#### Analysis of Variance Psychology 3256

#### Introduction

- We have t and z tests to deal with differences with one or two groups
- what if we have more than two groups?

## an example

	AI	A2	A3
	85	67	52
	90	80	60
	77	75	65
Ā	84	74	59

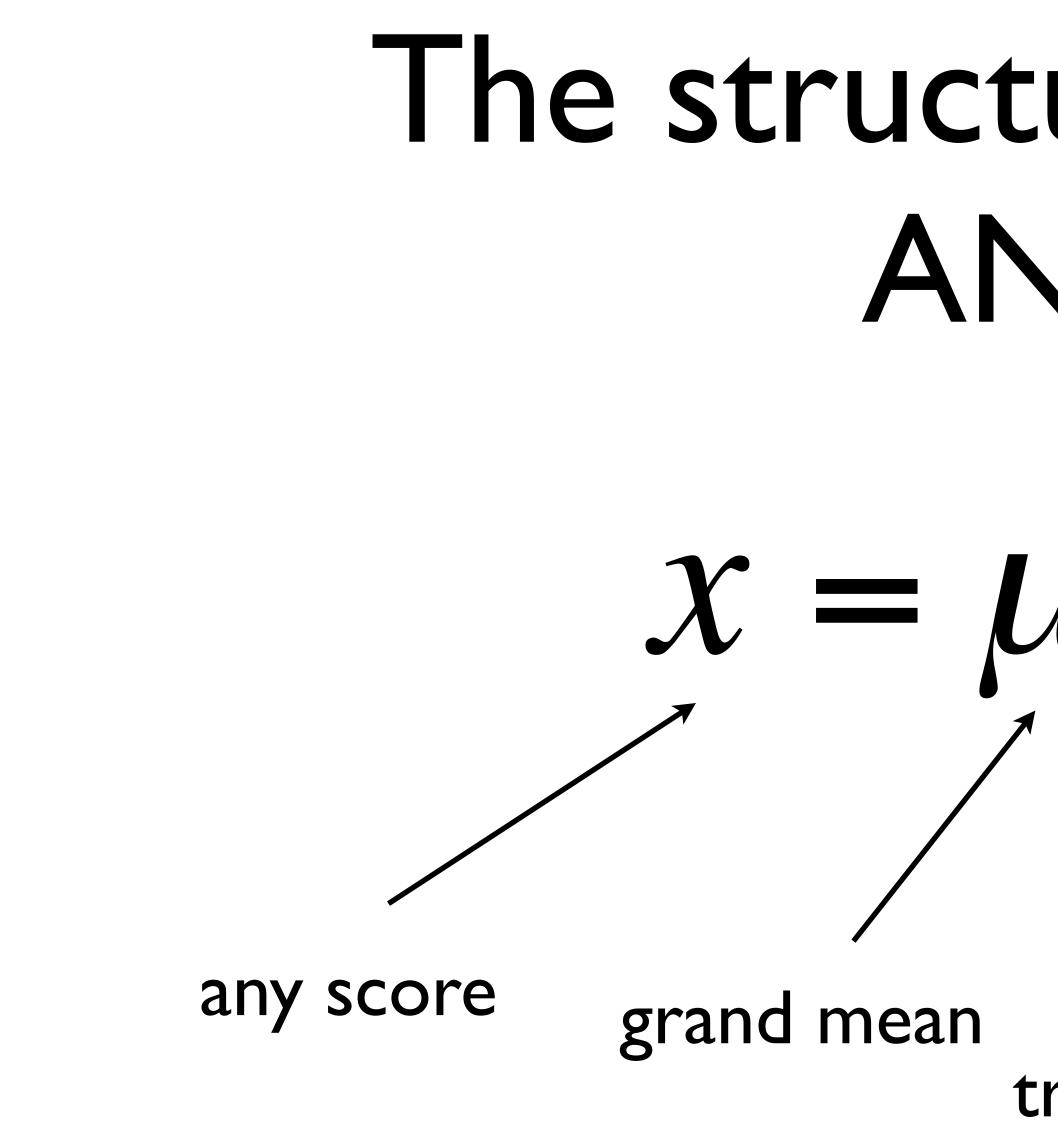
# Why do the scores vary?

- Or, what are the sources of variation
- Well individual difference
- and of course group differences

### I never said there'd be no math..

## any score = being human + group differences + individual difference

•  $x = \mu + \tau + \varepsilon$ 



# The structural model of ANOVA

# $x = \mu + \tau + \varepsilon$

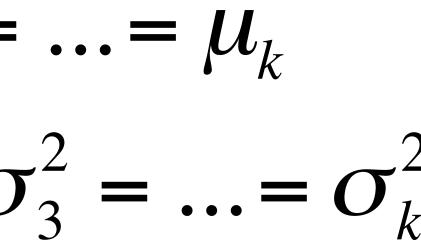
treatment effect

error

## Let's make an assumption

 $\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$  $: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = ... = \sigma_k^2$  $H_0$ true

#### This is the null hypothesis assumption



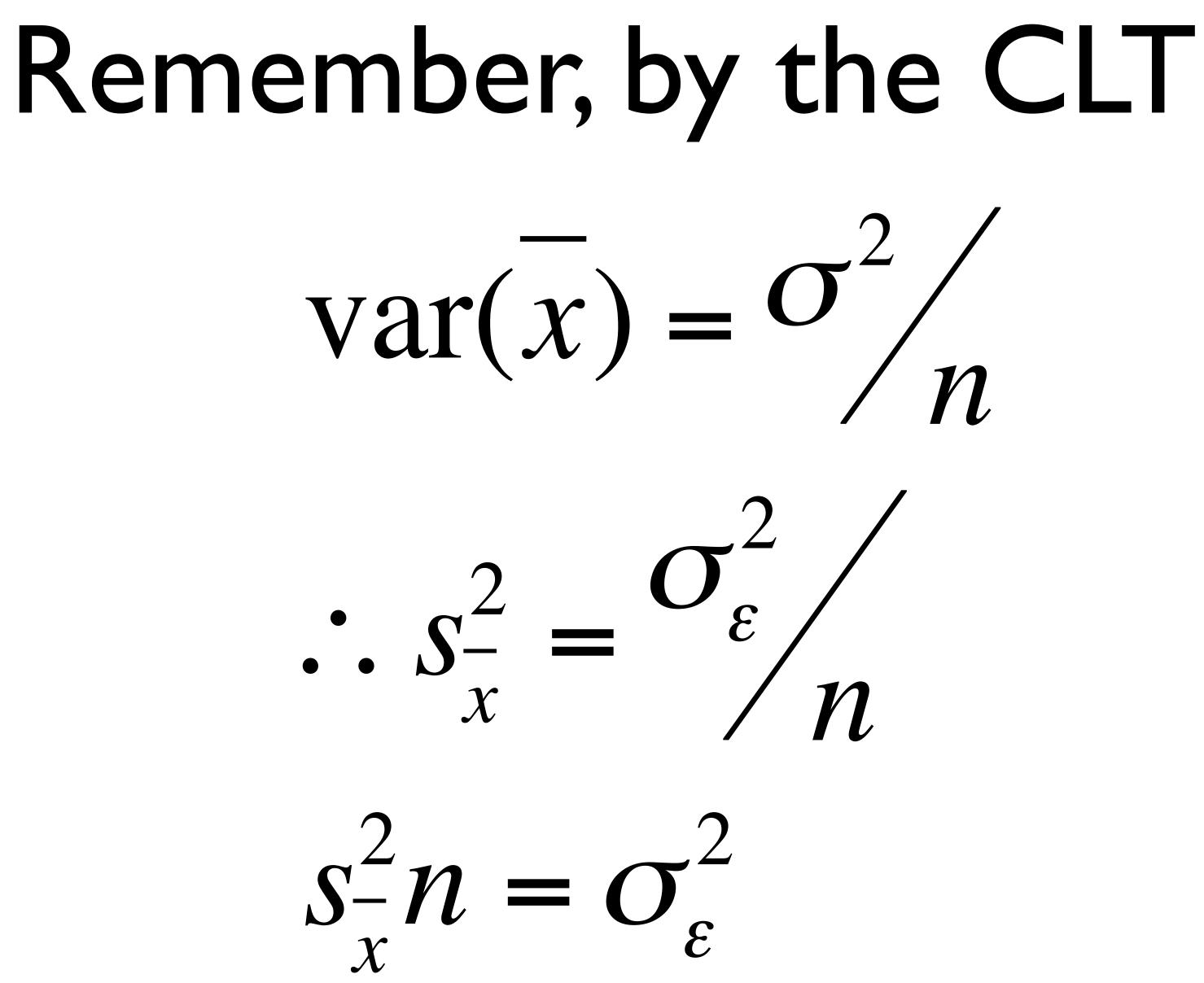
## More assumptions

- Scores are randomly and normally distributed around the grand mean
- Independent observations
- all sources of variation are in the model

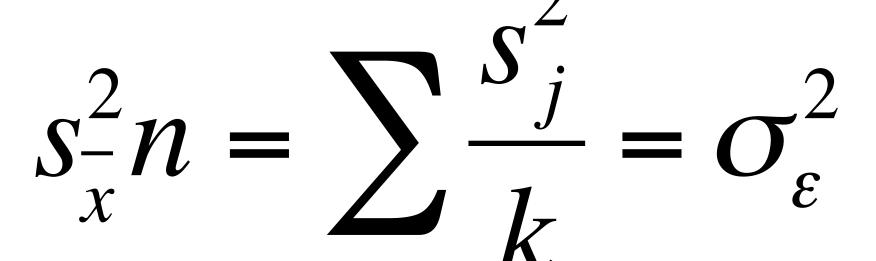
 $\sigma_1^2 \approx s_1^2$  $\sigma_2^2 \approx s_2^2$ etc

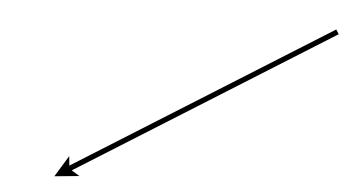
#### Let's look at variance

 $\sigma_{\varepsilon}^2 \approx s^2 = \overline{s}_j^2 = \sum \frac{s_j^2}{k}$ 

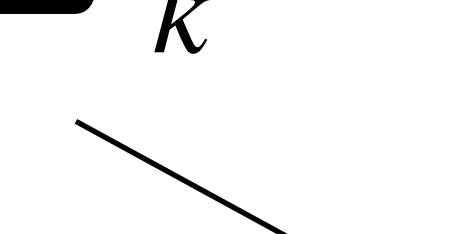


# We now have two estimates of $\sigma_{\varepsilon}^2$





**MS**<sub>Treat</sub>





## • E(MSE) = $\sigma_{\varepsilon}^{2}$ • E(MST) = $\sigma_{\varepsilon}^{2}$

#### So, if Ho is true..

- E(MSE) =  $\sigma_{\varepsilon}^{2}$  E(MST) =  $\sigma_{\varepsilon}^{2} + n\sigma_{\tau}^{2}$
- E(MSE) < E(MST)

#### If Ho is not true



- extra variation MST is measuring
- i.e. T

#### **SO...**

#### • If we were to divide MST by MSE (MST/MSE) we would have some estimate of how much

• This is precisely what is done in ANOVA

#### • F = MST / MSE

- E(F|Ho true) ?
- E(F|Ha true) ?

#### The F word

- If Ho is true, then MST/MSE will be distributed as  $F(df_t, df_e)$
- if not it will be distributed some other way
  - i.e. it will be unlikely to be distributed that way, or p < .05 (or whatever your alpha is)

#### The F word

## Partitioning SS and df

#### • SSTotal = SSTreatment + SSError

• dfTotal = dfTreatment +dfError

nent + SSError ent +dfError

 $\sum (x - \bar{x}_g)^2 = n \sum (\bar{x}_j - \bar{x}_g)^2 + \sum \sum (x - \bar{x}_j)^2$ N - 1 k - 1 N - k

More Precisely..

### ANOVA Summary Table

Source of Variation	df	MS	F
Between Groups	<b>k-I</b>	SSBG/k-I	MSBG/ MSWG
Within Groups	N-k	SSWG/N-k	
TOTAL	N-I		

#### Conclusions

- Easy to do
- Only tells you that two means are different, not which two, or three or whatever
- It's about the pattern overall
- You can figure that out, we'll do that next week.