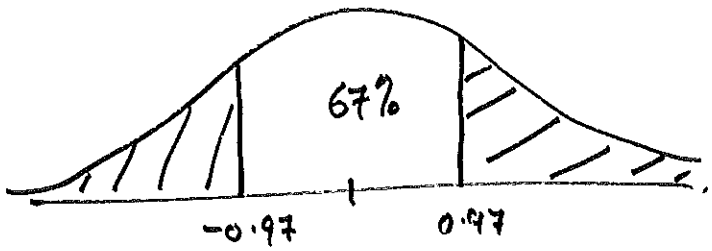
<u>independent</u> samples

$$SE_{diff} = \sqrt{(\text{first SE})^2 + (\text{second SE})^2}$$

$$SE_{diff} = \sqrt{(1.04)^2 + (0.68)^2} = \underline{1.24}$$

$$\text{2 sample } z\text{-statistic} = \frac{(33.7 - 34.9) - 0}{1.24} = \frac{-1.2}{1.24}$$

$$= \underline{-0.97}$$



The two shaded areas <u>together</u> represent the probability of obtaining a test statistic as extreme or more extreme than the one observed.

p-value is $1 - 0.67$

$= \underline{0.33}$.

Do not reject the null hypothesis (that there is no difference between total of 10 rolls between red & white dice)

Another Example.

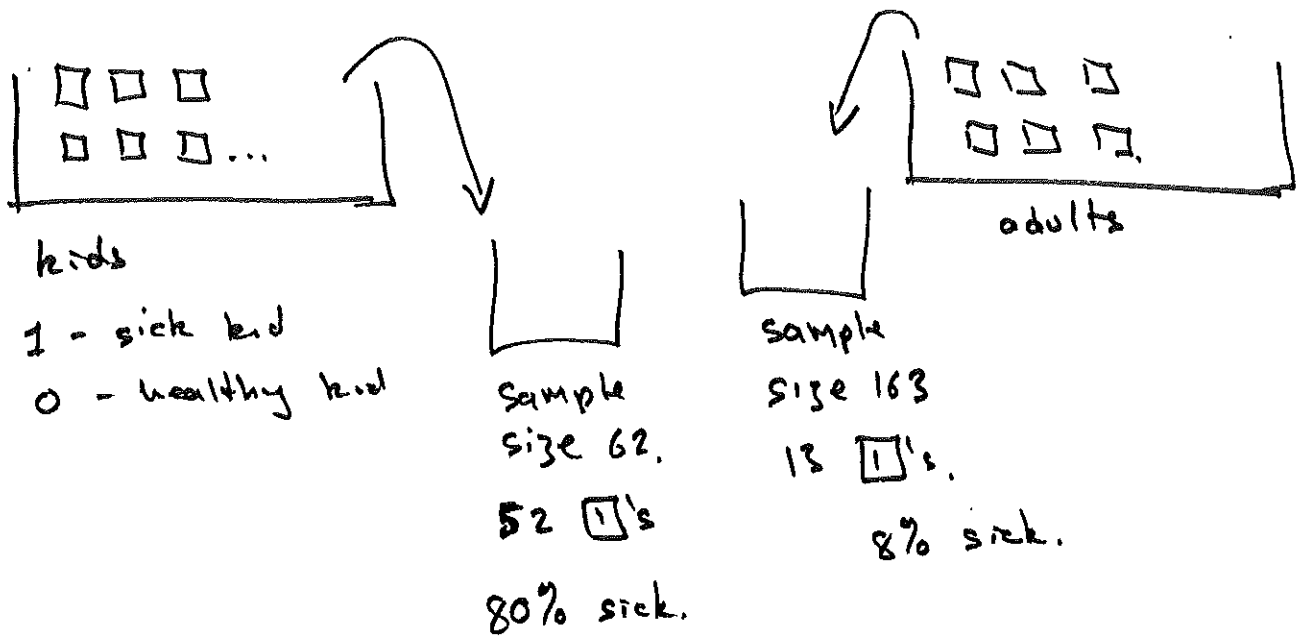
During a flu outbreak.

A sample (SRS) of 62 kids produced 52 sick kids

A sample (SRS) of 163 adults produced 13 sick adults.

Do kids and adults get sick at the same rate?

- box model.



H_0 : % of 1's in each box (adults / kids) is the same.

H_1 : % of 1's is different

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

kids.

$$SE_{\#\text{kids}} = \sqrt{\text{sample size}} \times SD_{\text{box}}$$

$$= \sqrt{62} \times (1-0)\sqrt{0.8 \times 0.2}$$

$$SE_{\%} = \frac{SE_{\#\text{kids}}}{\text{sample size}} \times 100$$

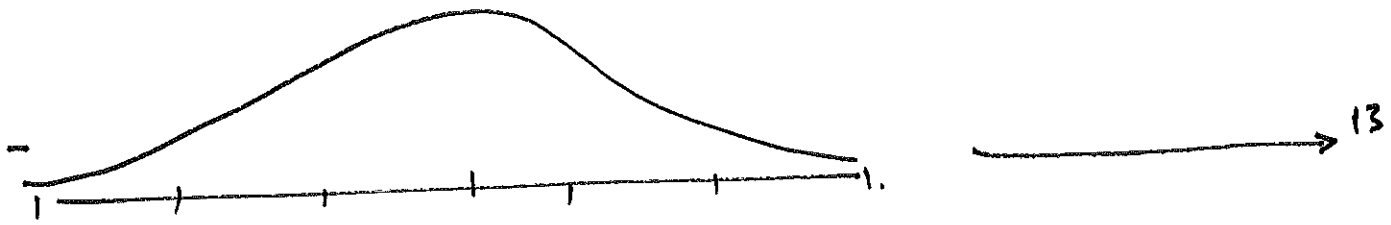
$$= \frac{\sqrt{62} \times \sqrt{0.8 \times 0.2}}{62} \times 100 = \underline{\underline{5.1}}$$

adults.

$$SE_{\%} = \frac{\sqrt{163} \times (1-0)\sqrt{0.08 \times 0.92}}{163} \times 100 = \underline{\underline{2.1}}$$

$$SE_{\text{diff in \%}} = \sqrt{5.1^2 + 2.1^2} = 5.5$$

$$z = \frac{(80-8) - 0}{5.5} = \underline{\underline{13}}$$



p-value is almost zero ($<< 0.01$)

Reject H_0 and conclude that adults + kids get sick at different rates.

What's wrong with what we've just done?

- the 62 kids were all from the same school
- the 163 adults were the adult family members of the 62 kids.

Hypothesis tests rests on some assumptions:

- samples are SRS from a population
- when calculating $SE_{difference}$, we assumed that the samples were independent.

Think carefully: is the analysis we're doing actually valid for the situation we're applying it to?

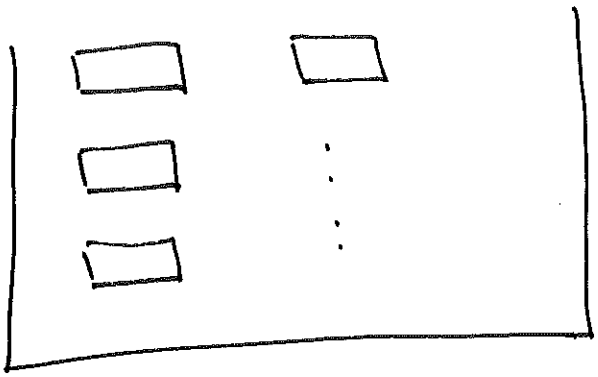
However:

there is one situation where we don't have independent simple random samples from two populations, but the test still works.

- Randomized Controlled Double Blind Experiment.

eg taking vitamin C to prevent colds.

- half get treatment, half get placebo
- measure: # colds each subject had.



1 ticket for each person in the trial.

what goes on the tickets?

 two numbers.

1st - what we would measure if this person was in treatment group.

2nd - what we would measure if this person was in control group

# colds on vit C	# colds on placebo

treatment.

chose 100 at random, look at only the 1st number

ave # colds. 2.3

SD 3.1

control

chose 100 at random and only look at the 2nd number.

ave # colds 2.6

SD 2.9

H_0: the average # of colds in the same in the two groups.

$$z = \frac{\text{observed diff} - \text{expected diff}}{SE_{diff}} = \frac{(2.3 - 2.6) - 0}{SE_{diff}}$$

assume that we have 2 independent samples, drawn w.th replacement.

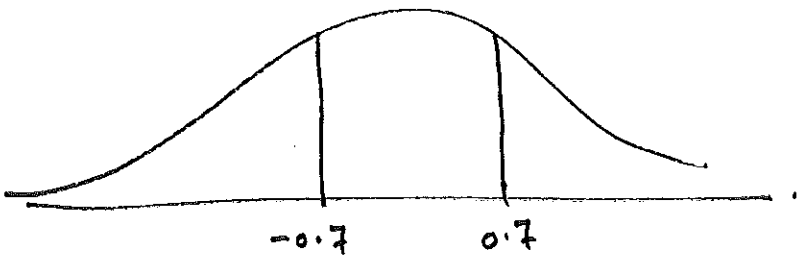
(actually: not independent drawn **without** replacement)

$$SE_{mean\ treatment} = \frac{3.1}{\sqrt{100}} = 0.31$$

$$SE_{mean\ control} = \frac{2.9}{\sqrt{100}} = 0.29$$

$$SE_{diff} = \sqrt{0.31^2 + 0.29^2} = 0.42.$$

$$z = \frac{-0.3}{0.42} = -0.7$$



p-value > 32%

The difference * could be due to chance.

Do not have enough evidence to reject H_0.

Why can we do this, when the assumptions do not hold?

draws made without replacement,

but SE computed ~~assuming~~
as if drawing with replacement

inflates SE

averages are dependent, but
we combine the SEs as if
they were independent

reduces SE

---

cancel.