ANOVA approaches to Repeated Measures

- univariate repeated-measures ANOVA (chapter 2)
- repeated measures MANOVA (chapter 3)

Assumptions

- Interval measurement and normally distributed errors (homogeneous across groups) - transformation may help
- Group comparisons
  - estimation and comparison of group means
  - not informative about individual growth
- Fixed time points
  - time treated as classification variable
  - timepoints can be evenly or unevenly spaced

- Better suited for short-term or moderate change
  esp. true for trend or growth-curve analysis
- LS techniques adversely affected by outliers
- Missing data
  - ANOVA can handle some - unbalanced ANOVA
  - MANOVA can't handle any
- Variance-covariance structure for $\boldsymbol{y}_i$
  - compound symmetry for ANOVA
  - unstructured for MANOVA

Univariate Repeated Measures ANOVA

One-sample case: randomized block ANOVA

$$y_{ij} = \mu + \pi_i + \tau_j + e_{ij} \qquad (i = 1, \ldots, N\ ; \qquad j = 1, \ldots, n)$$

- μ = grand mean
- π_i = individual difference component for subject i (constant over time)
- τ_j = effect of time - same for all subjects
- e_{ij} = error for subject i and time j

Random components

- $\pi_i \sim N(0, \sigma^2_\pi)$ subjects random; between-subjects variance
- $e_{ij} \sim N(0, \sigma^2_e)$ within-subjects variance

Data layout

subject	timepoint 1	2	$\ldots$	n
1	y_{11}	y_{12}	$\cdots$	y_{1n}
2	y_{21}	y_{22}	$\cdots$	y_{2n}
.	.	.	$\cdots$	.
.	.	.	$\cdots$	.
N	y_{N1}	y_{N2}	$\cdots$	y_{Nn}

- one observation per cell
- similar to randomized block design with subject as block
- identical to paired-t test when $n = 2$

Model assumptions

$$\sum_{j=1}^{n} \tau_j = 0 \qquad (n-1\ df \ \text{ for Time})$$

$$E(y_{ij}) = \mu + \tau_j$$

$$V(y_{ij}) = V(\mu + \tau_j + \pi_i + e_{ij}) = \sigma_\pi^2 + \sigma_e^2$$

$$C(y_{ij}, y_{i'j}) = 0 \ \text{ for } \ i \neq i' \qquad \text{(different subjects)}$$

$$C(y_{ij}, y_{ij'}) = \sigma_\pi^2 \ \text{ for } \ j \neq j' \qquad \text{(same subject)}$$

Intraclass correlation

$$Corr(y_{ij}, y_{ij'}) = \frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_e^2}$$

- same correlation for all longitudinal pairs (average correlation of y from any two timepoints)
- ranges from 0 to 1, so more like a proportion of variance
- proportion of total variance that is attributable to subjects (that is not accounted for by Time)

Compound symmetry structure of variance-covariance matrix

$$\Sigma \boldsymbol{y}_i = \begin{bmatrix} \sigma_e^2 + \sigma_\pi^2 & \sigma_\pi^2 & \sigma_\pi^2 & \dots & \sigma_\pi^2 \\ \sigma_\pi^2 & \sigma_e^2 + \sigma_\pi^2 & \sigma_\pi^2 & \dots & \sigma_\pi^2 \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_\pi^2 & \dots & \sigma_\pi^2 & \sigma_e^2 + \sigma_\pi^2 & \sigma_\pi^2 \\ \sigma_\pi^2 & \sigma_\pi^2 & \dots & \sigma_\pi^2 & \sigma_e^2 + \sigma_\pi^2 \end{bmatrix}$$

- variance homogeneous across time: $\sigma_e^2 + \sigma_\pi^2$
- covariances homogenous across time: σ_π^2
  correlation $= \sigma_\pi^2 / (\sigma_e^2 + \sigma_\pi^2)$

Not too realistic

- variances often change over time: increase is common in clinical trials (subjects more similar at start of trial)
- covariances close in time usually greater than covariances distant in time

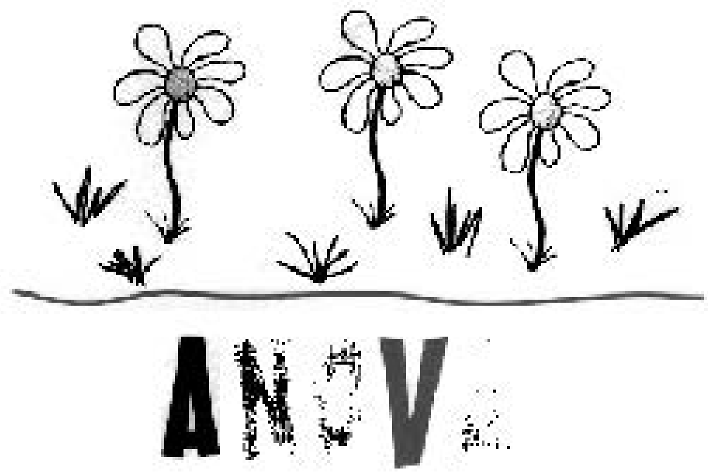
$\Rightarrow$ test of variance-covariance structure is necessary for the validity of significance tests (more on this later)

Chance (2005) 18:(3) 29-33

Should we quit using repeated measures analysis of variance?

Repeated Measures ANOVA, R.I.P.?

Charles E. McCulloch



ANOVA table: subjects random and time fixed (balanced)

Source	df	SS	MS	E(MS)
Subjects	$N-1$	$\mathrm{SS}_S = n\,\Sigma_{i=1}^{N}(\bar{y}_{i.} - \bar{y}_{..})^2$	$\frac{\mathrm{SS}_S}{N-1}$	$\sigma_e^2 + n\sigma_\pi^2$
Time	$n-1$	$\mathrm{SS}_T = N\,\Sigma_{j=1}^{n}(\bar{y}_{.j} - \bar{y}_{..})^2$	$\frac{\mathrm{SS}_T}{n-1}$	$\sigma_e^2 + N\,\Sigma(\tau_j - \tau_{.})^2$
Residual	$(N-1)$ $\times(n-1)$	$\mathrm{SS}_R = \Sigma_{i=1}^{N}\Sigma_{j=1}^{n}(y_{ij}$ $-\bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$	$\frac{\mathrm{SS}_R}{(N-1)(n-1)}$	σ_e^2
total	$Nn-1$	$\mathrm{SS}_y = \Sigma_{i=1}^{N}\Sigma_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2$		

- $\bar{y}_{..}$ = grand mean (averaged over time and subjects)
- $\bar{y}_{i.}$ = subject mean ($i = 1, \ldots, N$)
- $\bar{y}_{.j}$ = timepoint mean ($j = 1, \ldots, n$)

$$H_S: \qquad \sigma_\pi^2 = 0 \qquad\qquad F_S = \frac{\mathrm{MS}_S}{\mathrm{MS}_R} \overset{H_S}{\sim} F_{N-1,(N-1)(n-1)}$$

$$H_T: \tau_1 = \tau_2 = ... = \tau_n = 0 \qquad\qquad F_T = \frac{\mathrm{MS}_T}{\mathrm{MS}_R} \overset{H_T}{\sim} F_{n-1,(N-1)(n-1)}$$

note: $\mathrm{MS}_R =$ MSE

Intraclass correlation - relative magnitude of σ_π^2

- usually assume $\sigma_\pi^2 > 0$
- want to quantify degree of subject effect

$$r = \frac{\hat{\sigma}_\pi^2}{\hat{\sigma}_\pi^2 + \hat{\sigma}_e^2}$$

- $\hat{\sigma}_\pi^2 \; = \; (\ \mathrm{MS}_S - \ \mathrm{MS}_R) \ / \ n$
  - when $\mathrm{MS}_S \leq \mathrm{MS}_R$, then $\hat{\sigma}_\pi^2 = 0$
- $\hat{\sigma}_e^2 \; = \; \mathrm{MS}_R$

- proportion of (unexplained) variation due to subjects
- as subjects' data are highly correlated, r gets large
- can vary depending on model covariates

Contrasts for time differences

- $H_T : \tau_1 = \tau_2 = ... = \tau_n = 0$ is a very global test
- contrasts address more specific time-related comparisions
- $H_{j'} : Ł_{j'} = \Sigma_{j=1}^{n} c_{j'j}\, \mu_{.j} = 0$

Define the (estimated) contrast of the timepoint means as:

$$L_{j'} \;=\; \sum_{j=1}^{n} c_{j'j}\, \bar{y}_{.j} \qquad j' = 1, \ldots, n-1 \qquad \sum_{j=1}^{n} c_{j'j} = 0$$

e.g., suppose that there are two timepoints, then

$$H_1 : \mu_1 = \mu_2 \iff H_1 : \mu_1 - \mu_2 = 0 \iff H_1 : (1)\mu_1 + (-1)\mu_2 = 0$$

here, $c_{11} = 1$ and $c_{12} = -1$ and H_1 is tested by the contrast:

$$L_1 = (1)\bar{y}_1 + (-1)\bar{y}_2$$

suppose that there are three timepoints and interest is in:

$$H_1 : \mu_1 = \mu_2 \iff H_1 : (1)\mu_1 + (-1)\mu_2 + (0)\mu_3 = 0$$

$$H_2 : \mu_1 = \mu_3 \iff H_2 : (1)\mu_1 + (0)\mu_2 + (-1)\mu_3 = 0$$

these would be tested by the contrasts:

$$L_1 = (1)\bar{y}_1 + (-1)\bar{y}_2 + (0)\bar{y}_3$$

$$L_2 = (1)\bar{y}_1 + (0)\bar{y}_2 + (-1)\bar{y}_3$$

Inference for contrasts $H_{j'} : \text{Ł}_{j'} = \Sigma_{j=1}^{n} c_{j'j}\,\mu_{.j} = 0$

$$\mathrm{SS}_{j'} \; = \; \mathrm{MS}_{j'} \; = \; N L_{j'}^2 \,/\, \sum_{j=1}^{n} c_{j'j}^2$$

$$F_{j'} = \frac{\mathrm{MS}_{j'}}{\mathrm{MS}_R} \overset{H_{j'}}{\sim} F_{1,(N-1)(n-1)}$$

$$t_{j'} = \frac{L_{j'}}{\sqrt{\mathrm{MS}_R \left[\Sigma_{j=1}^{n} \frac{c_{j'j}^2}{N} \right]}} \overset{H_{j'}}{\sim} t_{(N-1)(n-1)}$$

If the set of $n-1$ contrasts are orthogonal, then

$$\mathrm{SS}_T = \sum_{j'=1}^{n-1} \mathrm{SS}_{j'}$$

$\Rightarrow$ independent partitioning of the variation due to time

Trend Analysis - orthogonal polynomials

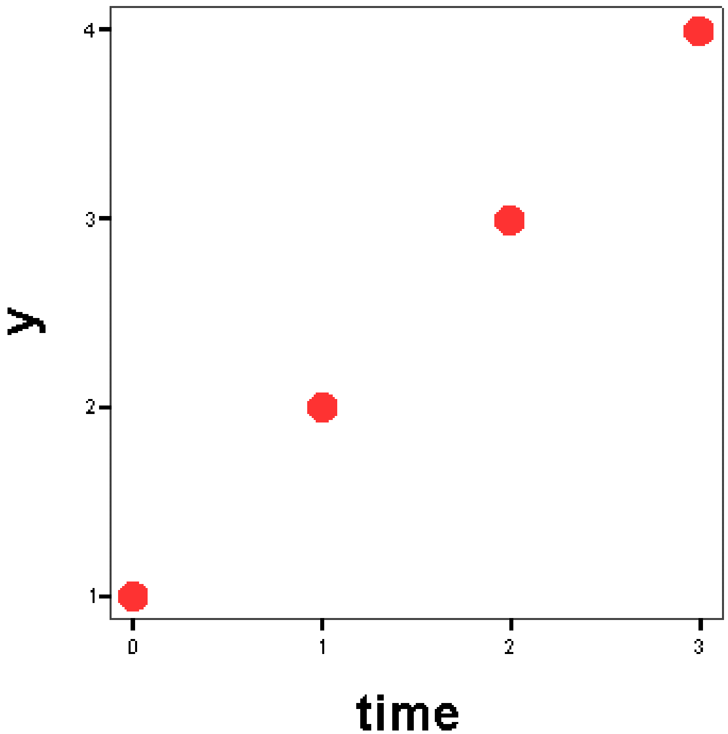
- characterize $n-1$ time effects as $n-1$ orthogonal polynomials
- tables in Bock (1975), Draper & Smith (1981), Fleiss (1986)

For example, the $n-1 \times n$ contrast matrix for $n=4$ is

$$
\boldsymbol{C} = \begin{bmatrix} -3 & -1 & 1 & 3 & \vdots & \div\sqrt{20} \\ 1 & -1 & -1 & 1 & \vdots & \div\sqrt{4} \\ -1 & 3 & -3 & 1 & \vdots & \div\sqrt{20} \end{bmatrix} \begin{array}{l} \text{linear} \\ \text{quad} \\ \text{cubic} \end{array}
$$

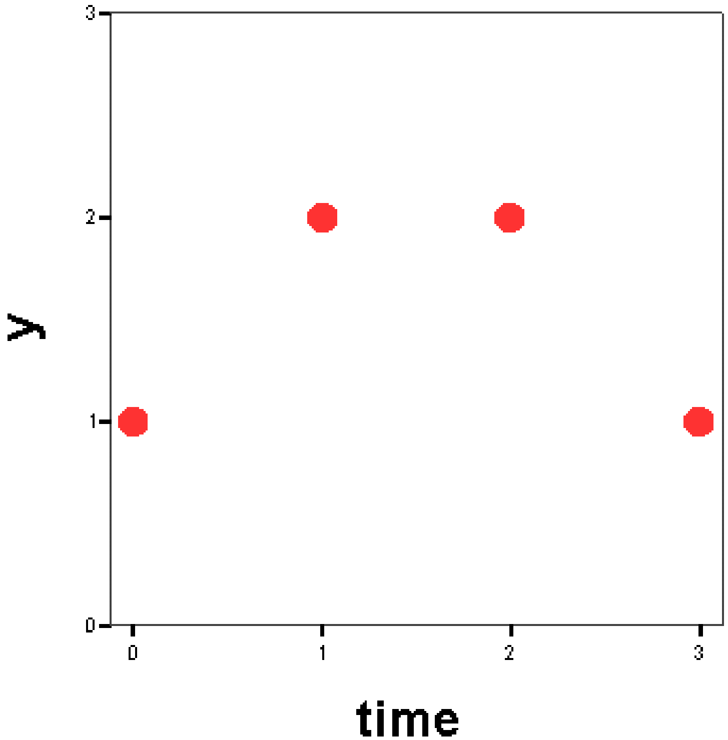
- orthogonal (and orthonormal with $\div$)
- useful for determining "degree" of change across time
- generalized for unequally spaced timepoints
- can specify fewer than $n-1$ contrasts

Purely linear change



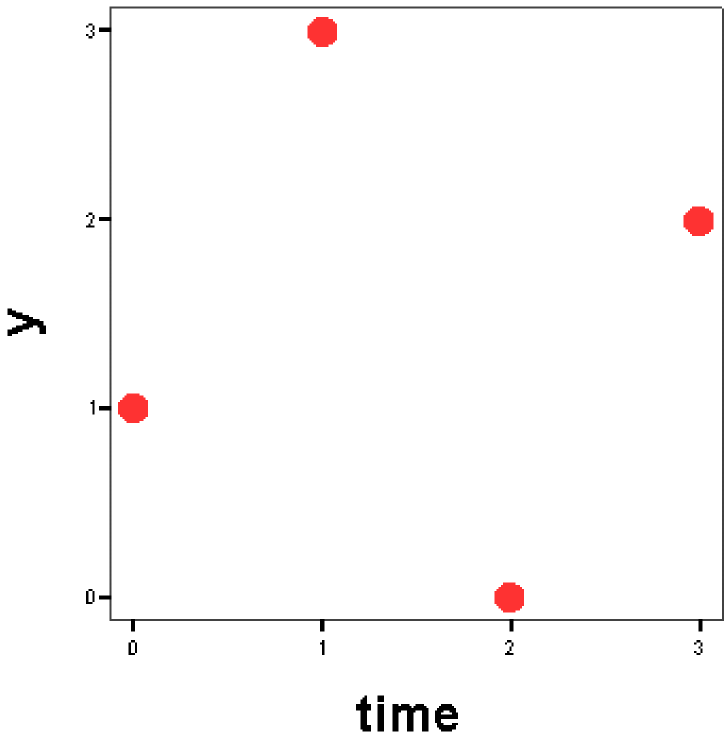
$$
\begin{aligned}
L_1 &= (-3)1 + (-1)2 + (1)3 + (3)4 = 10/\sqrt{20} = 2.24 \\
L_2 &= (1)1 + (-1)2 + (-1)3 + (1)4 = 0 \\
L_3 &= (-1)1 + (3)2 + (-3)3 + (1)4 = 0
\end{aligned}
$$

Purely quadratic change

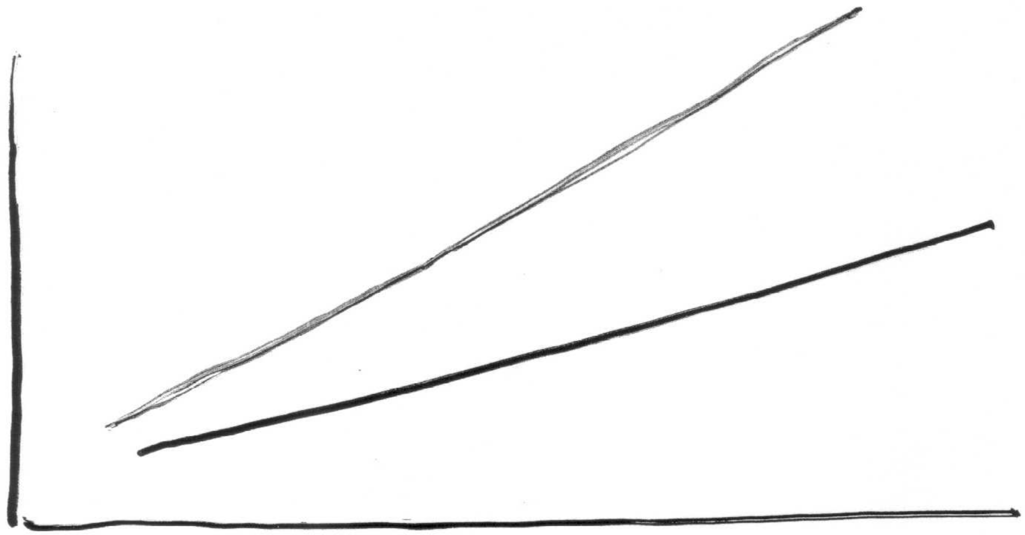


$$
\begin{aligned}
L_1 &= (-3)1 + (-1)2 + (1)2 + (3)1 = 0 \\
L_2 &= (1)1 + (-1)2 + (-1)2 + (1)1 = -2/\sqrt{4} = -1 \\
L_3 &= (-1)1 + (3)2 + (-3)2 + (1)1 = 0
\end{aligned}
$$

Purely cubic change

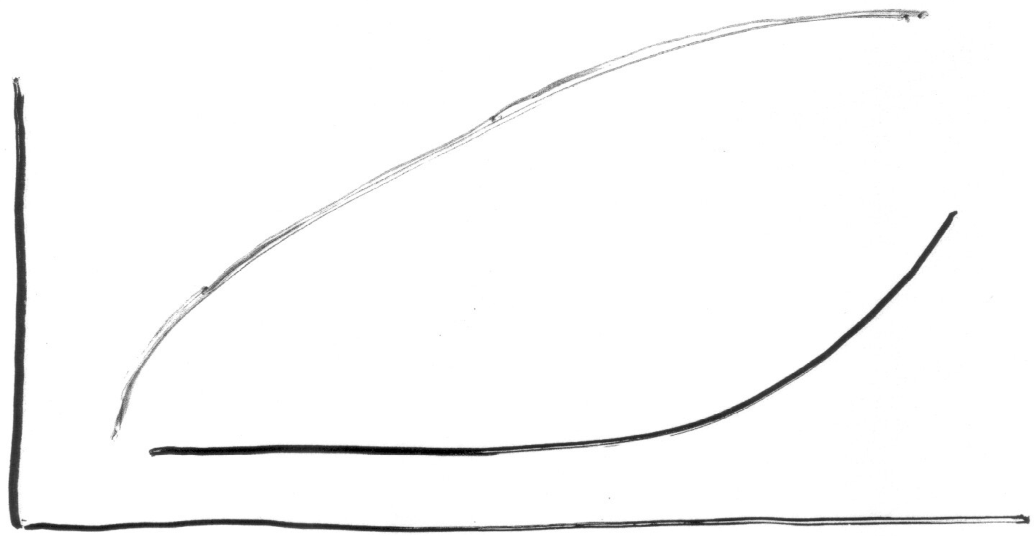


$$
\begin{aligned}
L_1 &= (-3)1 + (-1)3 + (1)0 + (3)2 = 0 \\
L_2 &= (1)1 + (-1)3 + (-1)0 + (1)2 = 0 \\
L_3 &= (-1)1 + (3)3 + (-3)0 + (1)2 = 10/\sqrt{20} = 2.24
\end{aligned}
$$



linear trends across time

cubic
trends



quadratic
trends

Example 1: IML code for generating orthogonal polynomial contrast matrix (SAS code and output)

BB: Example 1: ANOVA-Ex1-IML-orthpoly.txt

```
TITLE 'Generating Orthogonal Polynomial Matrix';

PROC IML;

 xraw = { 1 1 1  1   ,
          0 1 2  3   ,
          0 1 4  9   ,
          0 1 8 27 } ;

xorth =T(INV(ROOT(xraw*T(xraw))))*xraw;

PRINT 'Matrix of Time Polynomials', xraw
[FORMAT=10.4];

PRINT 'Orthonormalized Matrix of Time
Polynomials', xorth [FORMAT=10.4];
```

Generating Orthogonal Polynomial Matrix

Matrix of Time Polynomials

XRAW

1.0000	1.0000	1.0000	1.0000
0.0000	1.0000	2.0000	3.0000
0.0000	1.0000	4.0000	9.0000
0.0000	1.0000	8.0000	27.0000

Orthonormalized Matrix of Time Polynomials

XORTH

0.5000	0.5000	0.5000	0.5000
-0.6708	-0.2236	0.2236	0.6708
0.5000	-0.5000	-0.5000	0.5000
-0.2236	0.6708	-0.6708	0.2236

Change relative to 1st timepoint (*e.g.*, baseline)

$$\boldsymbol{C} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$$

for T2-T1, T3-T1, T4-T1, respectively

- not orthogonal
- sometimes called simple contrasts
- other cell can be reference cell (*e.g.*, last time)

Consectutive time comparisons (profile contrasts)

$$\boldsymbol{C} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

for T2-T1, T3-T2, T4-T3, respectively

- not orthogonal
- useful for identifying when change begins (and ends)

Contrasting timepoint to mean of subsequent timepoints (Helmert contrasts)

$$\boldsymbol{C} = \begin{bmatrix} 1 & -1/3 & -1/3 & -1/3 \\ 0 & 1 & -1/2 & -1/2 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

for T1 vs avg of T2, T3, T4; T2 vs avg of T3, T4; T3 vs T4

- orthogonal
- useful for "ordered" tests

Contrasting each timepoint to mean of others

(Deviational contrasts)

$$\boldsymbol{C} = \begin{bmatrix} 1 & -1/3 & -1/3 & -1/3 \\ -1/3 & 1 & -1/3 & -1/3 \\ -1/3 & -1/3 & 1 & -1/3 \end{bmatrix}$$

for T1 vs avg of T2, T3, T4; T2 vs avg of T1, T3, T4; T3 vs avg of T1, T2, T4

- not orthogonal
- useful for "vague knowledge" situation
- omitted timepoint can be other than last

Testing Time Differences

multiple comparisons issue, though often $n-1$ contrasts can be specified a-priori

- Bonferroni corrected α level
  - $\alpha^* = \alpha/(n-1)$
- Fisher protected test logic
  - use α, but only apply when $H_T : \tau_1 = \tau_2 = ... = \tau_n = 0$ is rejected
- For orthogonal polynomials
  - start with highest-order polynomial and eliminate (simplify) as polynomials are non-significant

Hummel & Sligo (1971) support Fisher protected test if n is not too large

Also, from Rosner (1995) *Fundamentals of Biostatistics*, p. 319:

> If a few linear contrasts, which have been specified in advance, are to be tested, then it may not be necessary to use a multiple-comparisons procedure, since if such procedures are used, there will be less power to detect differences for linear contrasts whose means are truly different from zero. Conversely, if many contrasts are to be tested, which have not been specified before looking at the data, then multiple-comparisons procedures may be useful in protecting against declaring too many significant differences.

Example 2: Analysis of vocabulary data from Bock (1975) using univariate repeated measures ANOVA

(SAS code and output)

BB: Example 2:
Plotting Data: ANOVA-Ex2a-bockvoc_ODS_plots.sas.txt
Analysis: ANOVA-Ex2b-bockv-analysis.txt

Multi-sample repeated measures ANOVA

- "split-plots" ANOVA model; Potthoff and Roy example

		timepoint			
group	subject	1	2	$\ldots$	n
1	1	y_{111}	y_{112}	$\ldots$	y_{11n}
1	2	y_{121}	y_{122}	$\ldots$	y_{12n}
1	.	.	.	$\ldots$	.
1	.	.	.	$\ldots$	.
1	N_1	y_{1N_11}	y_{1N_12}	$\ldots$	y_{1N_1n}
.	.	.	.	$\ldots$	.
.	.	.	.	$\ldots$	.
.	.	.	.	$\ldots$	.
.	.	.	.	$\ldots$	.
s	1	y_{s11}	y_{s12}	$\ldots$	y_{s1n}
s	2	y_{s21}	y_{s22}	$\ldots$	y_{s2n}
s	.	.	.	$\ldots$	.
s	.	.	.	$\ldots$	.
s	N_s	y_{sN_s1}	y_{sN_s2}	$\ldots$	y_{sN_sn}

$h = 1, \ldots, s$ groups; $i = 1, \ldots, N_h$ subjects in group h (with $N = \Sigma_{h=1}^{s} N_h$);
$j = 1, \ldots, n$ timepoints

Model

$$y_{hij} = \mu + \gamma_h + \tau_j + (\gamma\tau)_{hj} + \pi_{i(h)} + e_{hij}$$

- μ = grand mean
- γ_h = effect of group h $(\Sigma_h \gamma_h = 0)$
- τ_j = effect of time j $(\Sigma_j \tau_j = 0)$
- $(\gamma\tau)_{hj}$ = interaction of time j by grp h $[\Sigma_h \Sigma_j (\gamma\tau)_{hj} = 0]$
- $\pi_{i(h)}$ = individual difference component for sub i in grp h
- e_{hij} = error for subject i in group h at time j

- $\pi_{i(h)} \sim N(0, \sigma_\pi^2)$
- $e_{hij} \sim N(0, \sigma_e^2)$

$\Rightarrow$ compound symmetry structure for $V(\boldsymbol{y}_i)$

ANOVA table: subjects random and group & time fixed (balanced in terms of n, not in terms of N_h)

Source	df	SS	MS	E(MS)
Group	$s-1$	$\mathrm{SS}_G = n\,\Sigma_{h=1}^{s}\,N_h\,(\bar{y}_{h..}-\bar{y}_{...})^2$	$\frac{\mathrm{SS}_G}{s-1}$	$\sigma_e^2 + n\sigma_\pi^2 + D_G$
Time	$n-1$	$\mathrm{SS}_T = N\,\Sigma_{j=1}^{n}(\bar{y}_{..j}-\bar{y}_{...})^2$	$\frac{\mathrm{SS}_T}{n-1}$	$\sigma_e^2 + D_T$
Group $\times$ Time	$(s-1)$ $\times(n-1)$	$\mathrm{SS}_{GT} = \Sigma_{h=1}^{s}\,\Sigma_{j=1}^{n}\,N_h\,(\bar{y}_{h.j} -\bar{y}_{h..}-\bar{y}_{..j}+\bar{y}_{...})^2$	$\frac{\mathrm{SS}_{GT}}{(s-1)(n-1)}$	$\sigma_e^2 + D_{GT}$
Subjects in Grps	$N-s$	$\mathrm{SS}_{S(G)} = n\,\Sigma_{h=1}^{s}\,\Sigma_{i=1}^{N_h}(\bar{y}_{hi.}-\bar{y}_{h..})^2$	$\frac{\mathrm{SS}_{S(G)}}{N-s}$	$\sigma_e^2 + n\sigma_\pi^2$
Residual	$(N-s)$ $\times(n-1)$	$\mathrm{SS}_R = \Sigma_{h=1}^{s}\,\Sigma_{i=1}^{N_h}\,\Sigma_{j=1}^{n}\,(y_{hij} -\bar{y}_{h.j}-\bar{y}_{hi.}+\bar{y}_{h..})^2$	$\frac{\mathrm{SS}_R}{(N-s)(n-1)}$	σ_e^2
total	$Nn-1$	$\mathrm{SS}_y = \Sigma_{h=1}^{s}\,\Sigma_{i=1}^{N_h}\,\Sigma_{j=1}^{n}(y_{hij}-\bar{y}_{...})^2$		

- $h = 1, \ldots, s$ groups
- $i = 1, \ldots, N_h$ subjects in group h $\quad$ (with $N = \Sigma_{h=1}^{s} N_h$)
- $j = 1, \ldots, n$ timepoints

$D_G, D_T,$ and D_{GT} represent differences among groups, timepoints, and group $\times$ time interaction, respectively

Testing for Group by Time interaction

$$H_{GT} : D_{GT} = 0 \qquad F_{GT} = \frac{\mathrm{MS}_{GT}}{\mathrm{MS}_R} \overset{H_{GT}}{\sim} F_{(s-1)(n-1),(N-s)(n-1)}$$

Usually, test of primary interest

If rejected,

- group differences are not the same across time
- group curves across time are not parallel
- group and time effects are confounded with the interaction & can't be separately tested (or estimated)

If $H_{GT} : D_{GT} = 0$ accepted,

$$H_T : \tau_1 = \tau_2 = \ldots = \tau_n = 0 \qquad F_T = \frac{\mathrm{MS}_T}{\mathrm{MS}_R} \overset{H_T}{\sim} F_{n-1,(N-s)(n-1)}$$

$$H_G : \gamma_1 = \gamma_2 = \ldots = \gamma_s = 0 \qquad F_G = \frac{\mathrm{MS}_G}{\mathrm{MS}_{S(G)}} \overset{H_G}{\sim} F_{s-1,N-s}$$

H_T and H_G are separately and independently testable

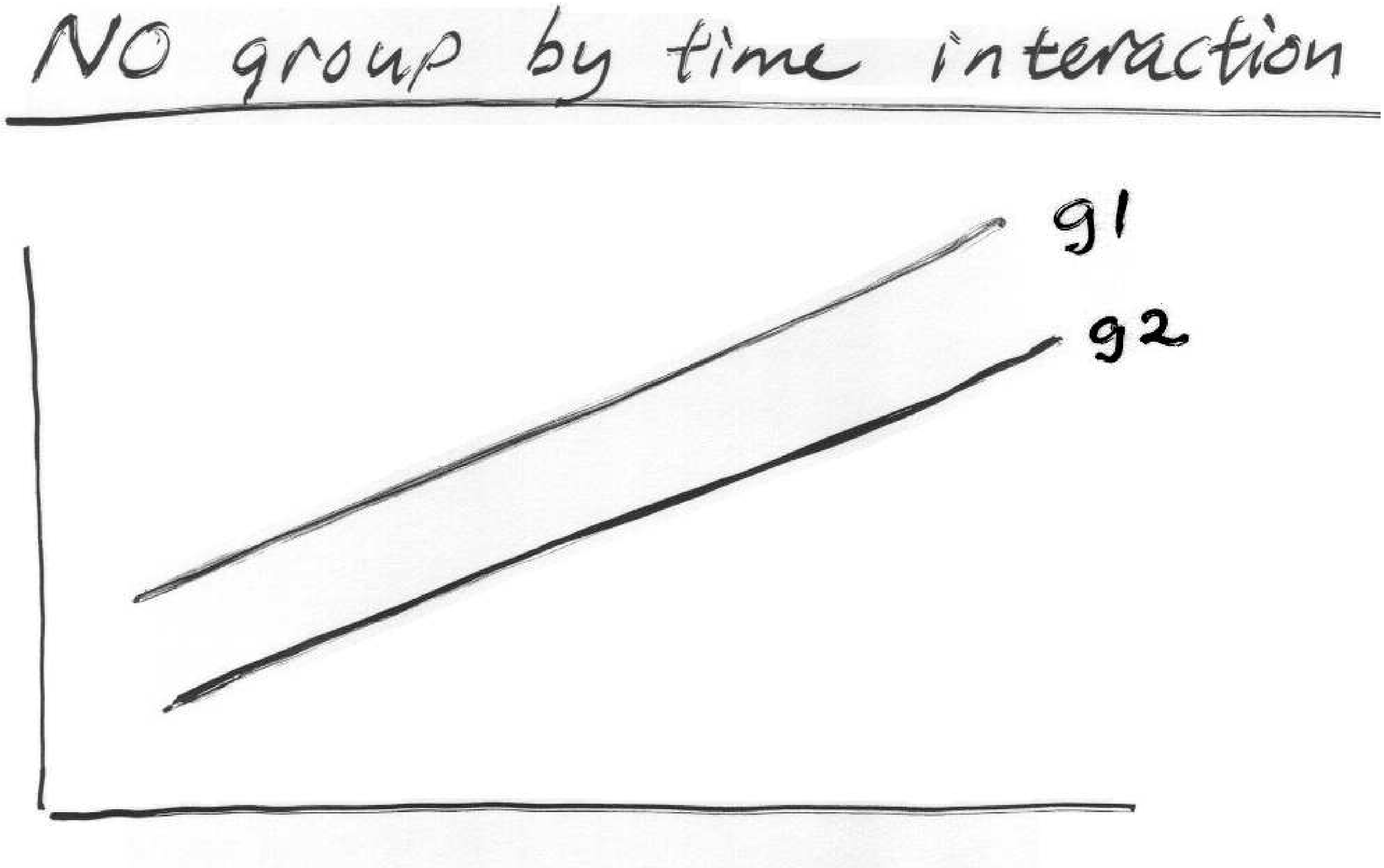


Figure 1: Caption for nogt1

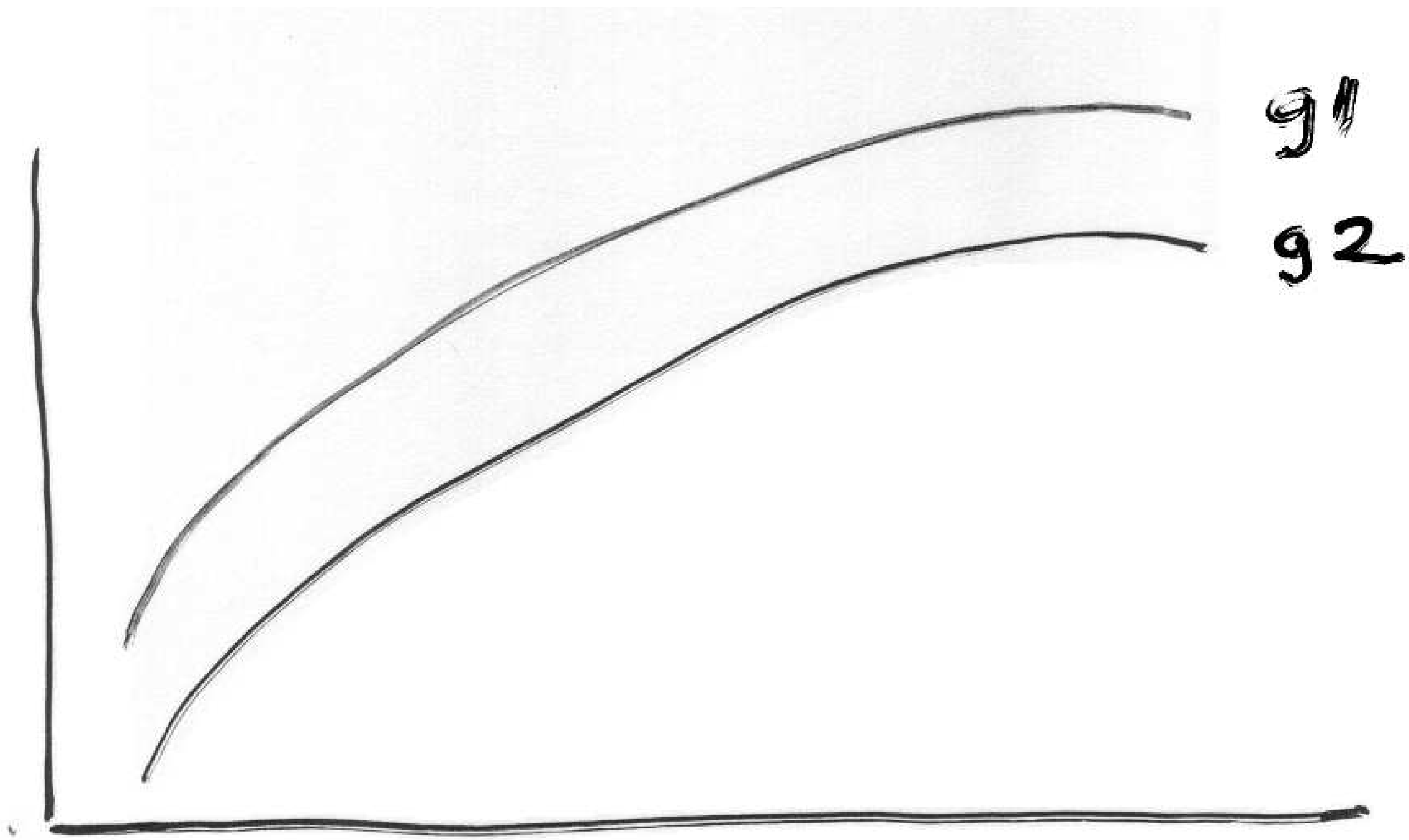


Figure 2: Caption for nogt2

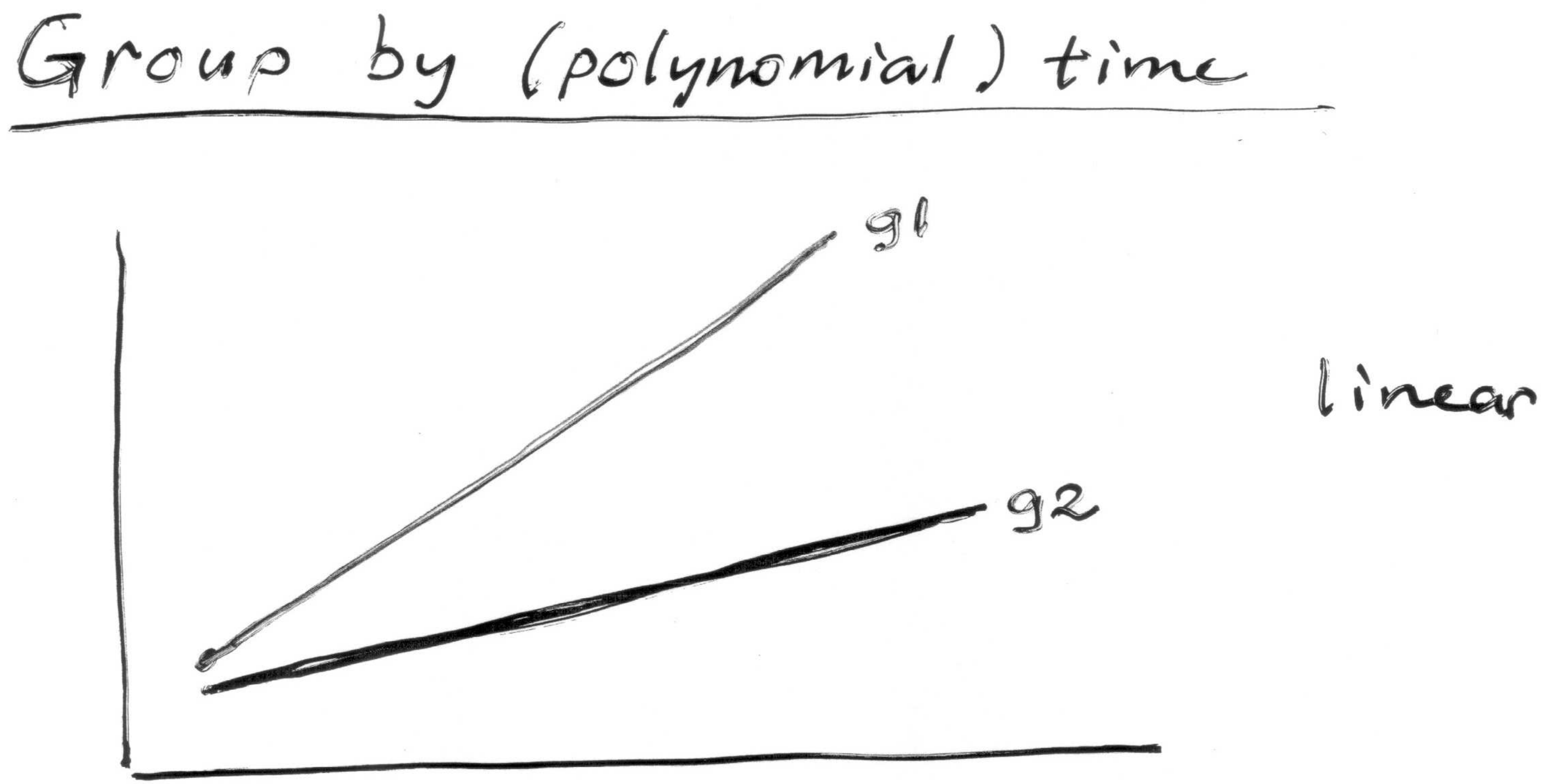


Figure 3: Caption for gt1

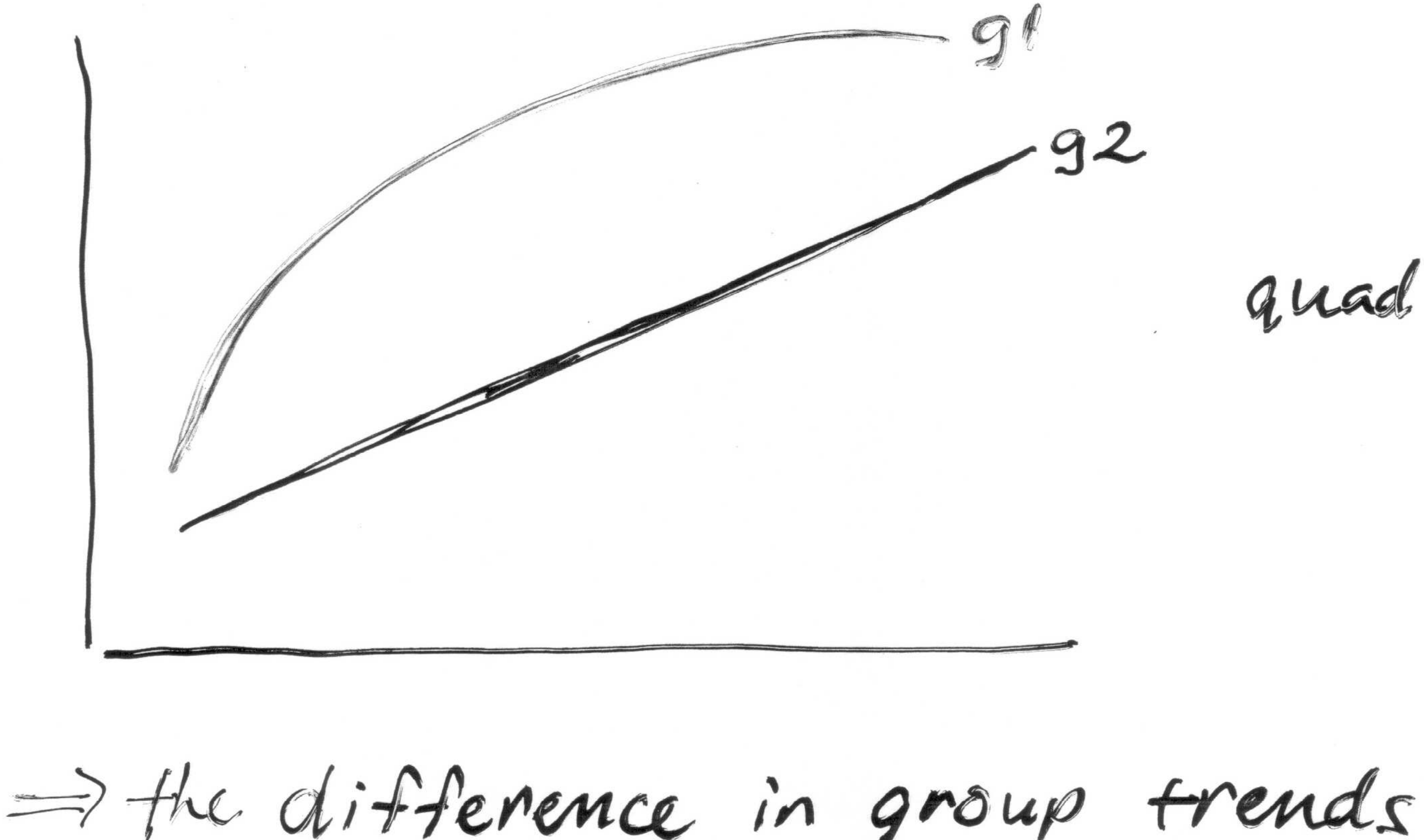


Figure 4: Caption for gt2

Testing for Subject effect

$$H_{S(G)} : \sigma_\pi^2 = 0 \qquad F_{S(G)} = \frac{\mathrm{MS}_{S(G)}}{\mathrm{MS}_R} \overset{H_{S(G)}}{\sim} F_{N-s,(N-s)(n-1)}$$

- usually assume $\sigma_\pi^2 > 0$
- estimate intraclass correlation
  - $r = \hat{\sigma}_\pi^2 / (\hat{\sigma}_\pi^2 + \hat{\sigma}_e^2)$
  - proportion of (unexplained) variation due to subjects

Contrasts for time effects

Again, it's advantageous to characterize time and group by time effects using time-related contrasts

- *e.g.*, orthogonal polynomials
- Fisher protected tests or Bonferroni correction for $(n-1)(s-1)$ group by time contrasts, or $n-1$ time contrasts

Orthogonal Polynomial Partition of SS

$$\boldsymbol{C}_4 = \begin{array}{c} \left[\begin{array}{rrrrcr} -3 & -1 & 1 & 3 & \vdots & \div\sqrt{20} \\ 1 & -1 & -1 & 1 & \vdots & \div\sqrt{4} \\ -1 & 3 & -3 & 1 & \vdots & \div\sqrt{20} \end{array}\right] \end{array} \begin{array}{c} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \boldsymbol{c}_3 \end{array}$$

Time	df	SS
linear	1	$\mathrm{SS}_{T_1} = N\,\boldsymbol{c}_1\,\bar{\boldsymbol{y}}_{..j}\,\bar{\boldsymbol{y}}'_{..j}\,\boldsymbol{c}'_1$
quadratic	1	$\mathrm{SS}_{T_2} = N\,\boldsymbol{c}_2\,\bar{\boldsymbol{y}}_{..j}\,\bar{\boldsymbol{y}}'_{..j}\,\boldsymbol{c}'_2$
.	.	.
$(n-1)$th	1	$\mathrm{SS}_{T_{n-1}} = N\,\boldsymbol{c}_{n-1}\,\bar{\boldsymbol{y}}_{..j}\,\bar{\boldsymbol{y}}'_{..j}\,\boldsymbol{c}'_{n-1}$
	$n-1$	SS_T

- $\bar{\boldsymbol{y}}_{..j} = n \times 1$ vector of timepoint means
- $\boldsymbol{c}_j = 1 \times n$ vector of contrasts of order j

$$
\begin{aligned}
F_{T_{n-1}} &= \ \mathrm{SS}_{T_{n-1}} / \ \mathrm{MS}_R \\
&\ \cdot \ \cdot \ \cdot \\
F_{T_1} &= \ \mathrm{SS}_{T_1} / \ \mathrm{MS}_R
\end{aligned}
$$

Determine polynomial of least degree working backwards

G × T	df	SS
linear	$s-1$	$\mathrm{SS}_{GT_1} = \Sigma_h \, N_h \, \boldsymbol{c}_1 \, \bar{\boldsymbol{y}}_{h.j} \, \bar{\boldsymbol{y}}'_{h.j} \, \boldsymbol{c}'_1 - \mathrm{SS}_{T_1}$
quadratic	$s-1$	$\mathrm{SS}_{GT_2} = \Sigma_h \, N_h \, \boldsymbol{c}_2 \, \bar{\boldsymbol{y}}_{h.j} \, \bar{\boldsymbol{y}}'_{h.j} \, \boldsymbol{c}'_2 - \mathrm{SS}_{T_2}$
.	.	.
$(n-1)$th	$s-1$	$\mathrm{SS}_{GT_{n-1}} = \Sigma_h \, N_h \, \boldsymbol{c}_{n-1} \, \bar{\boldsymbol{y}}_{h.j} \, \bar{\boldsymbol{y}}'_{h.j} \, \boldsymbol{c}'_{n-1} - \mathrm{SS}_{T_{n-1}}$
	$(s-1) \times (n-1)$	SS_{GT}

- $\bar{\boldsymbol{y}}_{h.j} = n \times 1$ vector of timepoint means for group h
- $\boldsymbol{c}_j = 1 \times n$ vector of contrasts of order j

$$
\begin{aligned}
F_{GT_{n-1}} &= \frac{\mathrm{SS}_{GT_{n-1}} \,/\, (s-1)}{\mathrm{MS}_R} \\
&\cdot\ \cdot\ \cdot \\
&\cdot\ \cdot\ \cdot \\
F_{GT_1} &= \frac{\mathrm{SS}_{GT_1} \,/\, (s-1)}{\mathrm{MS}_R}
\end{aligned}
$$

Determine polynomial of least degree working backwards

Example 3:
Analysis of Pothoff & Roy data using univariate repeated measures ANOVA model with two groups (boys and girls).

Plot of the means across time for the two groups:
BB: Example 3: ANOVA-Ex3a-pothoff_ODS_plots.sas.txt

Interaction Model: ANOVA-Ex3b-pothoff-interaction.txt

Example 4:
Analysis of Pothoff & Roy data using univariate repeated measures ANOVA.

This analysis is for a main effects model - no group by time interaction.

SAS code and output:
BB: Example 4: ANOVA-Ex4-pothoff-no-interaction.txt

SAS code for univariate repeated measures ANOVA with more than 2 groups

suppose 3 groups (control, tx1, tx2) and interest in comparisons to control

	control	tx1	tx2
contrast 1	1	-1	0
contrast 2	1	0	-1

suppose 4 timepoints and interest in trend analysis

	time1	time2	time3	time4
linear	-3	-1	1	3
quad	1	-1	-1	1
cubic	-1	3	-3	1

```
PROC GLM;
  CLASS group time subject;
  MODEL y = group subject(group) time group*time;
  TEST H=group E=subject(group);
  CONTRAST 'tx1' group 1 -1 0;
  CONTRAST 'tx2' group 1 0 -1;
  CONTRAST 'linear' time -3 -1 1 3;
  CONTRAST 'quad' time 1 -1 -1 1;
  CONTRAST 'cubic' time -1 3 -3 1;
  CONTRAST 'tx1*lin' group*time
            -3 -1 1 3   3 1 -1 -3   0 0 0 0;
  CONTRAST 'tx2*lin' group*time
            -3 -1 1 3   0 0 0 0   3 1 -1 -3;
  .........
  CONTRAST 'tx2*cub' group*time
            -1 3 -3 1   0 0 0 0   1 -3 3 -1;
```

Compound Symmetry and Sphericity

For both univariate RB and SP ANOVA models

$$V(\boldsymbol{y}_i) \;=\; \sigma_\pi^2 \mathbf{1}_n \mathbf{1}_n' \;+\; \sigma_e^2 \boldsymbol{I}_n$$

Compound symmetry (CS) structure

$$V(y_{ij}) \;=\; \sigma_\pi^2 + \sigma_e^2 \qquad\qquad \forall\; j$$

$$C(y_{ij}, y_{ij'}) \;=\; \sigma_\pi^2 \qquad\qquad\qquad \forall\; j \text{ and } j' \;\; (j \neq j')$$

- $\rho(y_{ij}, y_{ij'}) \;=\; \sigma_\pi^2 \,/\, (\sigma_\pi^2 + \sigma_e^2) \;=\; \text{ICC}$
- Highly restrictive assumption and often unrealistic (especially as n gets large)

- CS is special case of more general situation, sphericity, under which F-tests for time-related terms from ANOVA models are valid
  - if sphericity holds, then F-tests are valid
  - if sphericity doesn't hold, then F-tests are generally too liberal

Sphericity - sometimes called circularity

- Can be expressed in different ways
- *e.g.*, variances of all pairwise differences between variables are equal

$$\begin{aligned} V(y_{ij} - y_{ij'}) &= V(y_{ij}) + V(y_{ij'}) - 2C(y_{ij}, y_{ij'}) \\ &= \text{constant } \forall\, j \text{ and } j' \end{aligned}$$

- CS is a special case of sphericity

$$\begin{aligned} V(y_{ij}) &= \sigma_\pi^2 + \sigma_e^2 && \text{constant} \\ C(y_{ij}, y_{ij'}) &= \sigma_\pi^2 && \text{constant} \end{aligned}$$

More generally, Crowder and Hand (1990, page 50) note:

> MS ratios derived by the univariate approach follow exact F-distributions if and only if the covariance matrix of the orthonormal contrasts has equal variances and zero covariances

$$\underset{(n-1)\times n}{\boldsymbol{C}} \quad \underset{n\times n}{V(\boldsymbol{y}_i)} \quad \underset{n\times(n-1)}{\boldsymbol{C}'} \quad = \quad \text{constant} \quad \underset{(n-1)\times(n-1)}{\boldsymbol{I}_{n-1}}$$

- where $\boldsymbol{C}$ is matrix of orthonormal polynomials
- test in SAS using multivariate data setup

Running univariate repeated-measures ANOVA via multivariate setup

- univariate setup:

group	subject	time	y
1	1	1	y_{111}
1	1	2	y_{112}
1	1	3	y_{113}

- multivariate setup:

group	subject	y1	y2	y3
1	1	y_{111}	y_{112}	y_{113}

```
DATA unidat;
INPUT subject time y;
DATALINES;
1 1 1
1 2 2
1 3 3
2 1 4
2 2 5
2 3 6
;
DATA multdat;
ARRAY yv(3) y1 y2 y3;
DO time = 1 TO 3;
  SET unidat;
  yv(time) = y;
END;
DROP y subject time;
RUN;
```

SAS code for PROC GLM

```
proc glm;
  class group;
  model y1 y2 y3 = group / nouni;
  repeated time polynomial / summary printm
                             printe;
```

- `nouni` prevents separate tests for y1, y2, y3
- `summary` provides univariate repeated-measures results
- `printm` prints contrast matrix for $\boldsymbol{y}_i$
- `printe` gives sphericity information

Running univariate repeated-measures ANOVA via multivariate setup - notes

- Subjects with missing data on any timepoint are deleted from analysis (unlike in univariate setup)
- For the time-related contrasts, though SS and MS are the same, F-tests are not
  - denominator = MSE in univariate setup
  - denominator obtained from SSCP matrix in multivariate setup (more on this later)
- Only get sphericity test via multivariate setup
- Denominator $df >$ for univariate repeated-measures tests
  - generally greater power for univariate tests
- Sphericity is for pooled subjects-within-groups variance-covariance matrix (error var-covar)

- Get different results from sphericity test if contrast matrix $\boldsymbol{C}$ is not orthonormal
  - should use "applied to orthonormal components" test for assessing sphericity
- overall F-tests of main effects and interactions don't change if different contrasts are used
  - obviously, contrast results change

Sphericity Test - Mauchly, 1940

- low power for small sample
- for large sample, test is likely to be significant even though effect on F-test may be negligible
- sensitive to departures from normality and outliers
- use as a guide, not as a rule

If sphericity is rejected, or deemed implausible

- Multivariate repeated measures analysis (MANOVA)
  - allows for general $V(\boldsymbol{y}_i)$
  - but doesn't allow any missing data across time
- Adjusted univariate F-tests
  - Greenhouse-Geisser (1959) - too conservative
  - Huynh-Feldt (1976) - improved

Example 5:
Test of sphericity for the Pothoff & Roy data

SAS code and output
BB: Example 5: ANOVA-Ex5-pothoff-Sphericity.txt