- We use statistics to confirm effects, estimate parameters, and predict outcomes
- It usually rains when Im in Cape Town, but mostly on Sunday
- *Confirmation:* In Cape Town, it rains more on Sundays than other days
- *Estimation:* In Cape Town, the *odds* of rain on Sunday are 1.62.2 times higher than on other days
- *Prediction:* I am confident that it will rain at least one Sunday the next time I go

c

- How we interpret data like this necessarily depends on assumptions:
- Is it likely our observations occured by chance?
- Is it likely they *didnt*?

`![image](my`$_{i}mages/eight.jpg)height = $”0.8

¡span¿*Tessa Wessels, ¡span¿*Faces on a Train*¡/span¿*¡/span¿

- We measure the average heights of children raised with and without vitamin A supplements
- *Estimate:* how much taller (or shorter) are the treated children on average?
- *Confirmation:* are we sure that the supplements are helping (or hurting)?
- *Range of estimates:* how much do we think the supplement is helping?

Estimation ==========

- We use *P values* to say how sure we are that we have seen some effect
- We use *confidence intervals* to say what we think is going on (with a certain level of confidence)
- P values are *over-rated*
- *Never* use a high P value as evidence for anything, e.g.:
- that an effect is small
- that two quantities are similar
- We want to know if vitamin A supplements improve the health of village children
- Is height is a good measure of general health?
- How will we know height differences are due to our treatment?
- We want the two groups to start from the same point independent randomization of each individual
- We may measure *changes* in height
- Or *control for* other factors
- Is vitamin A good for these children?
- How sure are we?
- How good do we think it is?
- How sure are we about that?
- What does it mean if I find a significant P value for some effect in this experiment?
- The difference is unlikely to be due to chance
- So what! I already know vitamin A has strong effects on metabolism - If Im certain that the true answer isnt exactly zero, why do I want the P value anyway?

c

`![image](vitamins.Rout-0.pdf)`width=”100.00000%”

- What do these results mean?
- Which are significant?

c

`![image]`(vitamins.Rout-1.pdf)width=”100.00000%”
- A high P value means we cant see the sign of the effect clearly
- A low P value means we can
`![image]`(Lecture$_i mages/fog.jpg)height = $”0.8
`![image]`(Lecture$_i mages/clear.jpg)height = $”1
- More broadly, a P value measures whether we are seeing *something* clearly
- Its usually the sign ($\pm$) of some quantity, but doesnt need to be
- Type I (*False positive:*) concluding there is an effect when there isnt one
- This doesnt happen in biology. There is always an effect. - Type II (*False negative:*) concluding there is no effect when there really is
- This *should* never happen, because we should never conclude there is no effect
- Type I (*False positive:*) in the hypothetical case that the effect is exactly zero, what is the probability of falsely finding an effect
- Should be less than or equal to my significance value - Type II (*False negative:*) what is the probability of failing to find an effect that is there?
- Useful, but can only be asked for a specific hypothetical effect *size* - These are useful to analyze **power** and **validity** of a statistical design
- You should do these analyses *before* you collect data, not after

c

- *Sign error:* if I think an effect is positive, when its really negative (or vice versa)
- *Magnitude error:* if I think an effect is small, when its really large (or vice versa)
- Confidence intervals clarify all of this
`![image]`(Lecture$_i mages/error.jpg)width = $”100.00000%”

c

- If I have a low P value I can see something clearly
- But its usually better to focus on what I see than the P value
`![image]`(Lecture$_i mages/clear.jpg)width = $”100.00000%”

c

- If I have a high P value, there is something I *dont* see clearly
- It *may be* because this effect is small
- High P values should *not* be used to advance your conclusion
`![image]`(Lecture$_i mages/fog.jpg)width = $”100.00000%”
- Small differences
- Less data
- More noise
- An inappropriate model

- Less model resolution
- A lower P value means that your evidence for difference is better
- A higher P value means that your evidence for similarity is better or worse!

c

`![image](flu.Rout-0.pdf)`width="100.00000%"
- Why is weather not causing deaths at this time scale?

c

`![image](flu.Rout-1.pdf)`width="100.00000%"
- **Never** say: A is significant and B isnt, so $A > B$
- **Instead:** Construct a statistic for the hypothesis $A > B$
- May be difficult

c

- All men are mortal
- Jacob Zuma is mortal
- Therefore, Jacob Zuma is a man
`![image](Lecture`$_images/zuma.jpg)width = $"100.00000%"

c

- All men are mortal
- Fanny the elephant is mortal
- Therefore, Fanny is a man
`![image](Lecture`$_images/fanny.jpg)width = $"100.00000%"
- A lot of statistical practice works this way:
- bad logic in service of conclusions that are (usually) correct - This sort of statistical practice leads in the aggregate to bad science
- The logic can be fixed:
- Estimate a difference, or an interaction
- We cant build statistical confidence that something is small by failing to see it clearly
- We must instead see clearly that it is small
- This means we need a standard for what we mean by small
`![image](Lecture`$_images/N95.jpg)height = $"0.8
`![image](Lecture`$_images/surgical.jpg)height = $"0.7
- People who work in respiratory clinics sometimes have to wear bulky, uncomfortable, expensive masks
- They would like to switch to simpler masks, if those will do the job
- How can this be tested statistically? We dont want the masks to be different.
- Use a confidence interval
- Decide how big a level is acceptable, and construct a P value for the hypothesis that this level is excluded!
`![image](masks.Rout-2.pdf)`height="0.8

c

`![image](masks.Rout-0.pdf)width=”100.00000%”`

- Is the new mask good enough?

- Whats our standard for that?

c

`![image](masks.Rout-1.pdf)width=”100.00000%”`

- We can even attach a P value by basing it on the right” statistic.

- The right statistic is the thing whose sign we want to know:

- The difference between the observed effect and the standard we chose

Paradigms for inference ========================

Frequentist paradigm

- Make a null model

- Test whether the effect you see could be due to chance

- What is the probability of seeing exactly a 1.52 cm difference in average heights? - Test whether the effect you see ¡span¿*or a larger effect*¡/span¿ could be due to chance

- This probability is the P value

`![image](vitamins_plot.Rout − 0.pdf)height = ”0.8`

`![image](vitamins_plot.Rout − 1.pdf)height = ”0.8`

`![image](vitamins_plot.Rout − 2.pdf)height = ”0.8`

`![image](vitamins_plot.Rout − 3.pdf)height = ”0.8`

`![image](vitamins_plot.Rout − 4.pdf)height = ”0.8`

`![image](vitamins_scramble.Rout.pdf)height = ”0.8`

Bayesian paradigm

c

- Make a complete model world

- Use conditional probability to calculate the probability you want

`![image](Lecture_images/shanghai.jpg)width = ”100.00000%”`

c

- More assumptions $\implies$ more power

- With great power comes great responsibility

`![image](Lecture_images/spiderman.jpg)width = ”100.00000%”`

- We want to go from a *statistical model* of how our data are generated, to a probability model of parameter values

- Requires *prior* distributions describing the assumed likelihood of parameters before these observations are made

- Use Bayes theorem to calculate posterior distribution likelihood after taking data into account

c

- A frequentist can do a clear analysis right away
- A Bayesian needs a ton of assumptions will try to make uninformative assumptions
![image]$(\text{vitamins}_plot.Rout - 0.pdf)width = $"100.00000%"

$$c$$

- Frequentist: how unlikely is the observation, from a random perspective?
- Bayesian: whats my model world? What is my prior belief about weather-weekday interactions.
![image]$(\text{my}_images/eight.jpg)height = $"0.8
¡span¿*Tessa Wessels, ¡span¿*Faces on a Train*¡/span¿*¡/span¿

Conclusion ==========

- Statistics are not a magic machine that gives you the right answer
- If you are to be a serious scientist in a noisy world, you should have your own philosophy of statistics
- Be pragmatic: your goal is to do science, not get caught by theoretical considerations
- Be honest: its harder than it sounds.
- You can always keep analyzing until you find a significant result
- If you do this you will make a lot of mistakes - You may also keep analyzing until you find a result that you already know is true.
- This is confirmation bias; youre probably right, but your project is not advancing science - Good practice
- Keep a data-analysis journal
- Start *before* you look at the data