The Multivariate General Linear Model

The Univariate General Linear Model A Review

$$\mathbf{y}_{N\times 1} = \mathbf{X}_{N\times q}\boldsymbol{\beta}_{q\times 1} + \boldsymbol{\varepsilon}_{N\times 1}$$

Assumption: $\boldsymbol{\varepsilon} \sim \mathrm{N}_N\left(\mathbf{0}, \sigma^2\mathbf{I}_N\right)$

The Univariate General Linear Hypothesis

$$\mathrm{H}_0 : \mathbf{C}_{k\times q}\boldsymbol{\beta}_{q\times 1} = \boldsymbol{\gamma}_{k\times 1}$$

$$\mathrm{H}_1 : \mathbf{C}_{k\times q}\boldsymbol{\beta}_{q\times 1} \neq \boldsymbol{\gamma}_{k\times 1}$$

The $\mathbf{C}$ matrix defines linear combinations of the elements of $\boldsymbol{\beta}$. In ANOVA, it is used for linear contrasts and linear combinations among the means μ_i .

Test Statistic for Testing the General Linear Hypothesis

$$\mathrm{F} = \frac{\mathrm{SS}_H / \nu_H}{\mathrm{SS}_E / \nu_E} = \frac{\mathrm{MS}_H}{\mathrm{MS}_E} \sim \mathrm{F}_{\nu_H, \nu_E}$$

where $$\mathrm{SS}_H = \left(\mathbf{C}\boldsymbol{\beta} - \boldsymbol{\gamma}\right)' \left[\mathbf{C}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{C}'\right]^{-1}\left(\mathbf{C}\boldsymbol{\beta} - \boldsymbol{\gamma}\right)$$

$$\mathrm{SS}_E = \mathbf{y}'\left[\mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\right]\mathbf{y}$$

$$\nu_H = \mathrm{rank}(\mathbf{C}) = k \qquad \nu_E = N - k$$

The Univariate One-Sample Problem

The model for the univariate one-sample problem is often expressed as:

$$y_i = \mu + \varepsilon_i \qquad \text{for } i = 1, \cdots, N$$

The Univariate One-Sample Problem

In matrix notation, the model becomes:

$$\mathbf{y} = \mathbf{X}\ \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \mu + \boldsymbol{\varepsilon}$$

The Univariate
One-Sample Problem

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \boldsymbol{\gamma}$$

or

$$H_0 : 1\mu = \mu_0$$

The Univariate
Two-Sample Problem

The model for the univariate two-sample problem is often expressed as:

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \qquad \text{for } i = 1, \cdots, N;\ j = 1, 2$$

where

$$\sum_{j=1}^{2} \alpha_j = \alpha_1 + \alpha_2 = 0$$

The α's are sometimes referred to as (treatment) effects.

The Univariate Two-Sample Problem

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} y_{11} \\ \vdots \\ y_{1N_1} \\ y_{21} \\ \vdots \\ y_{2N_2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \end{bmatrix} + \boldsymbol{\varepsilon}$$

The Univariate Two-Sample Problem

$\mathrm{H}_0 : \mu_1 = \mu_2$

is equivalent to

$\mathrm{H}_0 : \alpha_1 = \alpha_2 = 0$

which in terms of the GLH

$\mathrm{H}_0 : \mathbf{C}\boldsymbol{\beta} = \boldsymbol{\gamma}$

becomes

$$\mathrm{H}_0 : \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \end{bmatrix} = 0$$

The Univariate One-Way ANOVA

The model for the k-sample one-way ANOVA is often expressed as:

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \qquad \text{for } i = 1, \cdots, N;\ j = 1, \cdots, k$$

where

$$\sum_{j=1}^{k} \alpha_j = 0$$

The α's are sometimes referred to as (treatment) effects.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} y_{11} \\ \vdots \\ y_{1N_1} \\ y_{21} \\ \vdots \\ y_{2N_2} \\ \vdots \\ y_{k1} \\ \vdots \\ y_{kN_k} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & -1 & -1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & -1 & -1 & \cdots & -1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{k-1} \end{bmatrix} + \boldsymbol{\varepsilon}$$

The Univariate One-Way ANOVA

$$\mathrm{H}_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

is equivalent to

$$\mathrm{H}_0 : \alpha_1 = \alpha_2 = \cdots \alpha_k = 0$$

The Univariate One-Way ANOVA

which in terms of the GLH becomes

$$\mathrm{H}_0 : \quad \mathbf{C} \quad \boldsymbol{\beta} = \boldsymbol{\gamma}$$

or

$$\mathrm{H}_0 : \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{k-1} \end{bmatrix} = 0$$

The Multivariate General Linear Model

$$\mathbf{Y}_{N\times p} = \mathbf{X}_{N\times q}\mathbf{B}_{q\times p} + \mathbf{E}_{N\times p}$$

Assumption: $\mathbf{E} \sim \mathrm{N}_{Np}\left(\mathbf{0}, \mathbf{I}_N \otimes \mathbf{\Sigma}\right)$

The Multivariate General Linear Hypothesis

$$\mathrm{H}_0 : \mathbf{C}_{k\times q}\mathbf{B}_{q\times p}\mathbf{A}_{p\times r} = \mathbf{\Gamma}_{k\times r}$$

$$\mathrm{H}_1 : \mathbf{C}_{k\times q}\mathbf{B}_{q\times p}\mathbf{A}_{p\times r} \neq \mathbf{\Gamma}_{k\times r}$$

The $\mathbf{C}$ matrix defines linear combinations of the rows of $\mathbf{B}$, and is used for between-subjects effects in MANOVA.

The $\mathbf{A}$ matrix defines linear combinations of the columns of $\mathbf{B}$, and is used for within-subjects effects (e.g., paired comparisons and repeated measures effects).

The Multivariate General Linear Hypothesis

$$\mathrm{H}_0 : \mathbf{C}_{k\times q}\mathbf{B}_{q\times p}\mathbf{A}_{p\times r} = \mathbf{\Gamma}_{k\times r}$$

$$\mathrm{H}_1 : \mathbf{C}_{k\times q}\mathbf{B}_{q\times p}\mathbf{A}_{p\times r} \neq \mathbf{\Gamma}_{k\times r}$$

Also:

$$\mathrm{rank}(\mathbf{C}) = k \leq q$$

$$\mathrm{rank}(\mathbf{A}) = r \leq p$$

Sums of Squares and Cross-Products Matrices associated with the GLH

$$\mathbf{H} = \left(\mathbf{CBA} - \mathbf{\Gamma}\right)' \left[\mathbf{C}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{C}'\right]^{-1}\left(\mathbf{CBA} - \mathbf{\Gamma}\right)$$

$$\mathbf{E} = \mathbf{A}'\mathbf{Y}'\left[\mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\right]\mathbf{YA}$$

Test Statistics for Testing the Multivariate General Linear Hypothesis

- In the univariate GLM, the test statistic was:

$$F = \frac{SS_H / \nu_H}{SS_E / \nu_E}$$

- In the multivariate GLM, the quantities SS_H and SS_E become the matrices $\mathbf{H}$ and $\mathbf{E}$ respectively.

Test Statistics for Testing the Multivariate General Linear Hypothesis

- Any test statistic for testing the multivariate general linear hypothesis will be a function of the elements of $\mathbf{H}$ and $\mathbf{E}$ (or $\mathbf{H}$ and $\mathbf{T}=\mathbf{H}+\mathbf{E}$).
- In fact, the test statistics that were developed for this purpose are all functions of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ of $\mathbf{HE}^{-1}$.

Wilk's Λ
(The Maximum Likelihood Test)

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{H} + \mathbf{E}|}$$

$$= \frac{1}{|\mathbf{I} + \mathbf{HE}^{-1}|}$$

$$= \prod_{i=1}^{s} (1 + \lambda_i)^{-1}$$

Pillai's Trace

$$\text{V} = \text{tr}\left[\mathbf{H}(\mathbf{H} + \mathbf{E})^{-1} \right]$$

$$= \text{tr}\left[\mathbf{I} + \left(\mathbf{HE}^{-1}\right)^{-1} \right]^{-1}$$

$$= \sum_{i=1}^{s} \frac{\lambda_i}{1 + \lambda_i}$$

Lawley-Hotelling Trace

$$\mathrm{U} = \mathrm{tr}\left(\mathbf{HE}^{-1}\right) = \sum_{i=1}^{s} \lambda_i$$

Roy's Largest Root (The Union-Intersection Test)

- Can be seen two ways in practice:

  (a) $\lambda_1 =$ largest eigenvalue of $\mathbf{HE}^{-1}$

  (b) $\theta_1 =$ largest eigenvalue of $\mathbf{H}\left(\mathbf{H}+\mathbf{E}\right)^{-1}$

- Since λ_1 and θ_1 are monotonic transformations of one another, the use of either one leads to the same conclusion.

Parameters associated with the Multivariate Test Criteria for testing the GLH

$$H_0 : \mathbf{C}_{k\times q}\mathbf{B}_{q\times p}\mathbf{A}_{p\times r} = \mathbf{\Gamma}_{k\times r}$$

$$s = \min(k, r)$$

$$m = \frac{1}{2}\left(|k-r|-1\right)$$

$$n = \frac{1}{2}\left(\nu_E - r - 1\right)$$

Simultaneous Confidence Intervals for $\mathbf{c}'\mathbf{B}\mathbf{a}$

A 100(1-a)% C. I. for $\mathbf{c}'\mathbf{B}\mathbf{a}$ is:

$$\mathbf{c}'\mathbf{B}\mathbf{a} \pm c_0\sqrt{\mathbf{a}'\left(\frac{\mathbf{E}}{\nu_E}\right)\mathbf{a}\;\mathbf{c}'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{c}}$$

Where c_0 is defined as:

Wilk's: $c_0^2 = \nu_E \frac{1-\Lambda_\alpha}{\Lambda_\alpha}$, where $\Lambda_\alpha = \frac{1}{\left(1+\frac{\nu_1}{\nu_2}F_{\nu_1,\nu_2;\alpha}\right)^t}$

Pillai: $c_0^2 = \nu_E \frac{V_\alpha}{1-V_\alpha}$, where $V_\alpha = s\frac{\nu_1}{\nu_2}F_{\nu_1,\nu_2;\alpha}$

Lawley-Hotelling: $c_0^2 = \nu_E U_\alpha$, where $U_\alpha = s\frac{\nu_1}{\nu_2}F_{\nu_1,\nu_2;\alpha}$

Roy: $c_0^2 = \nu_E \frac{\theta_\alpha}{1-\theta_\alpha}$

Bonferroni t: $c_0 = t_{\nu_E;\alpha/2r}$, where $r = \#\text{C.I.'}s$

The Multivariate One-Sample Problem

$$\mathbf{Y} = \mathbf{X} \quad \mathbf{B} + \mathbf{E}$$

$$\begin{bmatrix} \mathbf{y}_1' \\ \mathbf{y}_2' \\ \vdots \\ \mathbf{y}_N' \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_p \end{bmatrix} + \mathbf{E}$$

$$\mathbf{Y} = \mathbf{1} \quad \boldsymbol{\mu}' + \mathbf{E}$$

The Multivariate One-Sample Problem

$$\mathrm{H}_0 : \mathbf{CBA} = \boldsymbol{\Gamma}$$

or

$$\mathrm{H}_0 : 1\,\boldsymbol{\mu}'\mathbf{I}_p = \boldsymbol{\mu}_0'$$

which simplifies to

$$\mathrm{H}_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$$

The Multivariate One-Sample Problem

For constructing the test statistics for this hypothesis, the hypothesis and error sums of squares and cross-products matrices become:

$$\mathbf{H} = N\left(\overline{\mathbf{y}} - \boldsymbol{\mu}_0\right)\left(\overline{\mathbf{y}} - \boldsymbol{\mu}_0\right)'$$

$$\mathbf{E} = (N-1)\mathbf{S}$$

The Multivariate One-Sample Problem

The associated multivariate test statistics are:

Wilk's: $\Lambda = \dfrac{1}{\left|\mathbf{I} + \mathbf{HE}^{-1}\right|} = \dfrac{1}{1 + \dfrac{1}{N-1}\mathrm{T}^2}$

Pillai's: $V = \dfrac{\lambda_1}{1+\lambda_1} = \dfrac{1}{1 + \dfrac{1}{\mathrm{T}^2/(N-1)}}$

The Multivariate One-Sample Problem

The associated multivariate test statistics are:

Lawley-Hotelling: $U = \dfrac{\mathrm{T}^2}{(N-1)}$

Roy's Largest Root: $\lambda_1 = \dfrac{\mathrm{T}^2}{(N-1)}$

The Multivariate One-Sample Problem

Thus, since each of these four multivariate test criteria are monotonic functions of each other, as well as Hotelling's T^2, for the one-sample problem, all test criteria are equivalent.

(This is true for the one-sample problem and the two-sample problem, but not necessarily for other problems.)

The Multivariate Two-Sample Problem

$$\mathbf{Y} = \mathbf{X} \mathbf{B} + \mathbf{E}$$

$$\begin{bmatrix} \mathbf{y}'_{11} \\ \vdots \\ \mathbf{y}'_{1N_1} \\ \mathbf{y}'_{21} \\ \vdots \\ \mathbf{y}'_{2N_2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\alpha}'_1 \end{bmatrix} + \mathbf{E}$$

The Multivariate Two-Sample Problem

$\mathrm{H}_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$

is equivalent to

$\mathrm{H}_0 : \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \mathbf{0}$

which in terms of the GLH

$\mathrm{H}_0 : \quad \mathbf{C} \quad \mathbf{B} \quad \mathbf{A} = \boldsymbol{\Gamma}$

becomes

$$\mathrm{H}_0 : \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\alpha}_1' \end{bmatrix} \mathbf{I}_p = \mathbf{0}$$

The Multivariate Two-Sample Problem

For constructing the test statistics for this hypothesis, the hypothesis and error sums of squares and cross-products matrices become:

$$\mathbf{H} = \frac{N_1 N_2}{N_1 + N_2} \left(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2 \right) \left(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2 \right)'$$

$$\mathbf{E} = (N_1 + N_2 - 2) \mathbf{S}_p$$

The Multivariate One-Way ANOVA:

$$\underset{\mathbf{Y}}{\begin{bmatrix} \mathbf{y}'_{11} \\ \vdots \\ \mathbf{y}'_{1N_1} \\ \mathbf{y}'_{21} \\ \vdots \\ \mathbf{y}'_{2N_2} \\ \vdots \\ \mathbf{y}'_{k1} \\ \vdots \\ \mathbf{y}'_{kN_k} \end{bmatrix}} = \underset{\mathbf{X}}{\begin{bmatrix} 1 & 1 & 0 & & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & -1 & -1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & -1 & -1 & \cdots & -1 \end{bmatrix}} \underset{\mathbf{B}}{\begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\alpha}'_1 \\ \boldsymbol{\alpha}'_2 \\ \vdots \\ \boldsymbol{\alpha}'_{k-1} \end{bmatrix}} + \mathbf{E}$$

The Multivariate One-Way ANOVA

$$\mathrm{H}_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k$$

is equivalent to

$$\mathrm{H}_0 : \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \cdots \boldsymbol{\alpha}_k = \mathbf{0}$$

The Multivariate One-Way ANOVA

which in terms of the GLH becomes

$$\mathrm{H}_0: \quad \mathbf{C} \quad \mathbf{B} = \boldsymbol{\Gamma}$$

or

$$\mathrm{H}_0: \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\alpha}_1' \\ \boldsymbol{\alpha}_2' \\ \vdots \\ \boldsymbol{\alpha}_{k-1}' \end{bmatrix} = \mathbf{0}$$

The Multivariate One-Way ANOVA

Assume $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}$

Let $\mathbf{S}_i = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} \left(\mathbf{y}_{ij} - \overline{\mathbf{y}}_i\right)\left(\mathbf{y}_{ij} - \overline{\mathbf{y}}_i\right)'$

Let $N = \sum_{i=1}^{k} N_i$

Then $\mathbf{S}_p = \frac{1}{N-k} \sum_{i=1}^{k} \left(N_i - 1\right) \mathbf{S}_i$ estimates $\boldsymbol{\Sigma}$

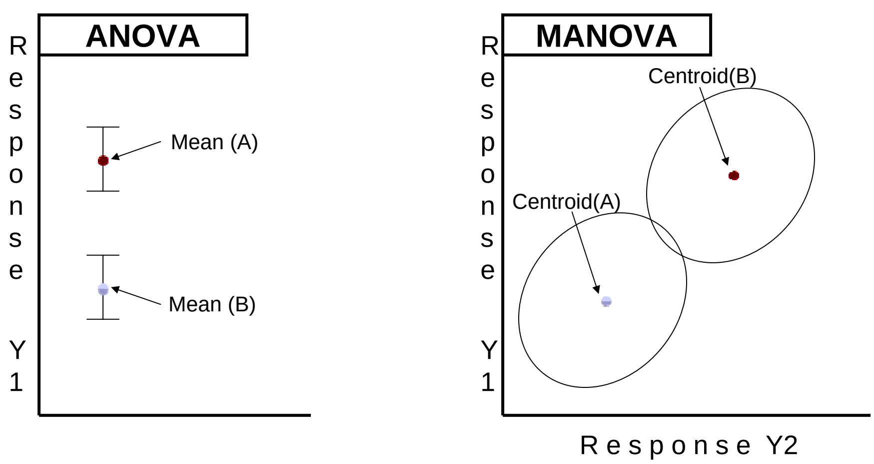
The Multivariate One-Way ANOVA

Let $\mathbf{E} = (N-k)\mathbf{S}_p$

$$\mathbf{H} = \sum_{i=1}^{k} N_i \left(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}\right)\left(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}\right)'$$

Then the multivariate test statistics can be constructed from the eigenvalues of $\mathbf{H}\mathbf{E}^{-1}$.

ANOVA versus MANOVA



Statistical Advantages of MANOVA

Compared to ANOVA with multiple dependent variables, MANOVA

- reduces overall type-I error rate
- accounts for important information such as correlation among the dependent variables
- accounts for joint effects in the responses that would be missed otherwise in univariate tests (MANOVA increases power)
- allows you to examine multiple scores to screen for overall differences without combining scores into a single composite.

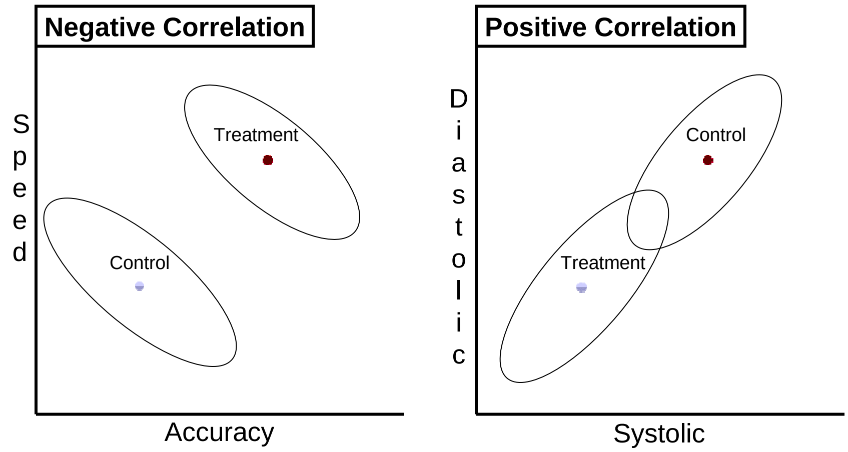
(Stevens 1996)

Assumptions of MANOVA

- Random samples
- Independent observations
- Multivariate normality
- Homogeneity of covariance matrices

Collinear Dependent Variables

A special problem in multivariate linear models:



Typing Skills **Blood Pressure**

Bray and Maxwell 1985

Collinear Dependent Variables

- Unlike ANOVA, MANOVA accounts for the relationships among the DVs as well as any effects of the IVs. It is possible to have multivariate significance and no univariate significance and vice versa.
- In the two pictures above, notice the relationships between the treatment method and the responses (this example is based on Bray and Maxwell 1985). The univariate ANOVAs would be the same for each of the two examples (the simple descriptive statistics are equal in the two examples).

Collinear Dependent Variables

- In the picture of the typing study results on the left, there is a negative correlation between the responses, accuracy and speed. This negative correlation actually magnifies the distance between the centroids of the control and the treatment groups. The multivariate effect is larger than either univariate effect.
- Now consider the blood pressure example in the picture on the right, above. Systolic and diastolic blood pressure measurements are positively correlated. Notice that the centroids are closer together in this example, even though the univariate statistics are the same as in the typing example. The multivariate effect is smaller than either univariate effect.

Sample Size

Most multivariate analyses are large-sample procedures.

Rules of thumb for **minimum** sample size:

– greater of 100 or 5 times the number of parameters

or

– 20+ observations per group.

For small-effects sizes and large variances, larger samples are necessary for adequate statistical power.