Analysis of Variance

What terminology do I need to know to understand Anova?

How can Anova handle within and between subject designs?

A case study of Anova usage

Saul Greenberg

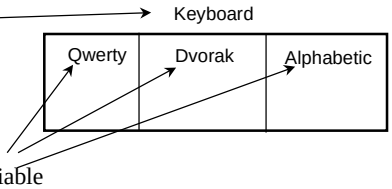
Analysis of Variance (Anova)

A Workhorse

- allows moderately complex experimental designs and statistics

Terminology

- Factor
  - independent variable
  - ie Keyboard, Toothpaste, Age
- Factor level
  - specific value of independent variable
  - ie Qwerty, Crest, 5-10 years old



Saul Greenberg

Anova terminology

- Between subjects (*aka* nested factors)
  - a subject is assigned to only one factor level of treatment
  - control is general population
  - problem: greater variability, requires more subjects

Keyboard		
Qwerty	Dvorak	Alphabetic
S1-20	*S21-40*	*S41-60*

- Within subjects (*aka* crossed factors)
  - subjects assigned to all factor levels of a treatment
  - subjects act as their own control
    requires fewer subjects
    less variability as subject measures are paired
  - problem:
    order effects (eg learning)
  - partially solved by counter-balanced ordering

Keyboard		
Qwerty	Dvorak	Alphabetic
S1-20	*S1-20*	*S1-20*

Saul Greenberg

Single Factor Analysis of Variance

Compare means between two or more factor levels within a single factor

example:

- dependent variable: mouse-typing speed
- independent variable (factor): keyboard
- between subject design

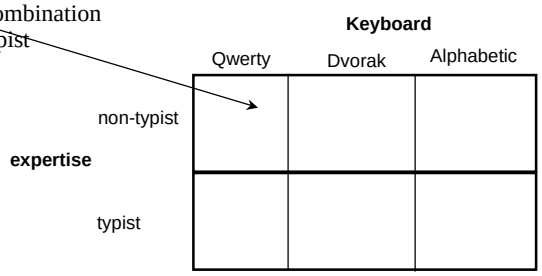
Qwerty	Alphabetic	Dvorak
S1: 25 secs S2: 29 … S20: 33	S21: 40 secs S22: 55 … S40: 33	S51: 17 secs S52: 45 … S60: 23

Saul Greenberg

Anova terminology

- Factorial design
  - cross combination of levels of one factor with levels of another
  - eg keyboard type (3) x expertise (2)

- Cell
  - unique treatment combination
  - eg qwerty x non-typist

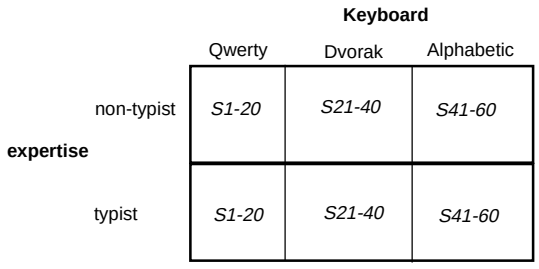


Saul Greenberg

Anova terminology

Mixed factor

- contains both between and within subject combinations



Saul Greenberg

Anova

Compares the relationships between many factors

Provides more informed results

- considers the interactions between factors
- eg
  - typists mouse-type faster on Qwerty, than on alphabetic and Dvorak
  - there is no difference in mouse-typing speeds for non-typists across all keyboards

	Qwerty	Alphabetic	Dvorak
non-typist	S1-S10	S11-S20	S21-S30
typist	S31-S40	S41-S50	S51-S60

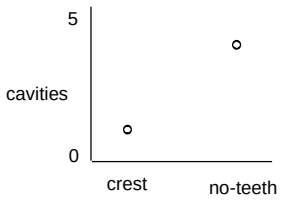
Saul Greenberg

Anova

In reality, we can rarely look at one variable at a time

Example:

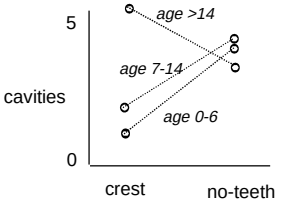
- t-test:
  subjects who use crest have fewer cavities



- anova: toothpaste x age

  Subjects who are 14 or less have fewer cavities with crest.

  Subjects who are older than 14 have fewer cavities with no-teeth.



Saul Greenberg

Anova case study

The situation

- text-based menu display for very large telephone directory
- names are presented as a range within a selectable menu item
- users navigate until unique names are reached

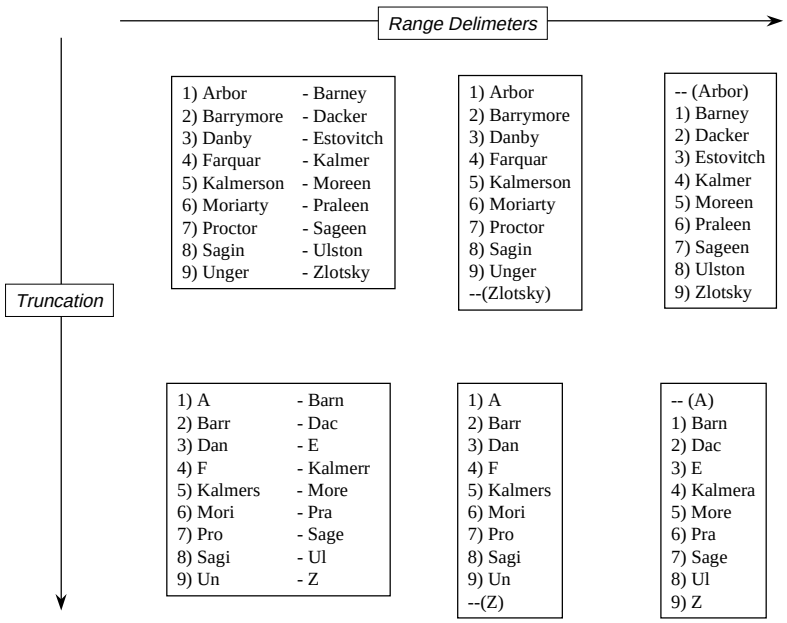
1) Arbor	- Kalmer	1) Arbor	- Farquar	...	1) Horace	- Horton
2) Kalmerson	- Ulston	2) Farston	- Hoover		2) Hoover, James	
3) Unger	- Zlotsky	3) Hover	- Kalmer		3) Howard, Rex	

- but several ways are possible to display these ranges

Question

- what display method is best?

Saul Greenberg



Saul Greenberg

Span

as one descends the menu hierarchy, name suffixes become similar

Wide Span	*Narrow Span*
1) Arbor 2) Barrymore 3) Danby 4) Farquar 5) Kalmerson 6) Moriarty 7) Proctor 8) Sagin 9) Unger --(Zlotsky)	1) Danby 2) Danton 3) Desiran 4) Desis 5) Dolton 6) Dormer 7) Eason 8) Erick 9) Fabian --(Farquar)

Saul Greenberg

Anova case study

Null hypothesis

- six menu display systems based on combinations of *truncation* and *delimiter* methods do not differ significantly from each other as measured by people's scanning speed and error rate
- *menu span* and *user experience* has no significant effect on these results

- 2 level (truncation) x
  2 level (menu span) x
  2 level (experience) x
  3 level (delimiter)

		Truncated		Not Truncated	
		narrow	wide	narrow	wide
Full	Novice	S1-8	S1-8	S1-8	S1-8
	Expert	S9-16	S9-16	S9-16	S9-16
Upper	Novice	S17-24	S17-24	S17-24	S17-24
	Expert	S25-32	S25-32	S25-32	S25-32
Lower	Novice	S33-40	S33-40	S33-40	S33-40
	Expert	S40-48	S40-48	S40-48	S40-48

Saul Greenberg

Statistical results

Scanning speed

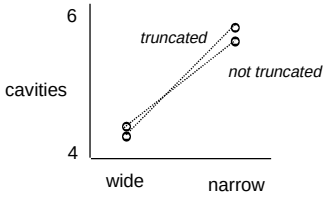
	F-ratio.	*p*
Range delimeter (R)	2.2*	<0.5
Truncation (T)	0.4	
Experience (E)	5.5*	<0.5
Menu Span (S)	216.0**	<0.01
RxT	0.0	
RxE	1.0	
RxS	3.0	
TxE	1.1	
TxS	14.8*	<0.5
ExS	1.0	
RxTxE	0.0	
RxTxS	1.0	
RxExS	1.7	
TxExS	0.3	
RxTxExS	0.5	

Saul Greenberg

Statistical results

Scanning speed:

- Truncation x Span



Main effects (means)

		Full	Lower	Upper
Full		----	1.15*	1.31*
Lower			----	0.16
Upper				----
Span:	Wide	4.35		
	Narrow	5.54		
Experience	Novice	5.44		
	Expert	4.36		

Results on Selection time

- Full range delimiters slowest
- Truncation has no effect on time
- Narrow span menus are slowest
- Novices are slower

Saul Greenberg

Statistical results

Error rate

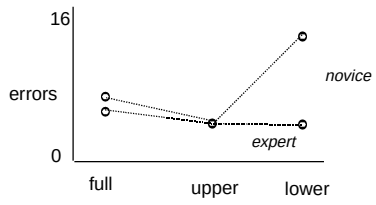
	F-ratio.	*p*
Range delimeter (R)	3.7*	<0.5
Truncation (T)	2.7	
Experience (E)	5.6*	<0.5
Menu Span (S)	77.9**	<0.01
RxT	1.1	
RxE	4.7*	<0.5
RxS	5.4*	<0.5
TxE	1.2	
TxS	1.5	
ExS	2.0	
RxTxE	0.5	
RxTxS	1.6	
RxExS	1.4	
TxExS	0.1	
RxTxExS	0.1	

Saul Greenberg

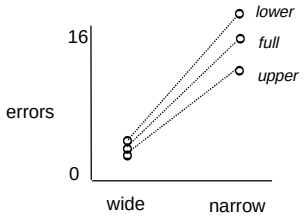
Statistical results

Error rates

- Range x Experience



Range x Span



Results on error rate

- lower range delimiters have more errors at narrow span
- truncation has no effect on errors
- novices have more errors at lower range delimiter

Saul Greenberg

Conclusions

- upper range delimiter is best
- truncation up to the implementers
- keep users from descending the menu hierarchy
- experience is critical in menu displays

Saul Greenberg

You now know

Anova terminology

- factors, levels, cells
- factorial design
  - between, within, mixed designs

Saul Greenberg