

sampling error (chance)

'replication' - means more than 1 datum gathered the same way. But there is also 'more' vs 'less' replication.

- How much replication is enough?

it depends . . .

- Sampling error can be quantified + is often reported if we know sample size + another property of the data we can say how uncertain we are about current data reported as margin of error, confidence limit

a statistical test is a test of sampling error

Why do a statistical test?

- to tell us if a difference or effect is meaningful
- we use the stat. test to tell us whether the difference/effect ~~error~~ can be attributed to sampling error (SE).
- if (possibly) due to SE, then we suspect difference/effect is not real.

a result that is 'not significant' is one in which sampling error can explain the difference/effect.

2) Human + technical error

Screw-ups. Data not recorded correctly, machines miscalibrated, badly designed machines.

AVOIDABLE in strict sense. We ~~do~~ take steps to reduce H+T error, but cannot eliminate all possibility.

detecting H+T error is an important step in reducing it.

fixing + detecting H+T error

- a) design procedures that ~~do~~ reduce H+T error
- b) use data standards ('knowns')  
e.g. Thermometer

proficiency test - is the test of a data-generating system using samples of known value (standards)

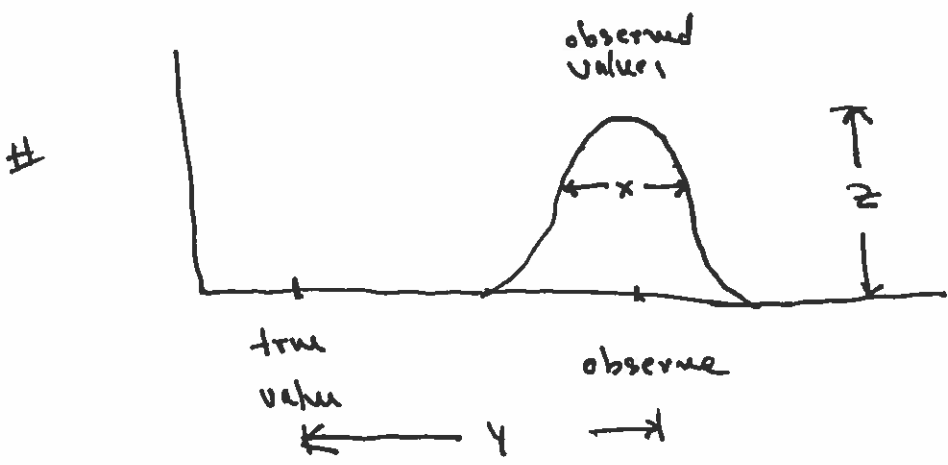
3) Bias (unintentional)

(random odd number exercise)

1	3	5	7	9
	.24	.13	.40	
	.22	.17	.44	

people don't think randomly

bias is a consistent, directional difference between the data you observe/gather + what they should represent



which of x, y, z represents bias?

## Contexts for unintentional bias

- a) choosing things
- b) when we are gathering subjective data or when presenting data
- c) confirmation bias - we usually look for support of our pre-existing views.

to 'fix' problems w/ bias

- a) when choosing things, do so randomly (random number generator, coin flip, die, draw from a hat, ...)
- b) blind - do things w/o knowing what's expected
- c) combine data w/ complementary sources of bias - success is NOT Assured
- d) confirmation bias - must look at opposing views

---

## 2 levels of blind

blind subjects subjects are those from wh/ data are being gathered  
means that the subjects don't know what's expected of them  
(only relevant if subjects are people)

in medical studies (tests of drugs, treatments) we make the control group (untreated subjects) ~~think~~ think they are (also) being treated by giving them a fake treatment (placebo)

blind observers - means those who are making the measurement how their subjects were treated

if both blind subjects + blind observers = double blind

something can be blind in one respect but not another

comment on random - randomization can be applied only if there is more than one assignment

## 4) RPA - rounding, precision, accuracy

quantitative measurements are never exact -  
always truncated in some fashion

speed

distance

time

mass/weight

RPA - due to limitations in our equipment

fix: better equipment

## 5) Flawed analysis + processing,

raw data almost never used directly. They are 'processed'  
+ the papers/reports usually only give us the processed data.

if the processing/analysis done incorrectly, may get a wrong  
conclusion

true fix - analyze properly

publishing the raw data allows others to reanalyze + maybe  
do it right

## (6) Protocols

a description (model) of how the data were or will be collected

explicit protocol - written or recorded protocol

protocols are critical to data quality (+ science) for 3 reasons

1) protocol affects types of errors present

2) allows the data to be gathered consistently on different occasions

\*\* 3) allows anyone who did not gather the data to identify  
the types of errors present

(inconsistent adherence to protocol = H+T error)

# Science + Criminal Justice System

much criminal evidence is about connecting a person to a crime.

what methods are used for matching?

DNA	blood type	shoe prints
<del>P</del> fingerprints	video	tire marks
eyewitness	<del>to</del> lie detector	odor (dogs)
bite mark	hair	