Inferential statistics

Statistics has 3+ components

- Data description & analysis
- Probability calculations
- Statistical inference
  - Inferential statistics
- Models ....

Hypothesis testing

Why test for differences?

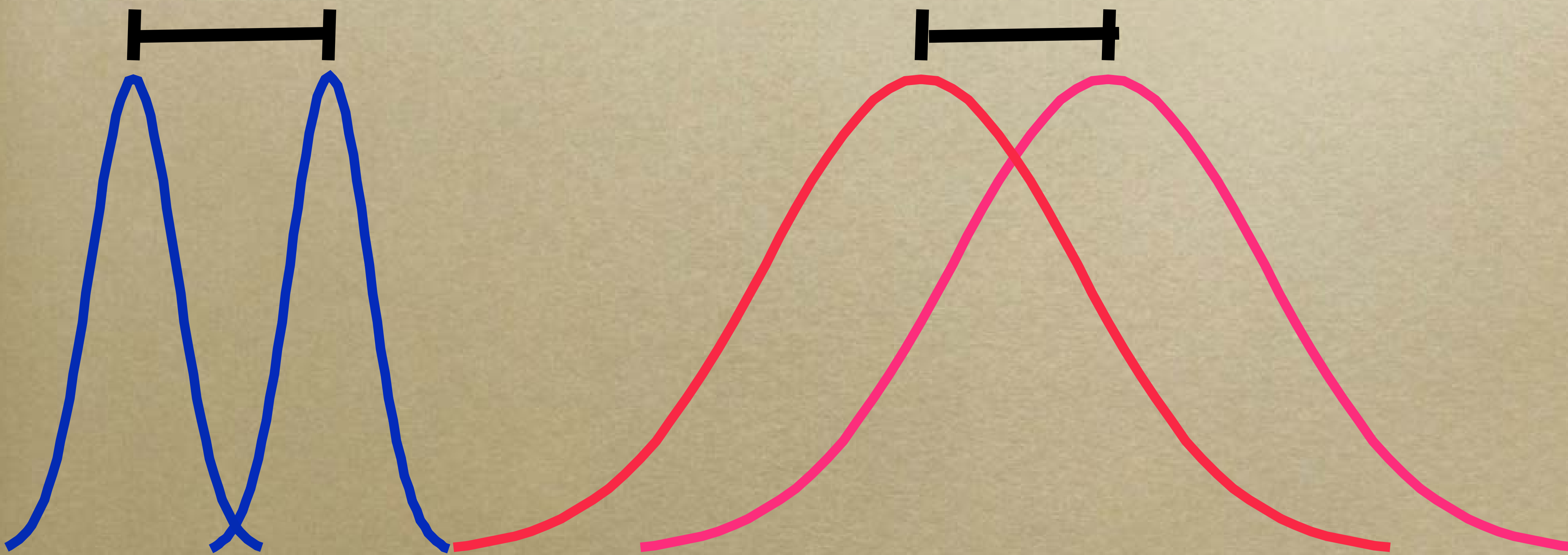
- If we find that females are rated as 8 on being nice and males are rated as 7 -- are they different?
- After all different things are different !
  - Why test these differences?

Why test for differences?

- Two main reasons:
  - Variance (measurement error, random error, other variables)
  - Making inferences beyond the sample to the population at large

Variance

We are living in a stochastic world



Hypotheses testing I

- How can we test that things are different?
  - Set 2 hypotheses that cover the entire possible range of outcomes
- H0 & H1
  - H0- No difference Group 1 [=, ≤, ≥] Group 2
  - H1-Difference Group 1 [≠, >, <] Group 2
-
- Examples

Hypotheses testing Ib

- H0 & H1-- Examples:
  - Is a coin fair
  - Gender and grades
  - Healing with a new medication
  - Ability to cheat
  - Marriage over time
    -
- For each please write H0 & H1

Hypothesis testing II

- Why test a hypothesis we don't believe in?
  - We do this because we can only show that something is wrong -- not that something is right
- What does it mean to reject H0?
- If H0 is correct, the probability of getting this result (or a more extreme result) is very low -- thus we reject H0 and (for now) accept H1

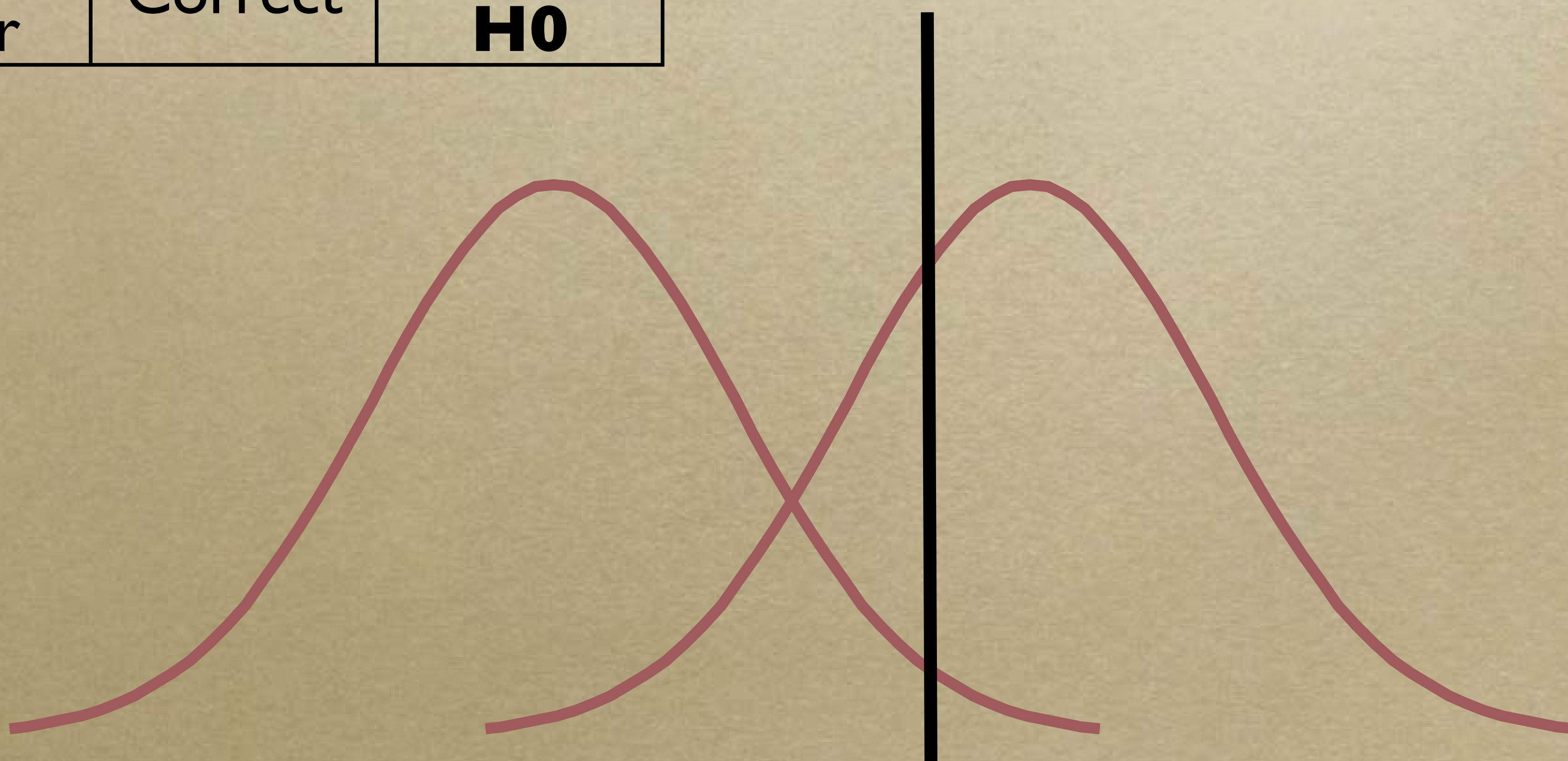
2 types of errors

- Not conservative and liberal just balancing 2 types of error

H0 is wrong	**H0 is correct**	
Correct	Type I error	**Reject H0**
Type 2 error	Correct	**Accept H0**

2 types of errors

H0 is wrong	H0 is correct	
Correct	Type I error	**Reject H0**
Type 2 error	Correct	**Accept H0**



The meaning of p

- What does p means?
- What is the difference between
  - $p = 0.03$, $p = 0.001$, & $p = 0.11$
- What is the relationship between p and confidence?
- What is the relationship between p and effect size?
- What is the relationship between p and number of subjects?

The importance of effect size

- Always give effect size measures
  - Mean difference
  - Quartile differences
  - etc.

Summary

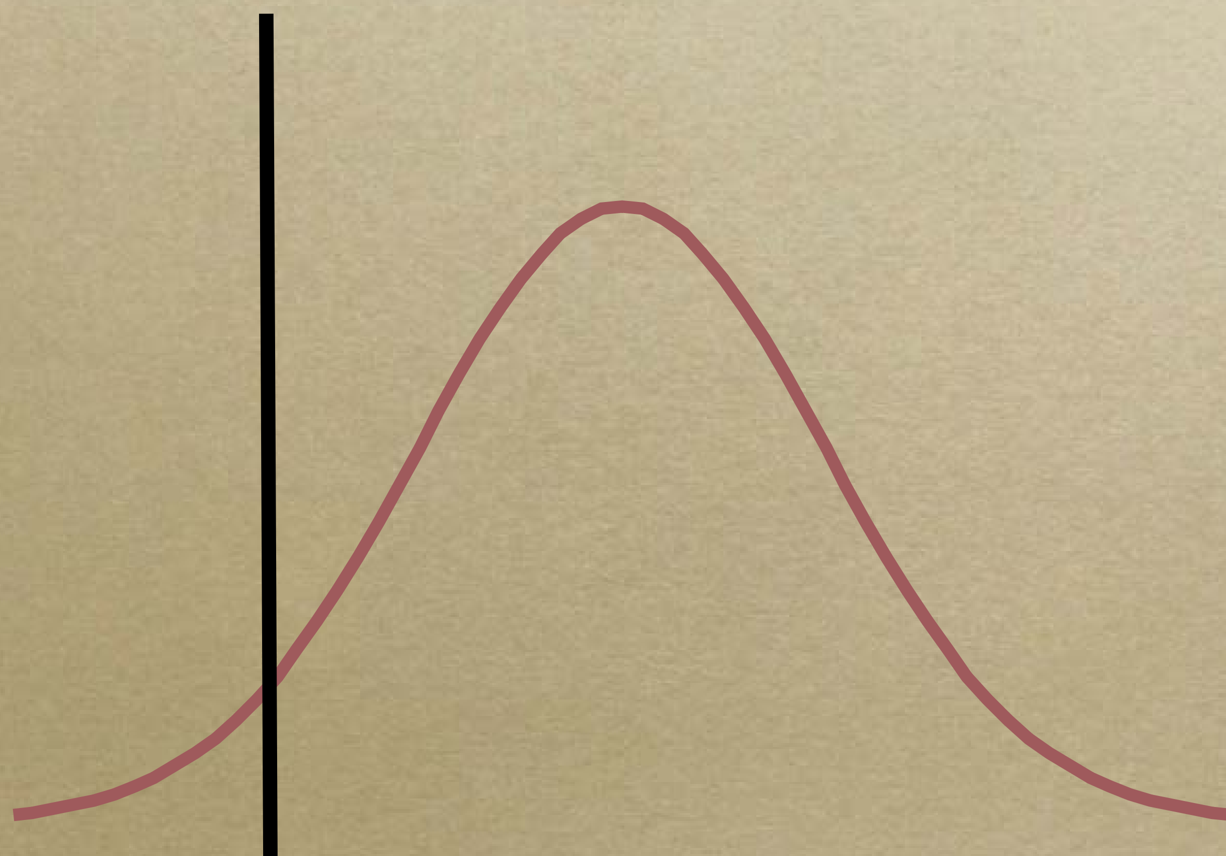
- Hypotheses testing
  - H1 & H0 -- setting the hypothesis to something you don’t believe in
  - The meaning of p
  - 2 types of errors
  - Effect size !

Statistical tests

- T-test for 1 sample
- T-test for 2 samples
- ANOVA
- Linear Regression
- Non-parametric tests

T test for 1 sample

One sample t test



One sample t test

Difference from the comparison

$$t = \frac{\mu - M}{\sqrt{\frac{\sum(xi - \mu)^2}{n-1}} / \sqrt{n}}$$

Standard deviation

Confidence

What do you do with “t”

- Compare it to the “t table”
-
- When there is more data, the t distribution gets closer to normal

Example step 1

Observation	Aggressive	xi - μ	(xi - μ)^2
1	24	4	16
2	22	2	4
3	23	3	9
4	18	-2	4
5	17	-3	9
6	16	-4	16
7	20	0	0
Sum	140	0	58

Example step 2

- *H0: average is 16*
- *H1: average ≠16*

$$\sigma = \sqrt{\frac{\Sigma(xi - \mu)^2}{n-1}} = 3.11$$

$$t = \frac{\mu - M}{\sigma / \sqrt{n}} = 3.42$$

T test for 2 samples

two samples t test

Test for dependent samples

$$t = \frac{\text{(between diff) – (expected diff)}}{\text{sd of diff} \;/ \sqrt{n}}$$

two samples t test

Test for 2 independent samples

Mean difference Hypothesized difference

$$t = \frac{(\mu 1 - \mu 2) - (M1 - M2)}{\sqrt{\frac{n1\,\sigma 1^2 + n2\,\sigma 2^2}{n1 + n2 - 2}\left(\frac{n1 + n2}{n1 \times n2}\right)}}$$

Example 1

- Who eats more lollipops males of females?
- 7 females; 5 males followed for a month
  - Females: $\mu = 27, \sigma^2 = 29.2$
  - Males: $\mu = 19, \sigma^2 = 24.57$
  -
- Is there a difference?

Calculating ...

$$t = \frac{(27 - 19) - (0)}{\sqrt{\dfrac{5 \times 24.57 + 7 \times 29.2}{5 + 7 - 2}\left(\dfrac{5 + 7}{5 \times 7}\right)}}$$

$$= 2.42$$

Example 2

- Does the sun creates freckles?
- Each subject has one side of their body in the sun and one in the shade
-
- H0 sun side ≤ non-sun side
- H1 sun side > non-sun side

Data

Subject	sun	shade	diff	d - μ	$(d - \mu)^2$
1	6	8	-2	-3	9
2	12	5	7	6	36
3	3	2	1	0	0
4	4	6	-2	-3	9
5	7	0	7	6	36
6	9	10	-1	-2	4
7	4	4	0	-1	1
8	0	2	-2	-3	9
9	4	3	1	0	0
Sum			**9**	**0**	**104**

Calculating ...

$$\sigma = \sqrt{\frac{104}{8}} = 3.606$$

$$t = \frac{(1) - (0)}{3.606/\sqrt{9}} = 0.831$$

t test summary

- t test as an example of inferential statistics
- Mean differences relative to variance

ANOVA

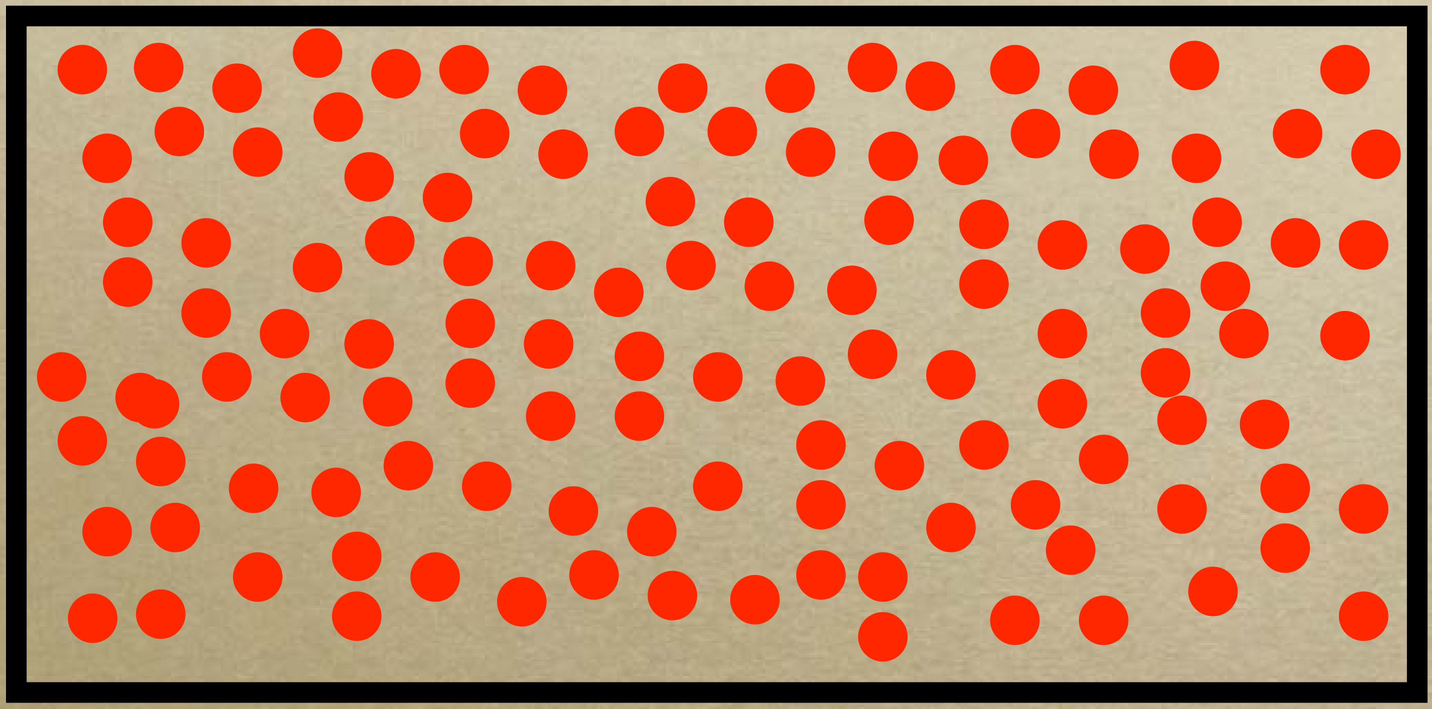
ANOVA I

- *Analysis of variance*
- *This is the same logic of the t test but allowing for more tests*
  - *So why not just do multiple t tests?*
  - *1 - doing many tests can cause errors*
  - *2 - there is benefit in pooling observations across cells*

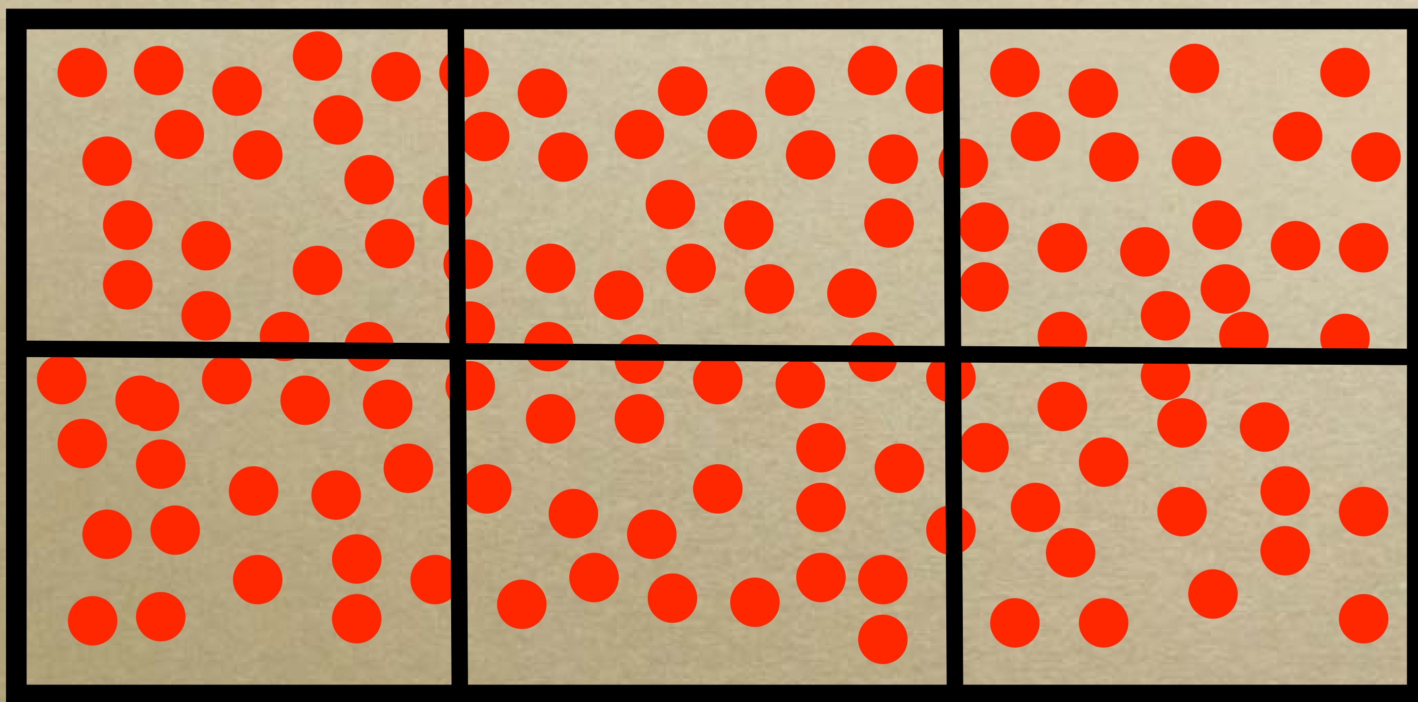
ANOVA II

- *The story is the same ...*
- *Looking at the variance within cells relative to the variance across cells (as if there were no treatments) and asking how much does the distinction between cells helps reduce the variance*

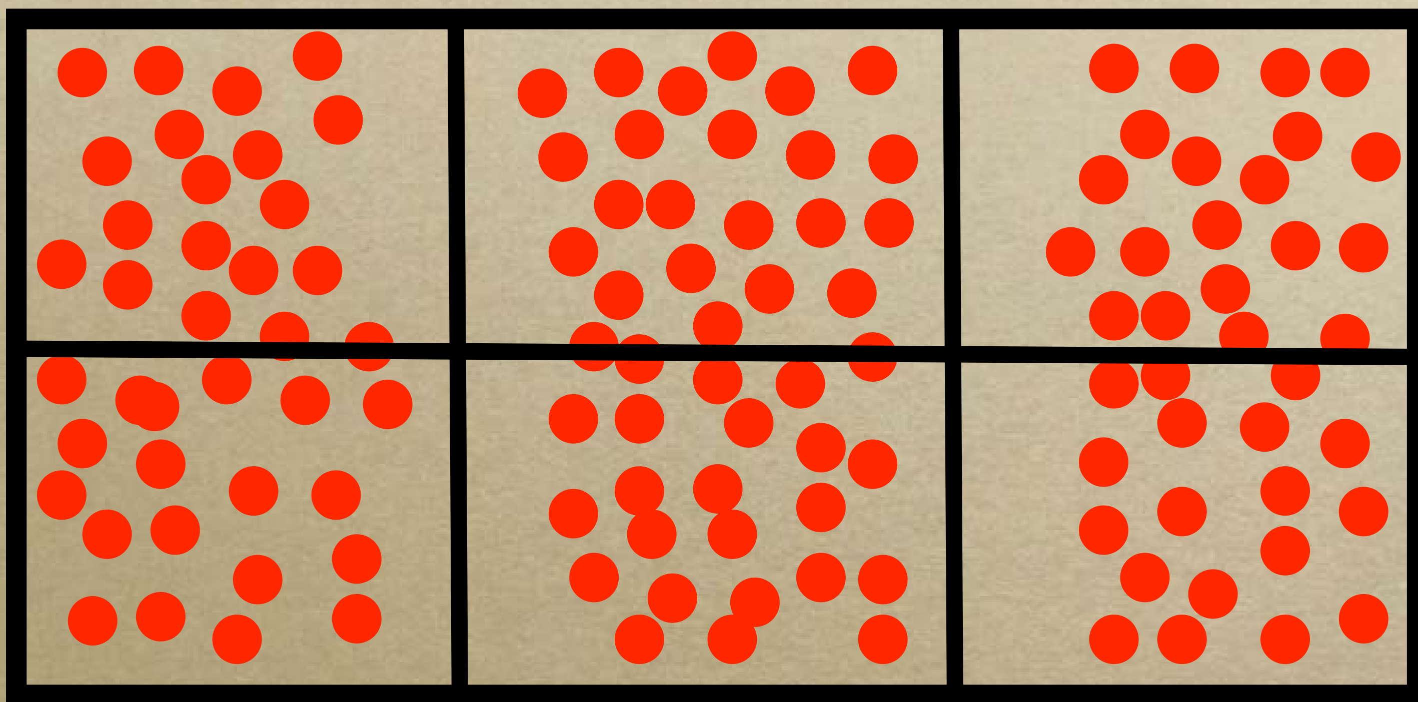
ANOVA IIa



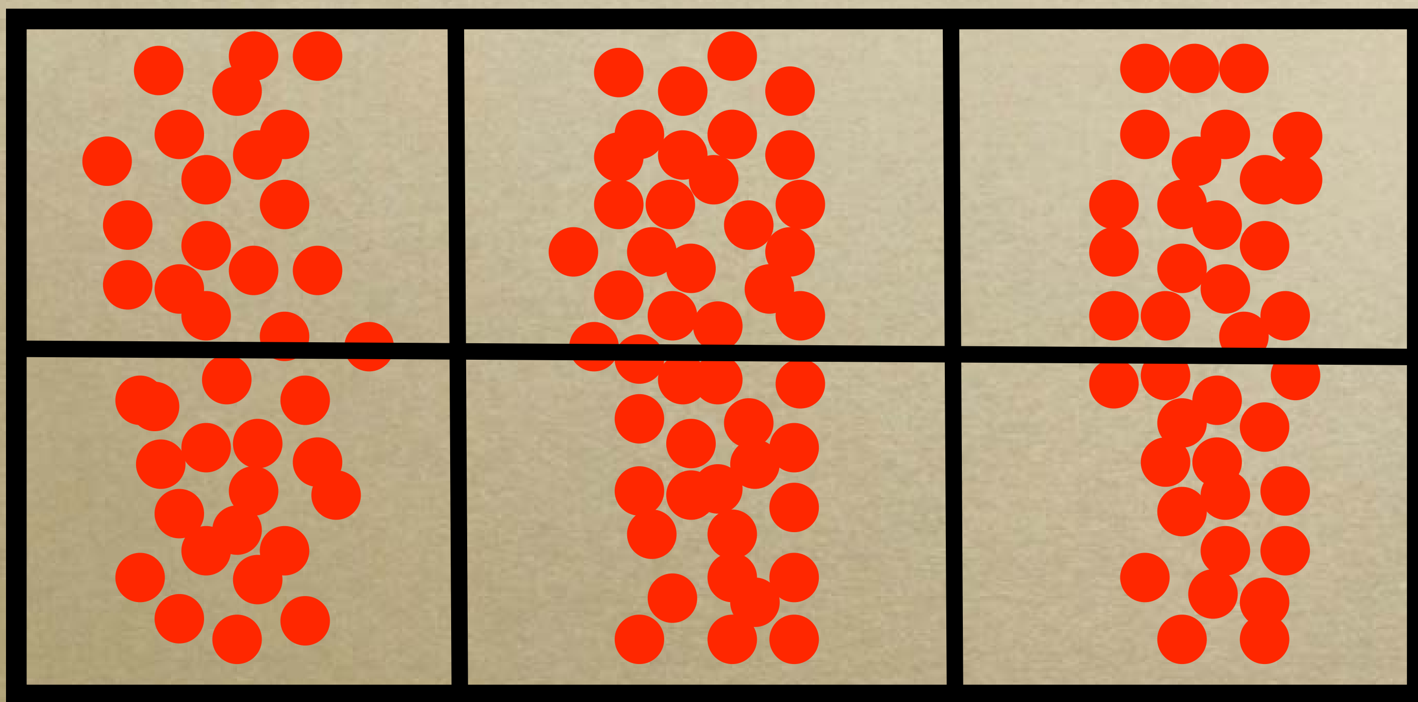
ANOVA IIb



ANOVA IIc



ANOVA IIc



A few examples I

- *No variance within or between groups*

	Group			
Case#	A	B	C	OA mean
1	40	40	40	
2	40	40	40	
3	40	40	40	
4	40	40	40	
Mean	**40**	**40**	**40**	**40**

A few examples II

- *No variance within groups, but variance between groups*

	Group			
Case#	A	B	C	OA mean
1	38	40	36	
2	38	40	36	
3	38	40	36	
4	38	40	36	
Mean	**38**	**40**	**36**	**38**

A few examples III

- *Variance within groups, but not between groups*

	Group			
Case#	A	B	C	OA mean
1	36	46	36	
2	40	37	38	
3	44	37	43	
4	40	40	43	
Mean	**40**	**40**	**40**	**40**

A few examples IV

- *Variance within & between groups*

	Group			
Case#	A	B	C	OA mean
1	38	45	36	
2	39	44	37	
3	39	49	35	
4	36	46	36	
Mean	**38**	**46**	**36**	**40**

The general formula

- *SS = Sum of squares*
- *SS total = SS within + SS between*
  - *SSt=SSw+SSb*
- *This means:*
  - *Take each of the samples, subtract it from the appropriate mean, and square it*

Sum of squares

- Once we have the sum of squares we compare the SSb and SSw
- As the SSb relative to SSw gets larger the results are more likely to be significant

Degrees of freedom

- This is a base to think about the amount of independent observations we have, and thus the strength of the results
- In general we lose a degree of freedom when we use a mean...

The formula

$$F = \frac{SSb / dfb}{SSw / dfw} = \frac{SSb / K-1}{SSw / n-k}$$

Once you have the F value use the F table with the correct df

ANOVAs

- *One way & multiple way ANOVAs*

	M	F
DV		

	M	F
E		
~E		

Summary I

- Hypothesis testing
  - Because we can never prove anything and only disprove things we set a hypothesis (H0) as one we do not believe in.
  - Once we reject H0 we are willing to accept H1

Summary II

- T test and ANOVA
  - It is all about variance and mean differences
  - How large is the mean difference relative to the variance!
  - These tests give us the probability of the data given that H0 is correct