Overview: Univariate & Multivariate Linear Models

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>1 Quantitative</th>
<th>2+ Quantitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent variables</td>
<td>( y = X \beta )</td>
<td>( Y = X B )</td>
</tr>
<tr>
<td>Quantitative</td>
<td>Regression</td>
<td>Multivariate regression</td>
</tr>
<tr>
<td>Categorical</td>
<td>ANOVA</td>
<td>MANOVA</td>
</tr>
<tr>
<td>Both</td>
<td>Reg. w/ dummy vars</td>
<td>General MLM</td>
</tr>
<tr>
<td></td>
<td>ANCOVA</td>
<td>Homogeneity of regression</td>
</tr>
<tr>
<td></td>
<td>Homogeneity of regression</td>
<td>MANCOVA</td>
</tr>
</tbody>
</table>

Today, just multivariate regression, with questions of homogeneity of regression. Once we learn how to do multivariate tests, extensions to other contexts are easy.

Multivariate regression

- When there are several \((p>1)\) criterion variables, we could just fit \(p\) separate models
  
  \[
  \begin{align*}
  y_1 &= X_1 \beta_1 \\
  y_2 &= X_2 \beta_2 \\
  & \quad \ldots \\
  y_p &= X_p \beta_p
  \end{align*}
  \]

  proc reg;
  model y1-y4 = x1 – x10;

- But this:
  - Does not give simultaneous tests for all regressions
  - Does not take correlations among the y’s into account

Why do multivariate tests?

- Avoid multiplying error rates, as in simple ANOVA
  - Overall test for multiple responses-- similar to overall test for many groups: \(g\) tests: \(\alpha_{\text{all}} = \alpha_g\)

- Often, multivariate tests are more powerful, when the responses are correlated
  - Small, positively correlated effects can pool power.
  - If responses are uncorrelated, no need for multivariate tests
    - But this is rarely so

- Multivariate tests provide a way to understand the structure of relations across separate response measures. In particular:
  - how many “dimensions” of responses are important?
  - how do the predictors contribute to these?
Multivariate regression model

- The **multivariate** regression model is

\[
\begin{bmatrix}
    y_1 \\
    \vdots \\
    y_p
\end{bmatrix} = \begin{bmatrix}
    x_1 \\
    \vdots \\
    x_q
\end{bmatrix} \begin{bmatrix}
    \beta_1 \\
    \vdots \\
    \beta_p
\end{bmatrix} + \begin{bmatrix}
    \varepsilon_{1,\ldots,p}
\end{bmatrix}
\]

\[
Y_{n\times p} = X_{n\times q}B_{q\times p} + \varepsilon_{n\times p}
\]

- The LS solution, \( B = (X'X)^{-1} X'Y \) gives **same coefficients** as fitting \( p \) models separately.

(Omitting here: consideration of model selection for each model)

Example: Rohwer data

- \( n=32 \) Lo SES kindergarten kids
- \( p=3 \) response measures of aptitude/achievement: SAT, PPVT, Raven
- \( q=5 \) predictors: PA tests: n, s, ns, na, ss

SAS:

```sas
proc reg data=lo_ses;
model sat ppvt raven = n s ns na ss;
M1: mtest /* all coeffs = 0 */
```

R:

```r
mod1<-lm(cbind(SAT, PPVT, Raven) ~ n+s+ns+na+ss, data=lo_ses)
Manova(mod1)
```

Rohwer data: univariate regressions

- **Separate univariate regressions**

<table>
<thead>
<tr>
<th></th>
<th>SAT</th>
<th>PPVT</th>
<th>Raven</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.208</td>
<td>0.511*</td>
<td>0.222</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>4.151</td>
<td>33.006</td>
<td>11.173</td>
</tr>
<tr>
<td>n</td>
<td>-0.609</td>
<td>-0.081</td>
<td>0.211</td>
</tr>
<tr>
<td>s</td>
<td>-0.050</td>
<td>-0.721</td>
<td>0.065</td>
</tr>
<tr>
<td>ns</td>
<td>-1.732</td>
<td>-0.298</td>
<td>0.216</td>
</tr>
<tr>
<td>na</td>
<td>0.495</td>
<td>1.470*</td>
<td>-0.037</td>
</tr>
<tr>
<td>ss</td>
<td>2.248*</td>
<td>0.324</td>
<td>-0.052</td>
</tr>
</tbody>
</table>

**Overall tests for each response:** \( H_0: \beta_i = 0 \)

Rohwer data: multivariate regression

- **Yet, the multivariate test is highly significant**
  - Overall test for the multivariate model: \( H_0: B = 0 \)
  - Positive correlations among responses have made this test more powerful – pooling power!

**Multivariate Statistics and F Approximations**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.343</td>
<td>2.54</td>
<td>15</td>
<td>80.46</td>
<td>0.0039</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>0.825</td>
<td>2.35</td>
<td>15</td>
<td>93</td>
<td>0.0066</td>
</tr>
<tr>
<td>Hotelling-Lawley</td>
<td>1.449</td>
<td>2.72</td>
<td>15</td>
<td>49.77</td>
<td>0.0042</td>
</tr>
<tr>
<td>Roy's Max Root</td>
<td>1.055</td>
<td>6.54</td>
<td>5</td>
<td>31</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

Publish or perish? Doesn't look like there is much predictive power here!
Multivariate General Linear Hypothesis (GLH)

- In addition to the overall test, $H_0: B = 0$, it is more often desired to test hypotheses about subsets of predictors or linear combinations of coefficients.
- The GLH is a single, general method for all such tests:

$$H_0: L_{r \times q} B_{q \times p} = 0_{r \times p}$$

where $L$ specifies $r$ linear combinations of the parameters.

**Examples of GLH:**

- **p=2 responses: $y_1, y_2$**
  - **q=3 predictors: $X_1 – X_3$**

(a) No effect of $X_2, X_3$

$$H_0: \beta_2 = \beta_3 = 0_{2 \times 1} \quad \Rightarrow \quad L = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \Rightarrow \quad LB = \begin{pmatrix} \beta_2^T \\ \beta_3^T \end{pmatrix} = 0$$

**mtest x2, x3;**

(b) Same coef. for $X_2, X_3$

$$H_0: \begin{cases} \beta_{21} = \beta_{31} \\ \beta_{22} = \beta_{32} \end{cases} \quad \Rightarrow \quad L = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix} \quad \Rightarrow \quad LB = \begin{pmatrix} \beta_2 - \beta_3 \end{pmatrix} = 0$$

**mtest x2-x3;**

(Makes sense only if $X_2, X_3$ are commensurate)

Extended GLH

- The GLH can be extended to test subsets or linear combinations of coefficients **across** the criteria:

$$H_0: L_{q \times q} B_{q \times p} M_{p \times t} = 0$$

where the post-factor, $M$, specifies $t$ linear combs. across criteria.

**Example: Coeffs for $Y1 = $ coeffs for $Y2**

$$L = I, \quad M = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \Rightarrow \quad LBM = \begin{pmatrix} \beta_{01} - \beta_{02} \\ \beta_{11} - \beta_{12} \\ \beta_{21} - \beta_{22} \\ \beta_{31} - \beta_{32} \end{pmatrix} = 0$$

**mtest y1-y2;**

(Again, makes sense only if $Y1$ and $Y2$ are commensurable)

**Note:** In MANOVA designs:

- $L$ specifies a set of contrasts or tests among 'between-group' effects
- $M$ specifies contrasts among 'within-subject' effects (e.g., orthogonal polynomials or other within-S comparisons)
Tests of multivariate hypotheses

- In the general linear model, \( Y = X \beta + \epsilon \), all hypotheses are tested in the same way.
- Calculate the \( q \times q \) sum of squares and products matrices:
  \[
  H = (LB)^T [L(X^TX)^{-1}L^T]^{-1}(LB) \\
  E = \hat{\epsilon}^T \hat{\epsilon}
  \]
  - Diag elements of \( H \) & \( E \) are just the univariate SS.
- The multivariate analog of the univariate F-test:
  \[
  F = \frac{MS_H}{MS_E} \rightarrow (MS_H - F \times MS_E) = 0 \quad \text{is} \quad |H - \lambda E| = 0
  \]

Tests of multivariate hypotheses

- All multivariate test statistics are based on latent roots, \( \lambda_i \) of \( H \) in the metric of \( E \) (or of \( HE^{-1} \)), or latent roots \( \theta_i \) of \( H(H+E)^{-1} \).
- These measure the "size" of \( H \) relative to \( E \) in up to \( p \) orthogonal dimensions.
- Various test statistics differ in how the information is combined across dimensions:
  - Wilks’ Lambda: product
  - Trace criteria: sum
  - Roy’s test: maximum

HE plots: Visualizing \( H \) & \( E \)

HE plots show the \( H \) & \( E \) matrices as data ellipsoids. It is difficult to judge naively the size of \( H \) relative to \( E \), but the eigenvalues of \( HE^{-1} \) capture the essential information.

Contributions of \( s = \min(p, df_H) \) dimensions can be summarized in different kinds of "means."

As explained later, this plot provides a visual test of significance, based on Roy’s test.

Multivariate test statistics: overview

- How big is \( H \) relative to \( E \) across one or more dimensions?
- All test statistics can be seen as kinds of means of the \( s = \min(p, df_H) \) non-zero eigenvalues of \( HE^{-1} \) or of \( H(H+E)^{-1} \).

Table 1: Multivariate test statistics as functions of the eigenvalues \( \lambda_i \) solving \( \det(H - \lambda E) = 0 \) or eigenvalues \( \rho_i \) solving \( \det(H - \rho(H + E)) = 0 \).

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Formula</th>
<th>“mean” of ( \rho )</th>
<th>Partial ( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks’ ( \Lambda )</td>
<td>( \Lambda = \prod_i \left( 1 - \frac{\lambda_i}{1 - \rho_i} \right) )</td>
<td>geometric</td>
<td>( \eta^2 = 1 - \Lambda^{1/s} )</td>
</tr>
<tr>
<td>Pillai trace</td>
<td>( V = \sum_i \frac{\lambda_i}{1 - \rho_i} )</td>
<td>arithmetic</td>
<td>( \eta^2 = \frac{V}{s} )</td>
</tr>
<tr>
<td>Hotelling-Lawley trace</td>
<td>( H = \sum_i \lambda_i )</td>
<td>harmonic</td>
<td>( \eta^2 = \frac{H}{tr} )</td>
</tr>
<tr>
<td>Roy maximum root</td>
<td>( R = \lambda_1 )</td>
<td>supremum</td>
<td>( \eta^2 = \frac{\lambda_1}{1 + \lambda_1} = \rho_1 )</td>
</tr>
</tbody>
</table>

(This table uses \( \rho \) instead of \( \theta \) for eigenvalues of \( H(H+E)^{-1} \).)
Multivariate test statistics: geometry

Easiest to see if we transform $H$ & $E$ to "canonical space" where

- $E \rightarrow E^* = I$ (standardized & uncorrelated)
- $(H+E) \rightarrow (H+E)^* = \sqrt{(H+E)} \ E^{1/2}$

Allows to focus on just size of $(H+E)^*$

Then,

- Wilks' $\Lambda \equiv$ test on area: $\sim \ (a \times b)^{-1}$
- HLT criterion $\sim$ test on $c$
- Pillai trace criterion $\sim$ test on $d$
- Roy's $\sim$ test on $a$ alone

Wilks' Lambda: details

- A likelihood-ratio test of $H_0$: $L \ B = 0$

\[
\Lambda = \frac{|E|}{|H+E|} = \prod_{i=1}^{s} \frac{1}{1 + \lambda_i} = \prod_{i=1}^{s} (1 - \theta_i) 
\]

- Rao's $F$ approximation (exact if $s \leq 2$)

\[
F = \frac{1 - \Lambda^{1/t}}{\Lambda^{1/t}} \times \frac{mt - 2k}{pn_n} \sim F(pn_n, mt - 2k) 
\]

- Association: $\eta^2 = 1 - \Lambda^{1/8} = \text{geometric mean}$

Multivariate test statistics: details

- $n_h = \text{df for hypothesis} = \# \text{rows of } L$
- $n_e = \text{df for error}$
- $s = \min(n_h, p) = \# \text{non-zero roots} = \text{rank}(H)$
- $\lambda_1, \lambda_2, \ldots, \lambda_s = \text{roots of } |H - \lambda \ E| = 0$
- $|HE^{-1} - \lambda I| = 0$

\[
\lambda_i = \frac{\theta_i}{1 - \theta_i} \quad \theta_i = \frac{\lambda_i}{1 + \lambda_i} 
\]

Pillai & Hotelling-Lawley trace criteria

- Based on sum (or average) of $\lambda_i$ or $\theta_i$
- Pillai:

\[
V = \text{tr} [H(H+E)^{-1}] = \sum_{i=1}^{s} \theta_i = \sum_{i=1}^{s} \frac{\lambda_i}{1 + \lambda_i} \quad \eta^2 = V/s 
\]

\[
F = \frac{2n + s + 1}{2m + s + 1} \times \frac{V}{s - V} \sim F[s(2m + s + 1), s(2n + s + 1)] 
\]

- Hotelling-Lawley:

\[
H = \text{tr} [HE^{-1}] = \sum_{i=1}^{s} \lambda_i = \sum_{i=1}^{s} \frac{\theta_i}{1 - \theta_i} \quad \eta^2 = H/(H+s) 
\]

\[
F = \frac{2(ns + 1)}{s^2(2m + s + 1)} \times H \sim F[s(2m + s + 1), 2(ns + 1)] 
\]
Roy's maximum root test

- Most powerful when there is one large dimension of $H$ relative to $E$
- $R = \lambda_1$
- $\eta^2 = R/(R+1)$
- $F = \frac{n_e + n_h - s}{s} \sim F(s, n_e + n_h - s)$ (Exact if $s=1$)
- Simplicity makes this useful for visual tests of significance in HE plots

Multivariate tests for individual predictors

- $H_0 : \beta_i = 0$ – simultaneous test, for all $p$ responses, of predictor $x_i$
  - $L = (0,0, \ldots, 1, 0, \ldots 0)_{(1 \times q)}$ in GLH
  - $H = \beta_i^T (X^T X)^{-1} \beta_i$ – a rank 1 matrix ($s=1$)
  - All multivariate tests are exact & have the same $p$-values
  - More parsimonious than separate univariate tests for each response.

Example: Rohwer data (SAS)

```sas
proc reg data=lo_ses;
   model sat ppvt raven = n s ns na ss ;
   Mn: mtest n; /* n=0 for all responses */
   Mna: mtest na; /* na=0 for all responses */
run;
```

Output for NA:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.733</td>
<td>3.53</td>
<td>3</td>
<td>29</td>
<td>0.0271</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>0.267</td>
<td>3.53</td>
<td>3</td>
<td>29</td>
<td>0.0271</td>
</tr>
<tr>
<td>Hotelling-Lawley</td>
<td>0.365</td>
<td>3.53</td>
<td>3</td>
<td>29</td>
<td>0.0271</td>
</tr>
<tr>
<td>Roy's Max Root</td>
<td>0.365</td>
<td>3.53</td>
<td>3</td>
<td>29</td>
<td>0.0271</td>
</tr>
</tbody>
</table>

Example: Rohwer data (R)

```r
> Manova(mod1)
```

Type II MANOVA Tests: Pillai test statistic

<table>
<thead>
<tr>
<th>Df</th>
<th>test stat</th>
<th>approx F</th>
<th>num Df</th>
<th>den Df</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1</td>
<td>0.0384</td>
<td>0.3856</td>
<td>3</td>
<td>29 0.76418</td>
</tr>
<tr>
<td>s</td>
<td>1</td>
<td>0.1118</td>
<td>1.2167</td>
<td>3</td>
<td>29 0.32130</td>
</tr>
<tr>
<td>ns</td>
<td>1</td>
<td>0.2252</td>
<td>2.8100</td>
<td>3</td>
<td>29 0.05696</td>
</tr>
<tr>
<td>na</td>
<td>1</td>
<td>0.2675</td>
<td>3.5294</td>
<td>3</td>
<td>29 0.02705</td>
</tr>
<tr>
<td>ss</td>
<td>1</td>
<td>0.1390</td>
<td>1.5601</td>
<td>3</td>
<td>29 0.22030</td>
</tr>
</tbody>
</table>

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Note: Manova() and Anova() in the car package are identical
They give a compact summary for all predictors, for a given test statistic
Gory details are available from the summary() method
Canonical analysis: How many dimensions?

- Sequential tests for the latent roots $\lambda_i$ of $HE^{-1}$ indicate the number of dimensions of the $y$s predicted by the $x$s.
- Canonical correlations: correlation of best linear combination of $y$s with best of $x$s

$$\lambda_i = \frac{\rho^2}{1 - \rho^2} \quad \rho^2 = \frac{\lambda}{1 - \lambda}$$

```
proc reg data=lo_ses;
model sat ppvt raven = n s ns na ss;
M1: mtest / canprint;    /* all coeffs = 0 */
run;
```

Canonical analysis: How many dimensions?

<table>
<thead>
<tr>
<th>Canonical Correlation</th>
<th>Adjusted Canonical Correlation</th>
<th>Approximate Standard Error</th>
<th>Squared Canonical Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.716526</td>
<td>0.655198</td>
<td>0.081098</td>
</tr>
<tr>
<td>2</td>
<td>0.490621</td>
<td>0.414578</td>
<td>0.126849</td>
</tr>
<tr>
<td>3</td>
<td>0.266778</td>
<td>0.211906</td>
<td>0.154905</td>
</tr>
</tbody>
</table>

Eigenvalues of $Inv(E)^*H = CanRsq/(1-CanRsq)$

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>Difference</th>
<th>Proportion</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0551</td>
<td>0.7283</td>
<td>0.7283</td>
</tr>
<tr>
<td>2</td>
<td>0.3170</td>
<td>0.2188</td>
<td>0.9471</td>
</tr>
<tr>
<td>3</td>
<td>0.0766</td>
<td>1.0000</td>
<td></td>
</tr>
</tbody>
</table>

Test of $H_0$: The canonical correlations in the current row and all that follow are zero

<table>
<thead>
<tr>
<th>Likelihood Ratio</th>
<th>Approximate F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.34316907</td>
<td>2.54</td>
<td>15</td>
<td>0.0039</td>
</tr>
<tr>
<td>2</td>
<td>0.70525204</td>
<td>1.43</td>
<td>8</td>
<td>0.2025</td>
</tr>
<tr>
<td>3</td>
<td>0.92882959</td>
<td>0.79</td>
<td>3</td>
<td>0.5078</td>
</tr>
</tbody>
</table>

Wilks’ Lambda only 1 signif. dim

Visualizing multivariate tests: HE plots

- The $H$ and $E$ matrices in the GLH summarize the (co)variation of the fitted values and residuals for a given effect

$$H = \hat{Y}_{eff}^T \hat{Y}_{eff} \quad E = \mathcal{E}^T \mathcal{E}$$

- For two variables, we can visualize their size & shape with data ellipses
- For $p=3$ these display as ellipsoids
- For $p>2$ can use an HE-plot matrix

Data ellipses for $H$ & $E$

How big is $H$ relative to $E$?
How to make them comparable?
HE plot: effect scaling

- Scale: \( E/dfe, H/dfe \)
- Center: shift to centroid
- Plot: 68\% data ellipses

For each predictor, the data ellipse degenerates to a line (rank \( H: s=1 \))

- **Orientation**: how \( x_i \) contributes to prediction of \( y_1, y_2 \)
- **Length**: relative strength of relation

HE plot: significance scaling

Scale:
- \( E: E/dfe \)
- \( H: H/dfe \lambda_\alpha \)

\( \lambda_\alpha \) = critical value of Roy statistic at level \( \alpha \)

\( \rightarrow \) any \( H \) ellipse will protrude beyond \( E \) ellipse \( iff \) effect is significant at level \( \alpha \)

Directions of \( H \)s show how predictors contribute to responses

HE plot: significance scaling

Rohwer data, low SES gp

- Overall test highly significant
- Only NA individually significant (in this view)
- NA & S contribute to predicting PPVT
- NS & SS contribute to SAT
- N doesn’t contribute at all

HE plot matrix

All pairwise views

An effect is significant if \( H \) projects outside \( E \) in any view

That applies to any rotation, not just the bivariate views shown here.
3D HE plot

In the R version (heplots package), 3D plots can be rotated dynamically.

In this view, we see NA poking out beyond the E ellipsoid.

Homogeneity of Regression

- When there are several groups, we often want to test hypotheses of "homogeneity":
  - equal slopes for the predictors (interactions)?
  - equal intercepts for groups (same means)?
  - equal slopes & intercepts (coincident regressions)?

```
*-- test equal slopes, by allowing interactions (separate slopes for each group);
  proc glm data=rohwer;
    class SES;
    model sat ppvt raven = SES|n SES|s SES|ns SES|na SES|ss /ss3 nouni;
    manova h=SES*n SES*s SES*ns SES*na SES*ss;
  run;

*-- MANCOVA model: test intercepts (means), assuming equal slopes;
  proc glm data=rohwer;
    class SES;
    model sat ppvt raven = SES n s ns na ss /ss3 nouni;
    manova h=_all_;  
  run;
```

NB: better than reporting separate results and making "eyeball" comparisons.

HE plots: Homogeneity of regression

Rohwer data: Lo (n=32) & Hi (n=37) SES groups:
- Fit separate regressions for each group
- Are slopes the same?
- Are intercepts the same?
- Are regressions coincident? (equal slopes and intercepts)

Here, slopes for NS are similar; most others seem to differ, but only NA is signif.
Intercepts (means) clearly differ.

HE plots: MANCOVA model

Alternatively, we can fit a model that assumes equal slopes for both SES groups, but allows unequal intercepts.

From the ANOVA view, this is equivalent to an analysis of covariance model, with group effect and quantitative predictors.

If the main interest is in the SES effect, the MANCOVA test relies on the assumption of equal slopes.
HE plots: MANCOVA model

Nature vs Nurture: IQ of adopted children
MMReg + Repeated measures

- Data from an observational, longitudinal, study on adopted children (n=62).
- Is child’s intelligence related to intelligence of the biological mother and the intelligence of the adoptive mother?
- The child’s intelligence was measured at age 2, 4, 8, and 13
- How does intelligence change over time?
- How are these changes related to intelligence of the birth and adoptive mother?

Data from an observational, longitudinal, study on adopted children (n=62).

- Is child’s intelligence related to intelligence of the biological mother and the intelligence of the adoptive mother?
- The child’s intelligence was measured at age 2, 4, 8, and 13
- How does intelligence change over time?
- How are these changes related to intelligence of the birth and adoptive mother?

Some (Adopted)
AMED BMIQ Age2IQ Age4IQ Age8IQ Age13IQ
3 14 89 126 115 113 90
6 8 64 125 109 96 87
23 6 92 116 121 119 109
30 13 78 108 90 86 80
31 16 87 113 97 101 109
32 15 63 127 121 119 101
40 8 95 140 130 126 118
42 15 65 110 111 114 95
52 13 74 121 132 132 113
58 11 88 112 107 110 103

Treat as multivariate regression problem:

\[
\begin{align*}
\text{AMED: Adoptive mother educ. (proxy for IQ)} \\
\text{BMIQ: Birth mother IQ}
\end{align*}
\]

What can we tell from this?

Scatterplots of child IQ vs. AMED and BMIQ
- Regression lines (red) show the fitted (univariate) relations
- Data ellipses: visualize strength of relations
- Blue lines: equality of child IQ and BMIQ

What can we tell from this?
Multivariate tests of each predictor: \( \beta_{\text{AMED}} = 0; \beta_{\text{BMIQ}} = 0 \)

Conclusions from this:
- Birth mother IQ significantly predicts child IQ at these ages: \( \beta_{\text{BMIQ}} \neq 0 \)
- Adoptive mother ED does not: \( \beta_{\text{AMED}} = 0 \)

How to understand the nature of these relations?

Repeated measures analysis

- Because Age is a quantitative factor, we can use it in a multivariate trend analysis.
- This amounts to analysis of \( \mathbf{Y} \mathbf{M} \), where \( \mathbf{M} \) comes from
  \[
  \mathbf{M} = \begin{bmatrix}
  2 & 2^2 & 3^2 \\
  4 & 4^2 & 4^3 \\
  8 & 8^2 & 8^3 \\
  13 & 13^2 & 13^3
  \end{bmatrix}
  \]

- This gives tests of linear, quadratic & cubic trends of IQ in relation to AMED and BMIQ
- Interactions– AMED*Age & BMIQ*Age test for equal slopes over Age
# Treat IQ at different ages as a repeated measure factor

```r
Age <- data.frame(Age=ordered(c(2,4,8,13)))
Anova(Adopted.mod, idata=Age, idesign=~Age, test="Roy")
```

Type II Repeated Measures MANOVA Tests: Roy test statistic

<table>
<thead>
<tr>
<th>Df</th>
<th>test stat</th>
<th>approx F</th>
<th>num Df</th>
<th>den Df</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMED</td>
<td>1</td>
<td>0.0019</td>
<td>0.1131</td>
<td>1</td>
<td>59</td>
</tr>
<tr>
<td>BMIQ</td>
<td>1</td>
<td>0.1265</td>
<td>7.4612</td>
<td>1</td>
<td>59</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>0.7120</td>
<td>13.5287</td>
<td>3</td>
<td>57</td>
</tr>
<tr>
<td>AMED:Age</td>
<td>1</td>
<td>0.0143</td>
<td>0.2718</td>
<td>3</td>
<td>57</td>
</tr>
<tr>
<td>BMIQ:Age</td>
<td>1</td>
<td>0.1217</td>
<td>2.3114</td>
<td>3</td>
<td>57</td>
</tr>
</tbody>
</table>

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

```r
heplot(Adopted.mod, idata=Age, idesign=~Age, iterm="Age", + hypoth=list("Reg"=c("AMED", "BMIQ")))
mark.H0()
```

## HE plots: software

- **SAS macros**

- **R packages**
  - heplots & car packages: [http://www.r-project.org/](http://www.r-project.org/)

## Summary

- **MMRA** → multivariate tests for a collection of \( p \) responses, each in up to \( s \) dimensions
  - Different test statistics combine these in different ways, to say **how big is \( H \) vs \( E \)**
  - Canonical analysis: **How many dimensions** of Ys are predicted by the Xs?
  - **HE plots** → visualize the relations of responses to the predictors
  - These methods generalize to all linear models: MANOVA, MANCOVA, etc.