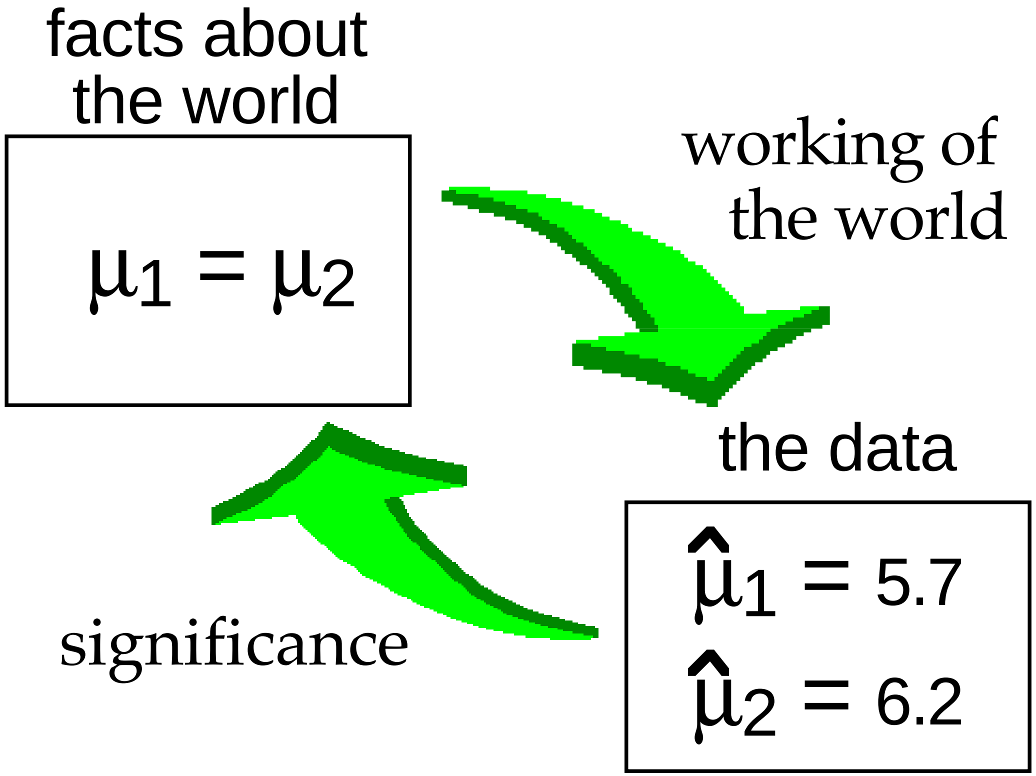
proving things

- proving things are different
  importance & significance
- proving things are the same
  confidence intervals

working backwards again



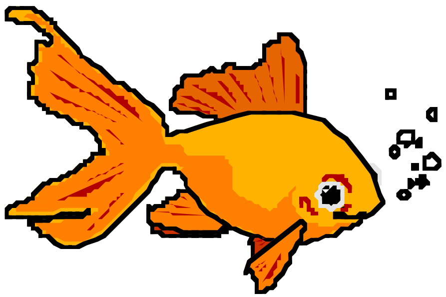
deduction

- all men are mortal
- Socrates is a man
- therefore Socrates is mortal

induction

- all men are mortal
- Socrates is mortal

? therefore Socrates is a man



scitsitats ↔ statistics

we know:

  if real means are different
  then sample means will be different

we measure:

  sample means are different

? what to deduce

significance test

hypotheses:

H_1 – what want to show
H_0 – null hypothesis (to disprove)

- if H_0 were true
  then observed effect is very unlikely

  therefore H_1 is (likely to be) true

5% significance level?

it says:

if	H_0 were true
then	probability observed effect happening by chance is less than 1 in 20 (5%)
so	H_0 is unlikely to be true
and	H_1 is likely to be true

does <u>not</u> say

- ✘ probability of H_0 is < 1 in 20
- ✘ probability of H_1 is > 0.95
- ✘ effect is important
  i.e. significant in the real sense

all it says

✔ if H_0 were true ...

... effect is unlikely

(prob. < 1 in 20)

proving differences

often:

H_0 – some things are equal

H_1 – they are different

statistically significance if

observed difference >> random variation

$$\hat{\mu} \gg \hat{\sigma}/\sqrt{n}$$

proving differences

$$\frac{\hat{\mu}}{\hat{\sigma}/\sqrt{n}}$$

bigger ratio $\Rightarrow$ smaller p

(less likely by chance)

top and bottom <u>both</u> estimates

(where ‘t’ test comes in)

reasons for significance

$$\frac{\hat{\mu}}{\hat{\sigma}/\sqrt{n}}$$

$\hat{\mu}$ large – large difference (?important?)

$\hat{\sigma}$ small – small natural variation (e.g. natural science)

n large – large sample

little things matter

$$\frac{\hat{\mu}}{\hat{\sigma}/\sqrt{n}}$$

what if small differences are important?
(i.e. $\hat{\mu}$ small)

$\hat{\sigma}$ fixed by the world
$\Rightarrow$ need large number of samples (n)

quick test

true or false?

① 5% significant ⇒
H_0 is false & H_1 is true
i.e. things are different

② <u>not</u> significant ⇒
H_0 is true & H_1 is false
i.e. things are the same

neither!!

① 5% significant ⇒

things are different

nearly true

② not significant ⇒

things are the same

NO! NO! NO!!!!!

① significant result

normally reason:

significant result ⇒ H_0 false

things are different

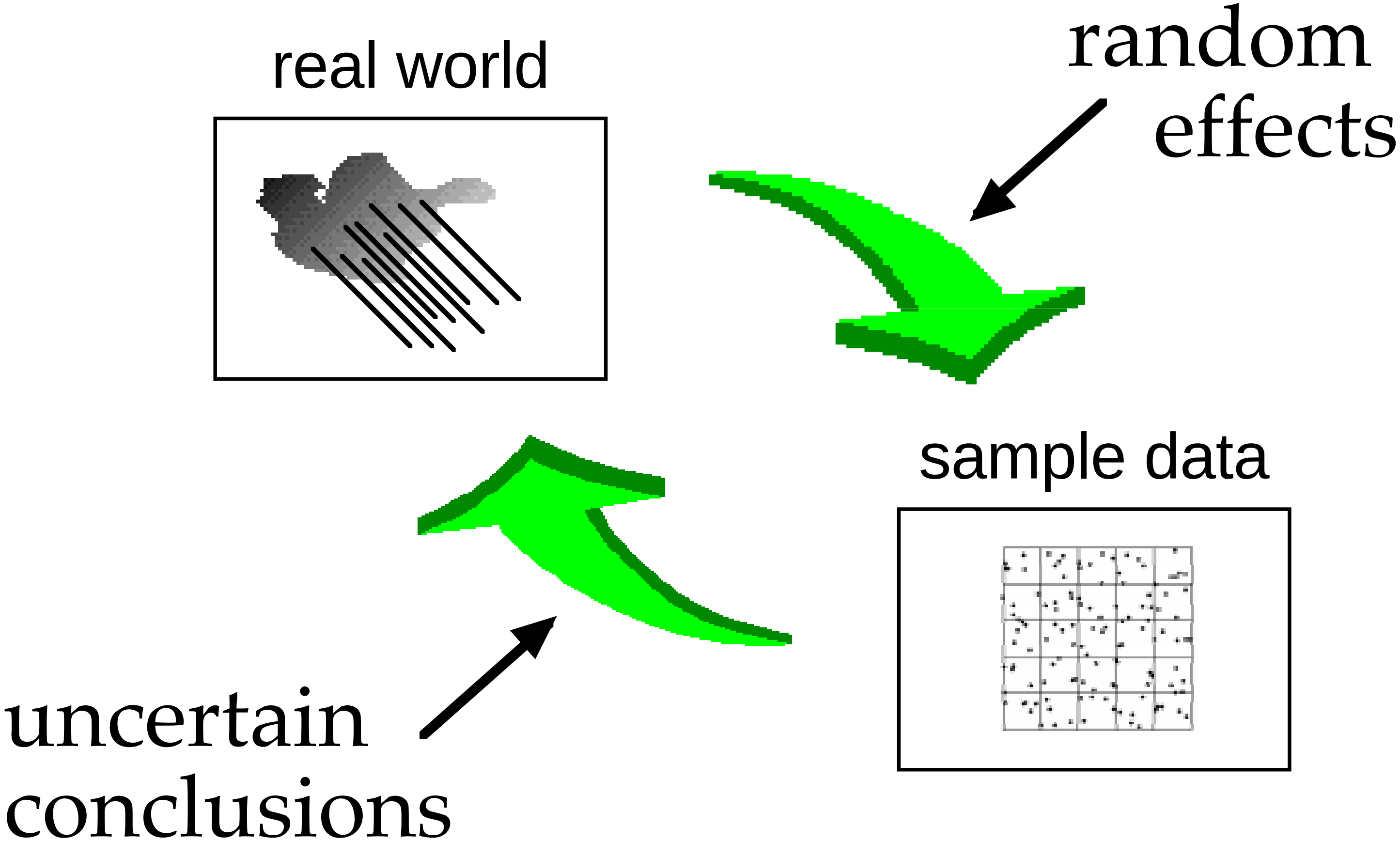
- 1 time in 20 you will be <u>wrong</u>

e.g. conclude drug improves health p<5%

1 in 20 chance it kills everyone

- 1 in 20 significant results are false

statistical proof?



statistical proof

<u>all</u> you can do is:

- say something is true
- know how often you will be wrong
  (on average)
- choose how often you will be wrong!
  (i.e. choose the significance level)

② non-significant result

can **<u>NEVER</u>** reason:

✘ non-significant result $\Rightarrow$ H_1 false
things are the same ✘

<u>all</u> you can say is:
H_1 is not statistically proven

reasons for non-significance

$$\frac{\hat{\mu}}{\hat{\sigma}/\sqrt{n}}$$

$\hat{\mu}$ too small – small effect

(possibly important)

$\frac{\hat{\sigma}}{\sqrt{n}}$ too large – insensitive experiment

not enough samples (n too small)

too much variability ($\hat{\sigma}$ too big)

lies ...

prosecution: and when did your dislike of the victim turn to hatred

defence: objection

prosecution: withdrawn

damn lies ...

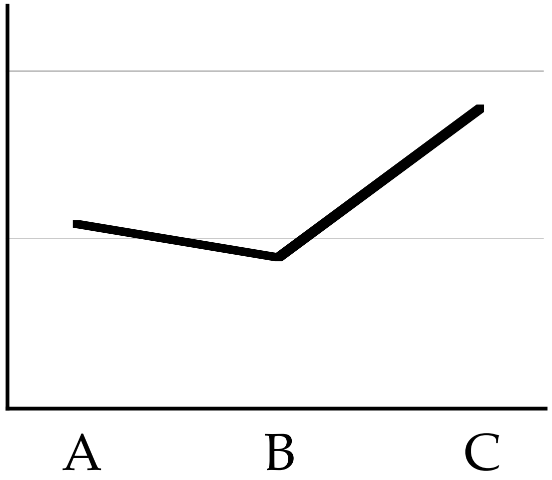


Figure 5.3

A	5.6
B	4.7
C	9.2

Table 5.7

The results of comparing the response times in the three conditions are shown in table 5.7 and figure 5.3. The response times in condition C were longer than those for A and B, but this difference was not significant (p=23%).

at InterCHI

presenter: (*shows slide with graph*)
... not significant

(*at end of presentation*)

questioner: the results in the graph
looked very interesting ...

proving equality

- non-significant – not <u>proved</u> different
- real difference may always be smaller than experimental error

  $\Rightarrow$ can <u>never</u> prove equality
- can put bounds on inequality

confidence interval

- bound on true value
- mean of data is 0.6
  95% confidence interval is [–0.7,1.3]
- says if you conclude:
  real mean is in the range [–0.7,1.3]

  95% of the time you will be right!

counterfactuals

- 95% confidence interval is [–0.7,1.3]
- does <u>not</u> say:
  there is 95% probability that the
  real mean is in the range [–0.7,1.3]
- it either is or it isn't!
- <u>all</u> it says:
  95% probability that you are right

proven!

H_0: no difference (real mean is zero)

experimental result:	mean is 0.6
significance test:	n.s. at 5%
	– so what!
95% confidence interval:	[–0.7,1.3]

? is 1.3 is an important difference