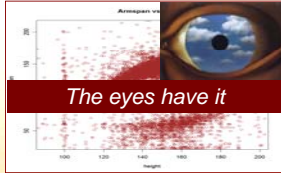


Statistical Informal Inference Revisited



Chris Wild
Dept of Statistics, University of Auckland
New Zealand

The ideas in this talk have developed

through a long series of brainstorming sessions about informal inference with:



Maxine Pfannkuch
U. of Auckland, NZ



Matt Regan



Nick Horton
Smith College, MA, USA

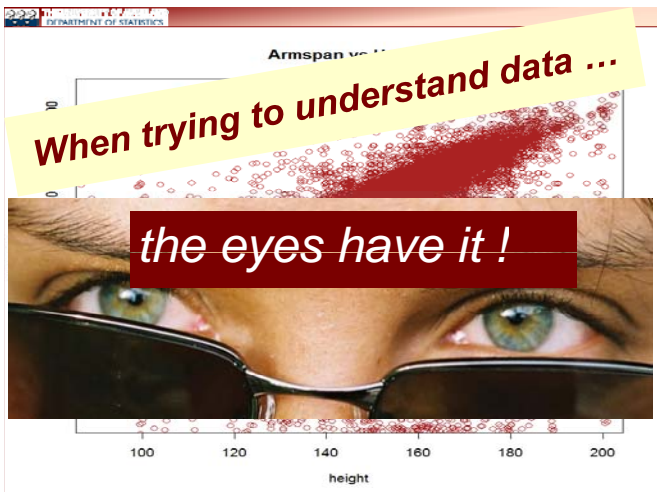
- important new element of the new curriculum

What is it?

- plain old statistical inference, but ...
 - **operated simply** enough for young students

We will ...

- Start with the big ideas of statistical inference
- Describe simple methods for students to apply **when looking at their own data**
 - Minimise steps that lead students to take their eyes off the data
 - **"Exploit the power of the visual sense"**

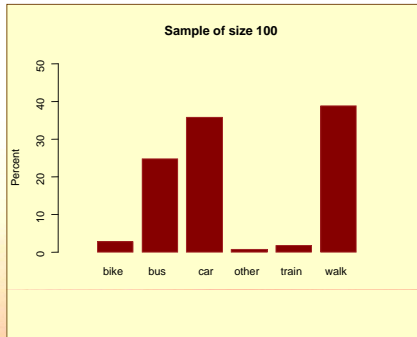


Let's look at some data from

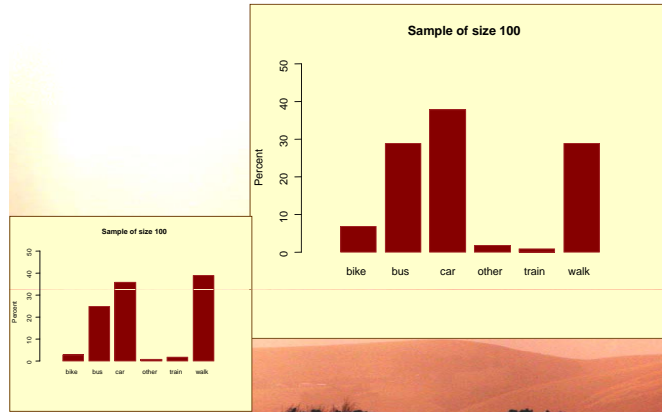


<http://www.censusatschool.org.nz/>

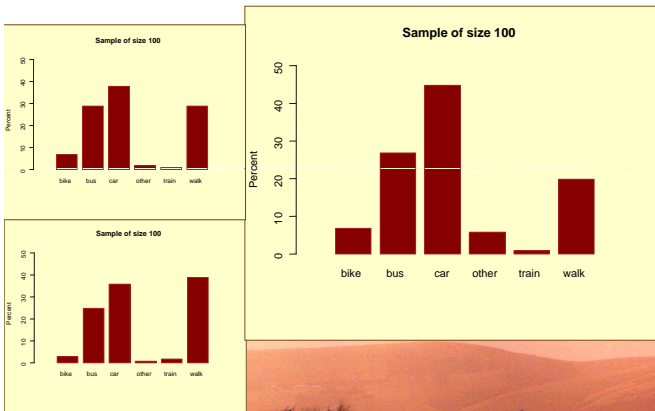
How did they travel to school ?



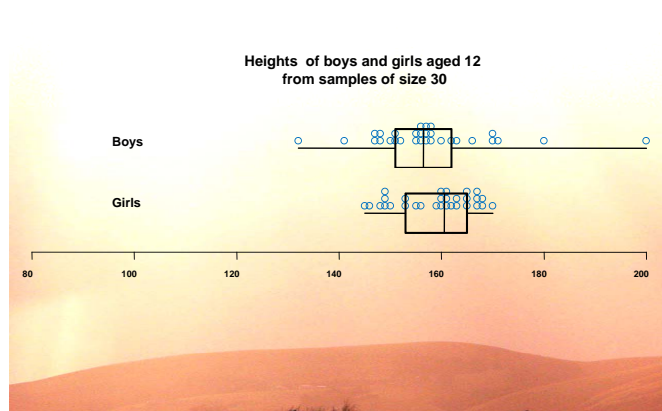
How did they travel to school ?



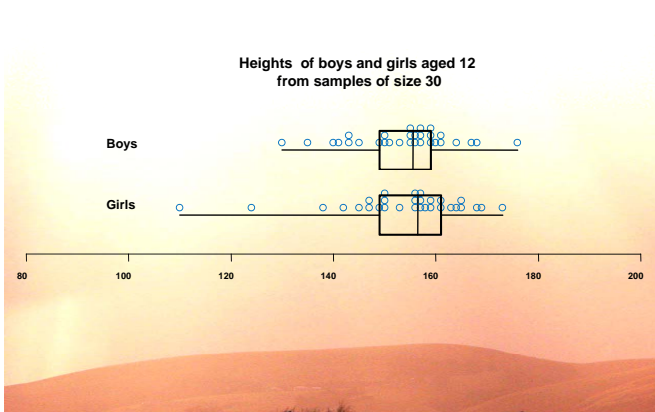
How did they travel to school ?



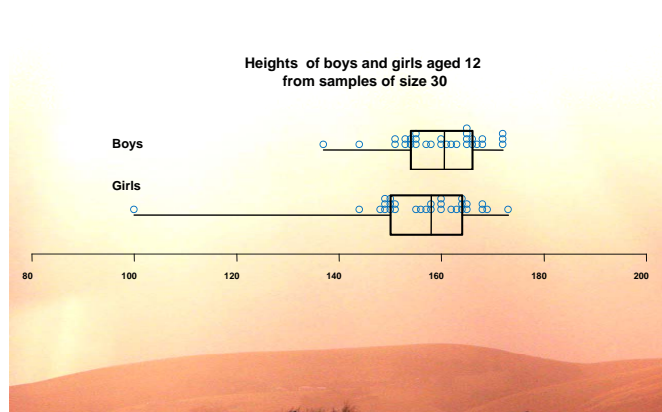
Comparing heights of boys and girls at age 12



Comparing heights of boys and girls at age 12

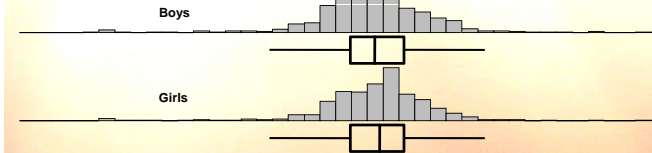


Comparing heights of boys and girls at age 12



Heights of boys and girls aged 12

Population distributions



A nationwide online survey for Year 5 - 13 students which provides real, relevant data and classroom activities to enhance statistical enquiry across the curriculum.

- survey
- data
- classroom activities
- new curriculum

Are you a masterpiece?

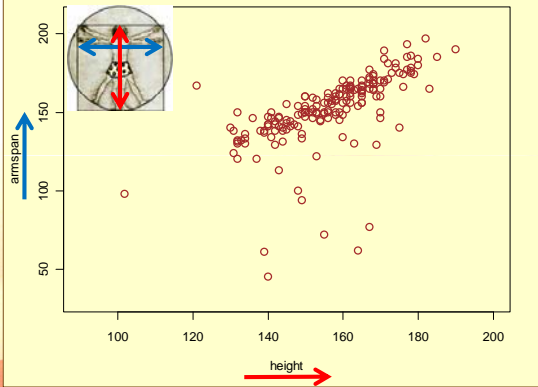
Leonardo da Vinci (1452-1519) was a scientist and an artist. He thought that the span of someone's arms was equal to their height. Why do you think he was interested in working out body proportions?

Do you think Leonardo's theories still work today?

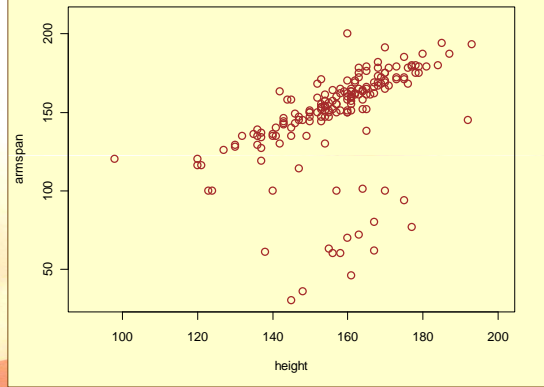
Focus: creating graphs and reasoning with them



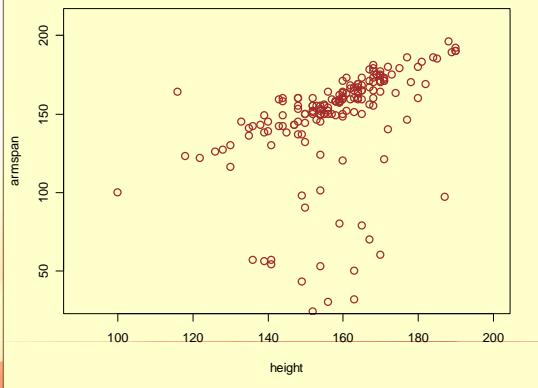
Armspan vs Height: Samples of size 200



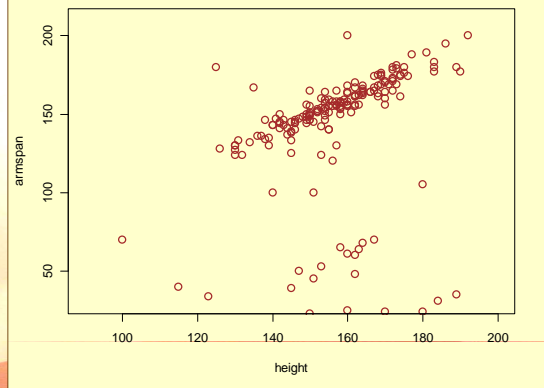
Armspan vs Height: Samples of size 200

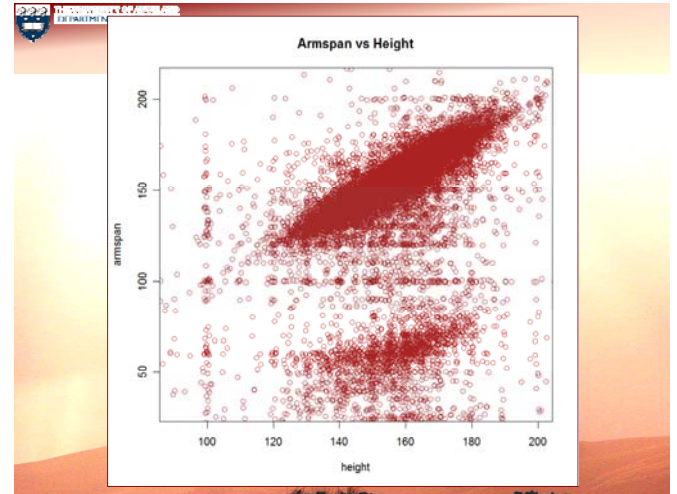
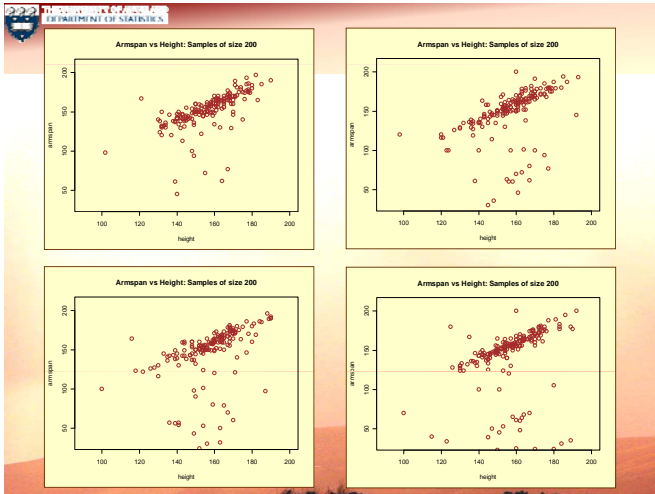


Armspan vs Height: Samples of size 200



Armspan vs Height: Samples of size 200





The nature of statistical inference

Description versus inference

- **Description** is *what I see* in the data in hand
 - Theme: “Right here, right now” – Fat Boy Slim
- **Inference** is what I think is likely to be happening *back in the populations*, back *where these data came from*
 - Theme: “Back in the USSR” – Beatles
 - We have a natural propensity to move early to inference
 - Many unclear in their thinking & communication when they are describing and when inferring

We will be concentrating on inference, but ...

To see the richness of the interplay between description and inference at work

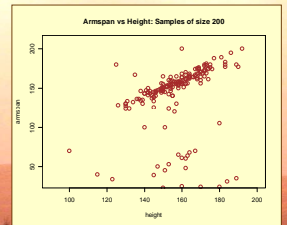
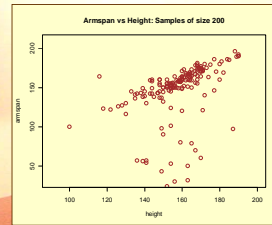
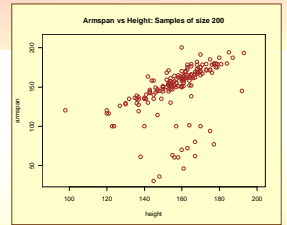
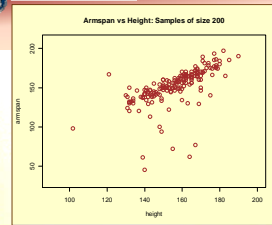
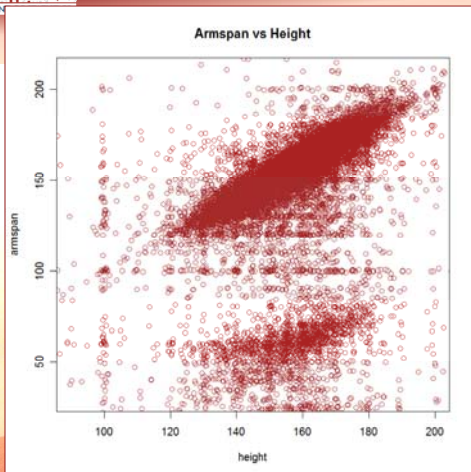
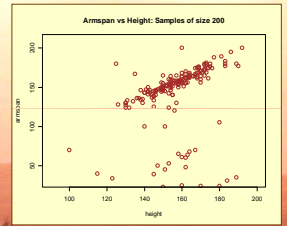
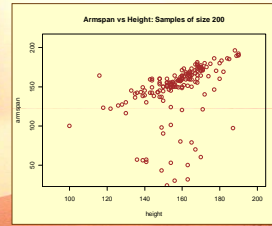
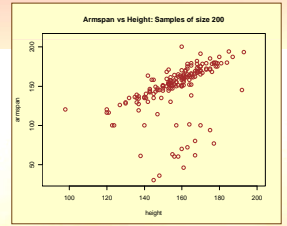
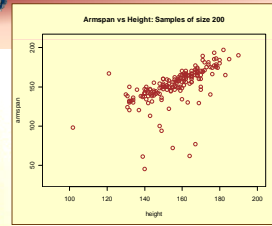
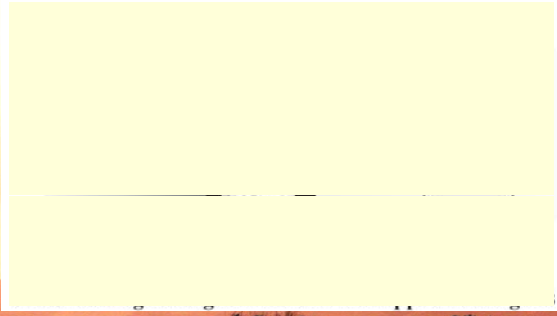
see Handout 2

How do we make inferences?

- Often from coming to believe that something I see in *these* data is a reflection of something occurring back in the populations
- Always know that what we see is, at best, an **imperfect reflection of the way it really is** back in the populations

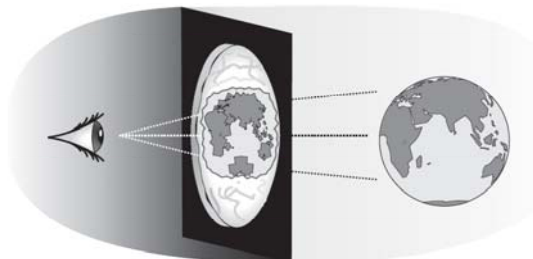
But ...
 "What I see ...
 is not quite the way it really is"

Looking at the world using data



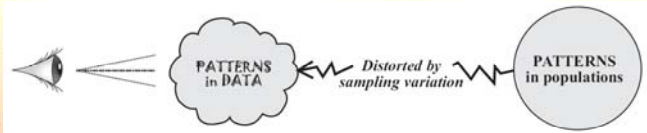
"What I see is not quite the way it really is"

Looking at the world using data

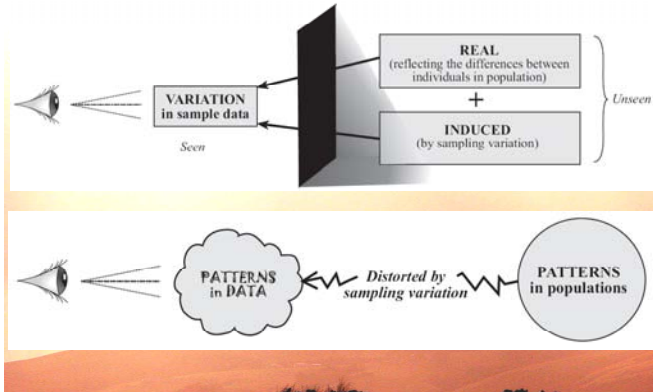


is like looking through a window with ripples in the glass

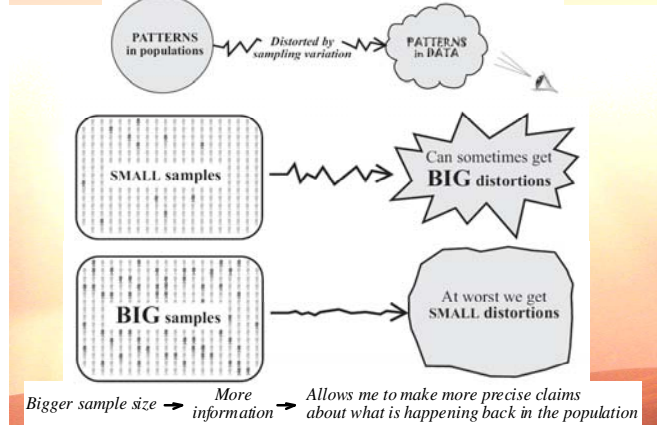
"What I see is not quite the way it really is"



“What I see is not quite the way it really is”

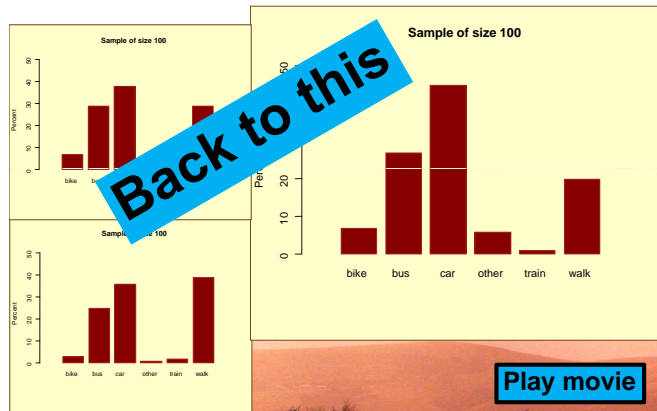


“What I see is not quite the way it really is”



Let's look at some sampling variation

How did they travel to school ?



Bar Chart Animations

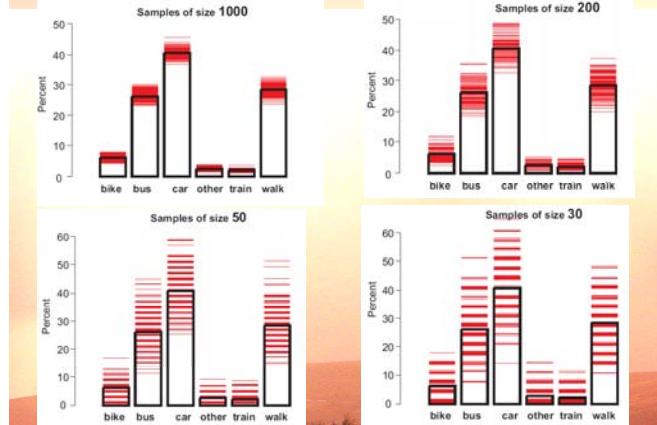


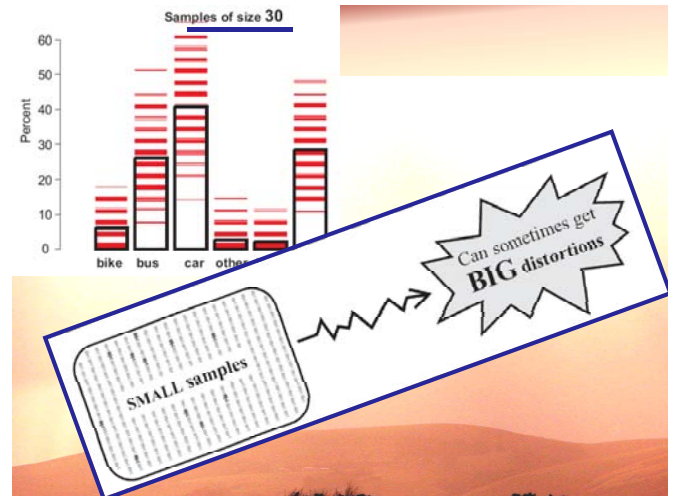
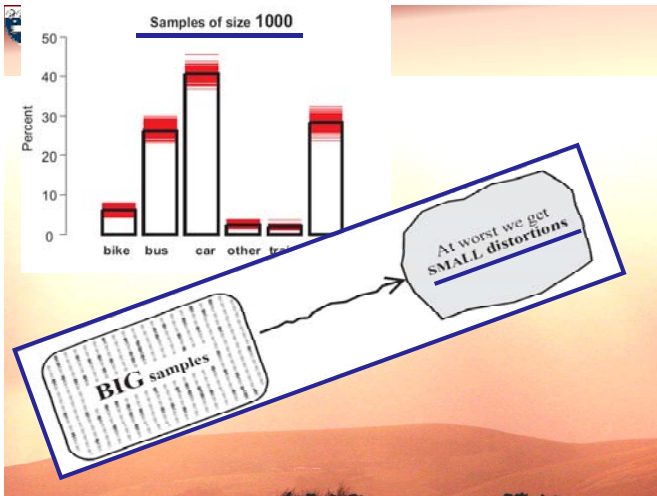
Play

- [Samples of 1000](#)
- [Samples of 200](#)
- [Samples of 100](#)
- [Samples of 50](#)
- [Samples of 30](#)
- [Samples of 30 without jitter](#)



“What I see is not quite the way it really is”





UNIVERSITY OF CALIFORNIA DEPARTMENT OF STATISTICS

"What I see is not quite the way it really is"

- What can we learn from proportions taken from samples of size 30?
 - Very little !!
- Information content of category data points
 - "Do you fall into this category? Yes/No" is very small
- Need very large samples before can say anything very useful
 - Unfortunate fact of life!
 - Situation better with measurement data

Samples of size 30

Percent

bike bus car other train walk

UNIVERSITY OF CALIFORNIA DEPARTMENT OF STATISTICS

Comparing heights of boys and girls at age 12

Back to this

Boys and girls aged 12 samples of size 30

Boys

Girls

80 100 120 140 160 180 200

Play movie

UNIVERSITY OF CALIFORNIA DEPARTMENT OF STATISTICS

Dot and Boxplot Animations

Play

- Original 2-sample
- Effect of sample size

Population distributions

Plots for samples of size 20 from different investigators

Population distribution: The unseen world

Samples of size

10 30 100

UNIVERSITY OF CALIFORNIA DEPARTMENT OF STATISTICS

Boxplots with a Memory I


Play 1-sample build-up, n=30

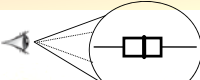
Population distribution: The unseen world

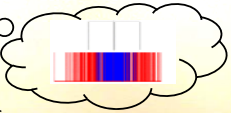
Plots for samples of size 30


100 120 140 160 180


Department of Statistics



Whenever I see ... 

I remember ... 


Mine could even be like this ... 

Or even this ... 

I must take this uncertainty about where it really should be into account when I make comparisons!

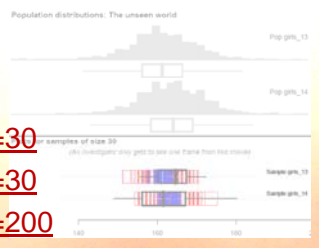
Department of Statistics

Boxplots with a Memory II




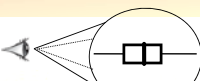
Play

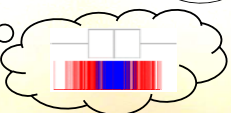
- 1-sample build-up, n=30
- 2-sample build-up, n=30
- 1-sample build-up, n=200
- 2-sample build-up, n=200





Department of Statistics



Whenever I see ... 

I remember ... 


Mine could even be like this ... 

Or even this ... 


I must take this uncertainty about where it really should be into account when I make comparisons!

Department of Statistics


Population distribution: The unseen world



Plots for samples of size 30
(An investigator only gets to see one frame from this movie)




Plots for samples of size 200
(An investigator only gets to see one frame from this movie)



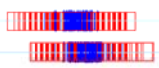
Height

Department of Statistics

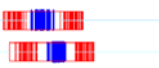
Population distributions: The unseen world



Plots for samples of size 30
(An investigator only gets to see one frame from this movie)



Plots for samples of size 200
(An investigator only gets to see one frame from this movie)

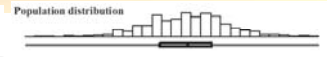


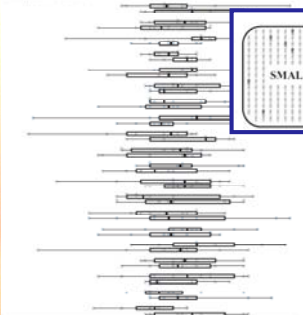
Height

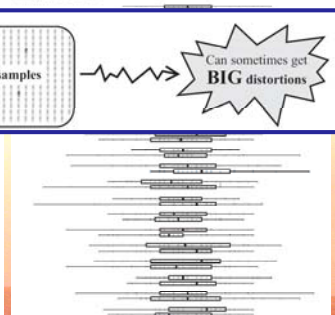
Department of Statistics

Heights of 12 year-olds

Examples of shifts called purely by sampling

Population distribution 

Samples of size 10 

Samples of size 30 

SMALL samples → Can sometimes get BIG distortions

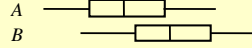
Samples of size 10 shows to demonstrate why we should not be working in this way

Handout 1, p6

“How to make the call” by Curriculum level

Handout 1, p4

Curriculum Level 5: the 3/4-1/2 rule



If the median for one of the samples lies outside the box for the other sample (“more than half of the B group are above three quarters of the A group”) make the claim “B tends to be bigger than A” back in the populations

[Restrict to samples sizes of between 20 and 40 in each group]

Majority of one to the right of “the great whack” of the other

Some notes about the rules

Handout 1, p5

Curriculum Level 5: the 3/4-1/2 rule

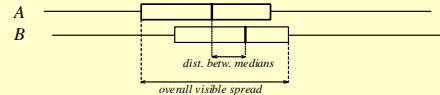
- The intuitive idea is that “the B group is bigger than the “great whack” of the A group”
- Technically, the rule is based on the fact that the median of B is greater than the 3/4 quantile of A. This is about 1/3 times in 100 for samples of size 20, 1/4 times in 100 for samples of size 30, 3 times in 100 for samples of 40, 1 times in 2,500 for samples of size 100.

See handout 1, p5 for discussion

“How to make the call” by Curriculum level

Handout 1, p. 4

Curriculum Level 6: distance between medians as proportion of “overall visible spread”



Make the claim “B tends to be bigger than A” back in the populations if distance between medians is greater than about ...

1/3 of overall visible spread for sample sizes of around 30

1/5 of overall visible spread for sample sizes of around 100

[Could also use 1/10 of overall visible spread for sample sizes of around 1000]

See Tech notes on p. 5

Stress “eye-ball judgements”

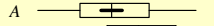
“How to make the call” by Curriculum level

Handout 1, p. 4

Curriculum Level 7: based on informal confidence intervals for the population median



Make the claim “B tends to be bigger than A” back in the populations



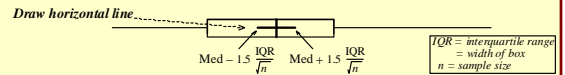
if these horizontal lines (intervals) do not overlap

See Tech notes on p. 5

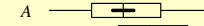
“How to make the call” by Curriculum level

Handout 1, p. 4

Curriculum Level 7: based on informal confidence intervals for the population median



Make the claim “B tends to be bigger than A” back in the populations



dist. = lower confidence limit for difference in population medians

dist. = upper confidence limit for difference in population medians

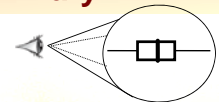
“How to make the call” by Curriculum level

Curriculum Level 8: on to formal inference

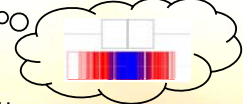
Quick Summary



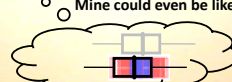
Whenever I see ...



I remember ...



Mine could even be like this ...



Or even this ...



I must take this uncertainty about where it really should be into account when I make comparisons!

Quick Summary

- **Sampling variation alone ...**
 - can produce shifts in our box plots
 - Small shifts with big samples
 - Sometimes quite big shifts with small samples
- **Makes no sense**
 - *to read meaning* into shifts in data of a size often produced by sampling variation
- **We have some rules** for signalling when a shift
 - is *big enough* that we can make a call on what group gives bigger values

Does the shift we see

look bigger than sampling variation would produce?

- **The rules**
 - Take sample size into account
 - Operated without taking the eyes off the data
 - Get more sophisticated over time
 - Converging towards the tools of formal inference

Is the data shift big enough?

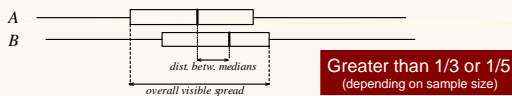
(for us to claim B bigger than A back in the populations)

Curriculum Level 5: the 3/4-1/2 rule



Majority of one to the right of "the great whack" of the other

Curriculum Level 6: distance between medians as proportion of "overall visible spread"



Curriculum Level 7: based on informal confidence intervals for the population median

Make the claim B tends to be bigger than A back in the populations



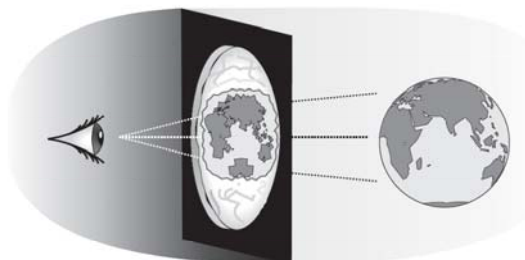
No overlap of constructed intervals

If the shift is not big enough ...

- **then we can't make a call ..** on "who is bigger" back in the populations?
 - Simply don't have enough information
- **Happens frequently when ...**
 - the **sample sizes** are small
 - very little data (very ripply window)
 - **differences between the populations** are small (looking for fine details rather than gross discrepancies)

But these are subjects for another talk

Looking at the world using data



is like looking through a window with ripples in the glass