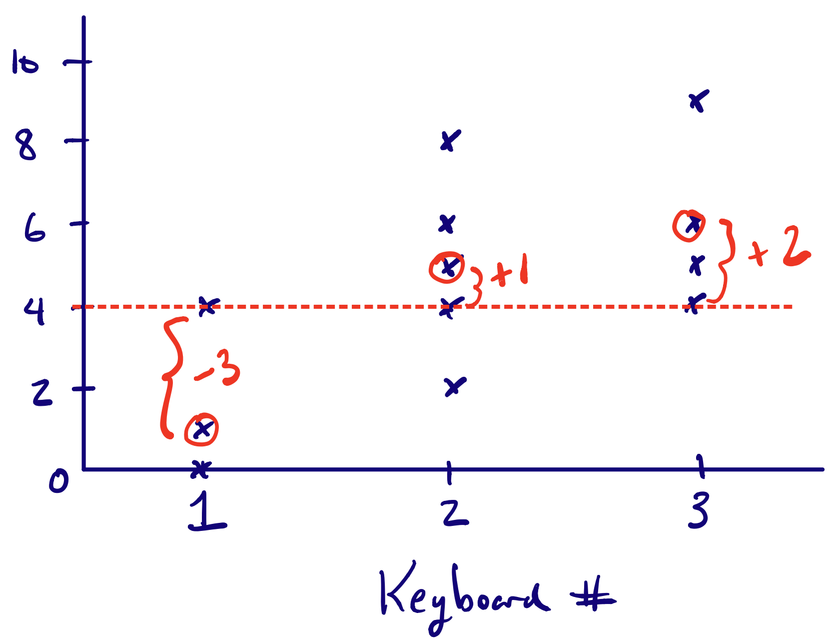
Lecture 4 – Repeated Measures Designs.

From last time: An I/O psychologist studied three computer keyboard designs. Subjects were randomly assigned to type on one of the three designs, and the # of errors was recorded.

Keyboard 1	Keyboard 2	Keyboard 3
0	6	6
4	8	5
0	5	9
1	4	4
0	2	6
$M = 1$	5	6
d_j -3	+1	+2

$\bar{X} = 4$



General principle of experimental design:

Data are nonsense without an underlying model to impose structure

Structure = "data story"

↳ how can we generate each data point?

Linear model:

$$X_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

Translation: each score X_{ij} can be obtained by:

(1) starting with the grand mean μ

(2) adding the relevant treatment effect α_j

(3) adding the remaining residual error ε_{ij} that is not already accounted for by μ and α_j

So I can write the full model of these data

as:

$$X_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

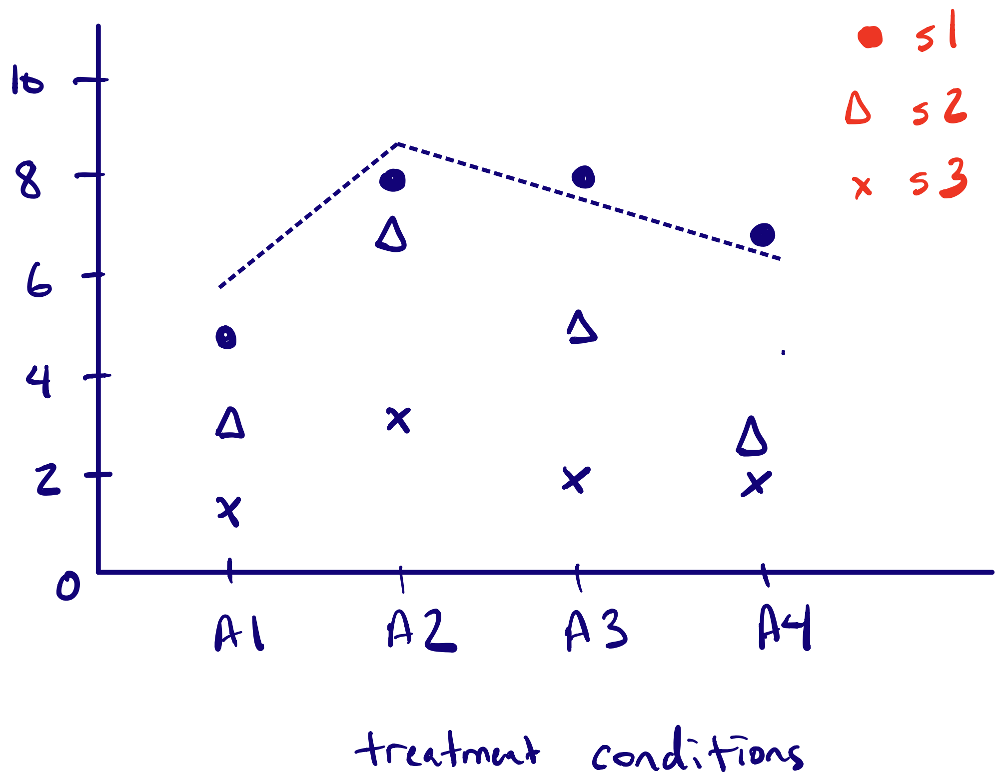
$$= 4 + \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix} + \varepsilon_{ij}$$

In this lecture, we consider designs where participants are measured multiple times

↳ "repeated measures designs"

Ex:

	treatment conditions			
subject	A1	A2	A3	A4
s1	5	8	8	7
s2	3	7	5	3
s3	1	3	2	2



What is our data story?

- not only is there variability between treatments, but there is also variability between subjects

- let us obtain each data point X_{ij} as:

  (1) grand mean μ

  (2) plus a treatment effect α_j

  (3) plus a subject effect π_i

  (4) plus residual error

<u>Model:</u>

$$X_{ij} = \mu + \alpha_j + \pi_i + \varepsilon_{ij}$$

let's estimate these effects from our data.

Estimating our model:

subject	A1	A2	A3	A4	M	π_i
s1	5	8	8	7	7	+2.5
s2	3	7	5	3	4.5	0
s3	1	3	2	2	2	−2.5
M	3	6	5	4	4.5	
α_j	−1.5	+1.5	+0.5	−0.5		

(treatment conditions: A1–A4)

So we can write

$$X_{ij} = \mu + \alpha_j + \pi_i + \varepsilon_{ij}$$

$$= 4.5 + \begin{bmatrix} -1.5 \\ 1.5 \\ 0.5 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 2.5 \\ 0 \\ -2.5 \end{bmatrix} + \varepsilon_{ij}$$

Hypothesis testing:

$\mathcal{H}_0: \alpha_j = 0$ ← "constrained model"

$\mathcal{H}_1:$ some $\alpha_j \neq 0$

Which better predicts the observed data?

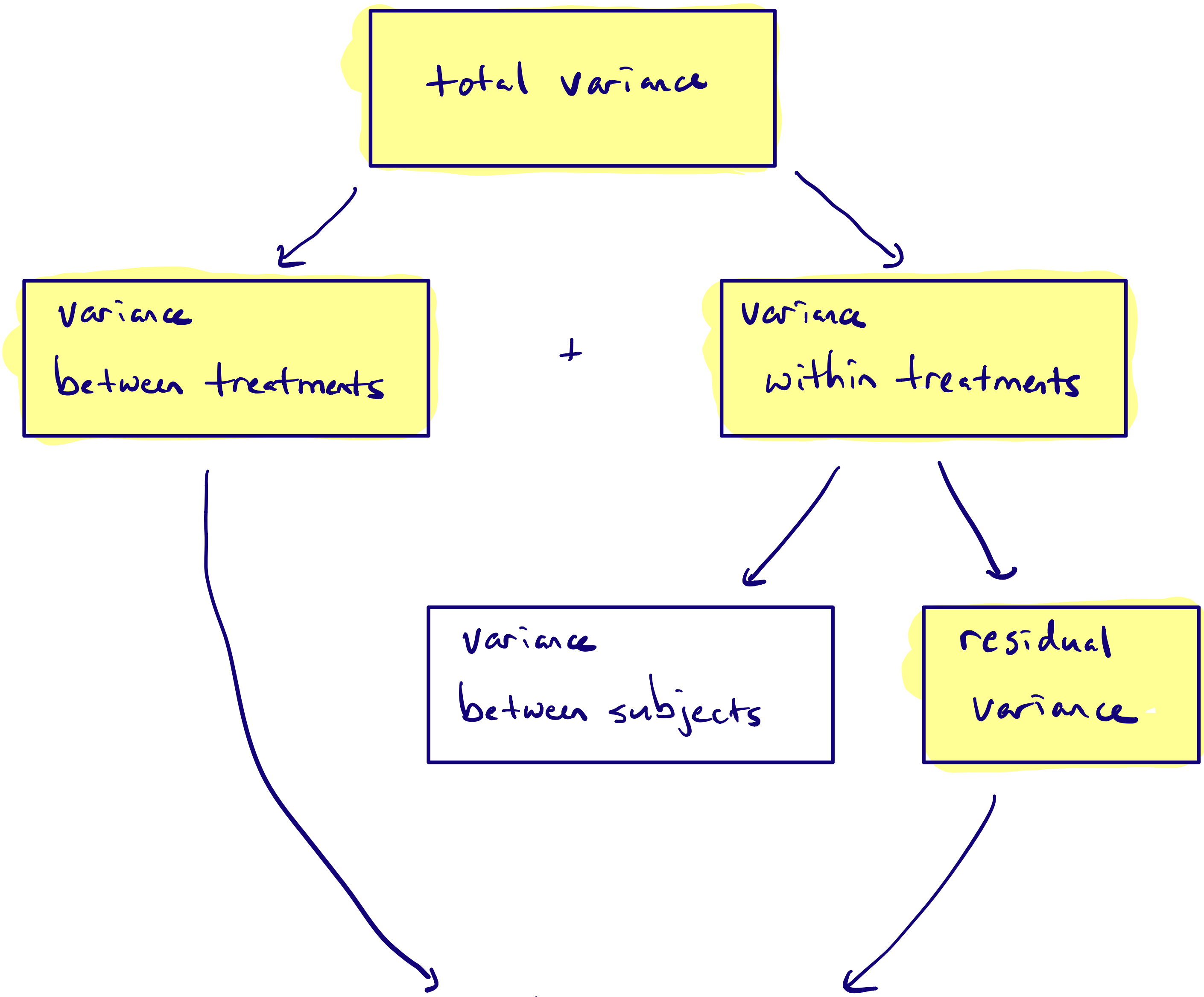
Recall: ANOVA works by partitioning variance

total variance

variance between treatments + variance within treatments

$$F = \frac{\text{variance between treatments}}{\text{variance within treatments}}$$

In repeated measures designs, we further partition the variance within treatments



$$F = \frac{\text{variance between treatments}}{\text{residual variance}}$$

ANOVA computations

Just like before we calculate

$$* \; SS_{total} = \sum x^2 - \frac{(\sum x)^2}{N}$$

$$* \; SS_{bet\ tmts} = N \sum_{j=1}^{4} \left(\overline{X}_{tmt_j} - \overline{X} \right)^2$$

We also need to compute:

$$* \; SS_{bet\ subj} = N \sum_{i=1}^{3} \left(\overline{X}_{sub_i} - \overline{X} \right)^2$$

Then SS_{resid} = whatever is left over

$$= SS_{total} - SS_{bet\ tmts} - SS_{bet\ subj}$$

Computations

$$SS_{\text{bet tmts}} = N \sum_{j=1}^{4} \left(\bar{X}_{\text{tmt}_j} - \bar{X}\right)^2$$

$$= 3\left[(3-4.5)^2 + (6-4.5)^2 + (5-4.5)^2 + (4-4.5)^2\right]$$

$$= 3(2.25 + 2.25 + 0.25 + 0.25)$$

$$= 15$$

$$SS_{\text{bet subj}} = N \sum_{i=1}^{3} \left(\bar{X}_{\text{sub}_i} - \bar{X}\right)^2$$

$$= 4\left[(7-4.5)^2 + (4.5-4.5)^2 + (2-4.5)^2\right]$$

$$= 4(6.25 + 0 + 6.25)$$

$$= 50$$

$$SS_{\text{total}} = \sum X^2 - \frac{\left(\sum X\right)^2}{N}$$

$$= 312 - \frac{54^2}{12} = 69$$

ANOVA table:

source	SS	df	MS	F
bet trmts	15	3	5	7.5
residual	4	6	0.67	
bet subj	50	2	25	
total	69	11		

Report: $F(3,6) = 7.50$, $p = 0.0187$,

$BF_{10} = 41.09$

What if we didn't account for subject variability?

source	SS	df	MS	F
bet tmts	15	3	5	0.74
residual	54	8	6.75	
total	69	11		

← $p = 0.557$

$BF_{01} = 9.56$

Moral: by accounting for subject variability, repeated measures ANOVA can better uncover latent treatment effects