MANOVA

M.Sinan İyisoy

ANOVA

- Tek değişkenli ANOVA da etki modeli

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases}$$

- μ genel ortalama
- τ_i işlem (treatment) etkisi
- ε_{ij} hata

$$\sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2$$

$$SS_T = SS_{\text{Treatments}} + SS_E \qquad F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$$

MANOVA

	Sample 1 from $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$	Sample 2 from $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$	...	Sample k from $N_p(\boldsymbol{\mu}_k, \boldsymbol{\Sigma})$
Vektör	$\mathbf{y}_{11}$	$\mathbf{y}_{21}$	$\cdots$	$\mathbf{y}_{k1}$
	$\mathbf{y}_{12}$	$\mathbf{y}_{22}$	$\cdots$	$\mathbf{y}_{k2}$
	$\vdots$	$\vdots$		$\vdots$
	$\mathbf{y}_{1n}$	$\mathbf{y}_{2n}$	$\cdots$	$\mathbf{y}_{kn}$
Total	$\mathbf{y}_{1.}$	$\mathbf{y}_{2.}$	$\cdots$	$\mathbf{y}_{k.}$
Mean	$\overline{\mathbf{y}}_{1.}$	$\overline{\mathbf{y}}_{2.}$	$\cdots$	$\overline{\mathbf{y}}_{k.}$

- Etki modeli

$$\begin{aligned}\mathbf{y}_{ij} &= \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\varepsilon}_{ij} \\ &= \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{ij}, i = 1, 2, \ldots, k; \quad j = 1, 2, \ldots, n.\end{aligned}$$

- Her vektör aşağıdaki gibi yazılır

$$\begin{pmatrix} y_{ij1} \\ y_{ij2} \\ \vdots \\ y_{ijp} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} + \begin{pmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \vdots \\ \alpha_{ip} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijp} \end{pmatrix} = \begin{pmatrix} \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{ip} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijp} \end{pmatrix}$$

- Bu durumda bir vektördeki her bir gözlem

$$y_{ijr} = \mu_r + \alpha_{ir} + \varepsilon_{ijr} = \mu_{ir} + \varepsilon_{ijr}. \quad (r = 1, 2, \ldots, p)$$

şeklinde yazılabilir.

Yokluk hipotezi

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k$$

$$H_0 : \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1p} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{2p} \end{pmatrix} = \cdots = \begin{pmatrix} \mu_{k1} \\ \mu_{k2} \\ \vdots \\ \mu_{kp} \end{pmatrix}.$$

SSH, SSE ve SST

- ANOVA daki grup içi kareler toplamı (SSH) ve gruplar arası kareler toplamı (SSE) ve toplam kareler toplamı

$$SSH = n\sum_{i=1}^{k}(\bar{y}_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^{k}\frac{y_{i.}^2}{n} - \frac{y_{..}^2}{kn},$$

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2 = \sum_{ij} y_{ij}^2 - \sum_{i}\frac{y_{i.}^2}{n}.$$

$$\mathbf{SST} = \sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n} {y_{ij}}^2 - \frac{\bar{y}_{..}}{nk}$$

MANOVA da benzer şekilde tanımlanır.

$$\mathbf{H} = n\sum_{i=1}^{k}(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})'$$
$$= \sum_{i=1}^{k}\frac{1}{n}\mathbf{y}_{i.}\mathbf{y}_{i.}' - \frac{1}{kn}\mathbf{y}_{..}\mathbf{y}_{..}',$$

$$\mathbf{E} = \sum_{i=1}^{k}\sum_{j=1}^{n}(\mathbf{y}_{ij} - \overline{\mathbf{y}}_{i.})(\mathbf{y}_{ij} - \overline{\mathbf{y}}_{i.})'$$
$$= \sum_{ij}\mathbf{y}_{ij}\mathbf{y}_{ij}' - \sum_{i}\frac{1}{n}\mathbf{y}_{i.}\mathbf{y}_{i.}'.$$

$$\mathrm{T} = \sum_{i=1}^{k}\sum_{j=1}^{n}(\mathbf{y}_{ij} - \bar{\mathbf{y}}_{..})(\mathbf{y}_{ij} - \bar{\mathbf{y}}_{..})'$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(\mathbf{y}_{ij} - \bar{\mathbf{y}}_{..})(\mathbf{y}_{ij} - \bar{\mathbf{y}}_{..})' = n\sum_{i=1}^{k}(\bar{\boldsymbol{y}}_{i} - \bar{\boldsymbol{y}}_{..})(\bar{\boldsymbol{y}}_{i} - \bar{\boldsymbol{y}}_{..})' + \sum_{i=1}^{k}\sum_{j=1}^{n}(\mathbf{y}_{ij} - \bar{\mathbf{y}}_{i.})(\mathbf{y}_{ij} - \bar{\mathbf{y}}_{i.})'$$

Toplam çarpımlar ve kareler toplamı	=	İşlem çarpımlar ve kareler toplamı	+	Artık çarpımlar ve kareler toplamı

y ler birer vektör

$$\mathbf{H} = \begin{pmatrix} \mathrm{SSH}_{11} & \mathrm{SPH}_{12} & \cdots & \mathrm{SPH}_{1p} \\ \mathrm{SPH}_{12} & \mathrm{SSH}_{22} & \cdots & \mathrm{SPH}_{2p} \\ \vdots & \vdots & & \vdots \\ \mathrm{SPH}_{1p} & \mathrm{SPH}_{2p} & \cdots & \mathrm{SSH}_{pp} \end{pmatrix},$$

$$\mathrm{SSH}_{22} = n \sum_{i=1}^{k} (\overline{y}_{i.2} - \overline{y}_{..2})^2 = \sum_i \frac{y_{i.2}^2}{n} - \frac{y_{..2}^2}{kn},$$

$$\mathrm{SPH}_{12} = n \sum_{i=1}^{k} (\overline{y}_{i.1} - \overline{y}_{..1})(\overline{y}_{i.2} - \overline{y}_{..2}) = \sum_i \frac{y_{i.1} y_{i.2}}{n} - \frac{y_{..1} y_{..2}}{kn}.$$

$$\overline{\mathbf{y}}_{i.} = \begin{pmatrix} \overline{y}_{i.1} \\ \overline{y}_{i.2} \\ \vdots \\ \overline{y}_{i.p} \end{pmatrix}$$

$$\mathbf{E} = \begin{pmatrix} \mathrm{SSE}_{11} & \mathrm{SPE}_{12} & \cdots & \mathrm{SPE}_{1p} \\ \mathrm{SPE}_{12} & \mathrm{SSE}_{22} & \cdots & \mathrm{SPE}_{2p} \\ \vdots & \vdots & & \vdots \\ \mathrm{SPE}_{1p} & \mathrm{SPE}_{2p} & \cdots & \mathrm{SSE}_{pp} \end{pmatrix}$$

$$\mathrm{SSE}_{22} = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij2} - \overline{y}_{i.2})^2 = \sum_{ij} y_{ij2}^2 - \sum_{i} \frac{y_{i.2}^2}{n},$$

$$\mathrm{SPE}_{12} = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij1} - \overline{y}_{i.1})(y_{ij2} - \overline{y}_{i.2}) = \sum_{ij} y_{ij1} y_{ij2} - \sum_{i} \frac{y_{i.1} y_{i.2}}{n}$$

$$\overline{\mathbf{y}}_{i.} = \begin{pmatrix} \overline{y}_{i.1} \\ \overline{y}_{i.2} \\ \vdots \\ \overline{y}_{i.p} \end{pmatrix}$$

Wilk's Lambda

- ANOVA daki F istatistiğine benzer olarak MANOVA da Wilk's istatistiği aşağıdaki gibi hesaplanır

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

- $\Lambda \leq \Lambda_{\alpha,p,\nu_H,\nu_E}$ ise yokluk hipotezi reddedilir.
- $\mathbf{E}^{-1}\mathbf{H}$ matrisinin özdeğerleri kullanılarak

$$\Lambda = \prod_{i=1}^{s} \frac{1}{1+\lambda_i}.$$

MANOVA Tablosu

Kareler Toplamı	Serbestlik Derecesi
$\mathrm{H} = n\sum_{i=1}^{k}(\bar{\boldsymbol{y}}_i - \bar{\boldsymbol{y}}_{..})(\bar{\boldsymbol{y}}_i - \bar{\boldsymbol{y}}_{..})'$	k-1
$\mathrm{E} = \sum_{\mathrm{i}=1}^{\mathrm{k}}\sum_{\mathrm{j}=1}^{\mathrm{n}}(\mathbf{y}_{\mathrm{ij}} - \bar{\mathbf{y}}_{\mathrm{i.}})(\mathbf{y}_{\mathrm{ij}} - \bar{\mathbf{y}}_{\mathrm{i.}})'$	nk-k
$T = \sum_{\mathrm{i}=1}^{\mathrm{k}}\sum_{\mathrm{j}=1}^{\mathrm{n}}(\mathbf{y}_{\mathrm{ij}} - \bar{\mathbf{y}}_{..})(\mathbf{y}_{\mathrm{ij}} - \bar{\mathbf{y}}_{..})'$	nk-1

Varsayımlar

- Gruplar bağımsız olmalı.
- Tüm gruplar aynı varyans-kovaryans matrisine sahip olmalı (Box’s M Testi).
- Her grup çok değişkenli normal olmalı.

Diğer İstatistikler

- Roy's largest root

$$\theta = \frac{\lambda_1}{1 + \lambda_1}$$

- Pillai's trace

$$V^{(s)} = \mathrm{tr}[(\mathbf{E} + \mathbf{H})^{-1}\mathbf{H}] = \sum_{i=1}^{s} \frac{\lambda_i}{1 + \lambda_i}.$$

- Lowley-Hotelling's trace

$$U^{(s)} = \mathrm{tr}(\mathbf{E}^{-1}\mathbf{H}) = \sum_{i=1}^{s} \lambda_i$$

İstatistiklerin Karşılaştırılması

- Ortalama vektörlerinin doğrusallığı $\mathbf{E}^{-1}\mathbf{H}$ matrisinin özdeğerleri ile ilgilidir.
- Bir özdeğer büyük, diğerleri küçükse ortalama vektörleri doğrusaldır. İki özdeğer büyükse iki boyutlu bir yapı sözkonusudur…
- Ortalama vektörleri doğrusalsa Roy testi daha güçlü Roy>Lawley>Wilks>Pillai
- Ortalama vektörleri doğrusal değil ya da varyans kovaryans matrisleri benzer değilse Pillai>Wilks>Lawley>Roy

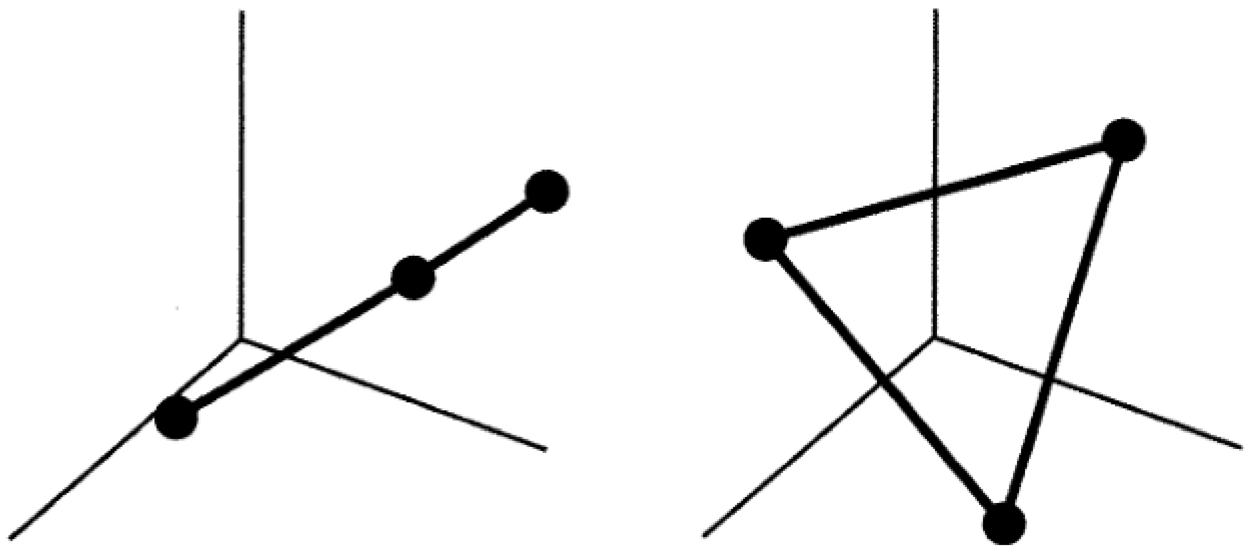

Figure 6.2 Two possible configurations for three mean vectors in 3-space.

- Grup büyüklükleri eşitse testler varyans-kovaryans matrislerinin eşitsizliğine karşı yeterince robust tır.
- Grup büyüklükleri farklı ve varyans-kovaryans matrisleri eşit değilse
- büyük gruplarda varyanslar büyük olduğunda testler tutucudur
- küçük gruplarda varyanslar büyük olduğunda testler daha liberaldir.
- Doğrusallık durumunda Roy testi tavsiye edilir.
- Varyanslar heterojen olma durumunda Pillai testi tavsiye edilir.

Örnek

$$\begin{pmatrix} \begin{bmatrix}9\\3\end{bmatrix} & \begin{bmatrix}0\\4\end{bmatrix} & \begin{bmatrix}3\\8\end{bmatrix} \\ \begin{bmatrix}6\\2\end{bmatrix} & \begin{bmatrix}2\\0\end{bmatrix} & \begin{bmatrix}1\\9\end{bmatrix} \\ \begin{bmatrix}9\\7\end{bmatrix} & & \begin{bmatrix}2\\7\end{bmatrix} \end{pmatrix}$$

$$\bar{y}_{1.} = \begin{bmatrix}8\\4\end{bmatrix} \qquad \bar{y}_{2.} = \begin{bmatrix}1\\2\end{bmatrix} \qquad \bar{y}_{3.} = \begin{bmatrix}2\\8\end{bmatrix}$$

$$\bar{y}_{..} = \begin{bmatrix}4\\5\end{bmatrix}$$

$$\begin{pmatrix} 9 & 0 & 3 \\ 6 & 2 & 1 \\ 9 & & 2 \end{pmatrix} = \begin{pmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & & 4 \end{pmatrix} + \begin{pmatrix} 4 & -3 & -2 \\ 4 & -3 & -2 \\ 4 & & -2 \end{pmatrix} + \begin{pmatrix} 1 & -1 & 1 \\ -2 & 1 & -1 \\ 1 & & 0 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 4 & 8 \\ 2 & 0 & 9 \\ 7 & & 7 \end{pmatrix} = \begin{pmatrix} 5 & 5 & 5 \\ 5 & 5 & 5 \\ 5 & & 5 \end{pmatrix} + \begin{pmatrix} -1 & -3 & 3 \\ -1 & -3 & 3 \\ -1 & & 3 \end{pmatrix} + \begin{pmatrix} -1 & 2 & 0 \\ -2 & -2 & 1 \\ 3 & & -1 \end{pmatrix}$$

$$\begin{bmatrix}9\\3\end{bmatrix} = \begin{bmatrix}4\\5\end{bmatrix} + \begin{bmatrix}4\\-1\end{bmatrix} + \begin{bmatrix}1\\-1\end{bmatrix}$$

(gözlem) = (ortalama) + (işlem) + (artık)

$$\begin{pmatrix} \begin{bmatrix}9\\3\end{bmatrix} & \begin{bmatrix}0\\4\end{bmatrix} & \begin{bmatrix}3\\8\end{bmatrix} \\ \begin{bmatrix}6\\2\end{bmatrix} & \begin{bmatrix}2\\0\end{bmatrix} & \begin{bmatrix}1\\9\end{bmatrix} \\ \begin{bmatrix}9\\7\end{bmatrix} & & \begin{bmatrix}2\\7\end{bmatrix} \end{pmatrix} \quad \bar{y}_{1.} = \begin{bmatrix}8\\4\end{bmatrix} \quad \bar{y}_{2.} = \begin{bmatrix}1\\2\end{bmatrix} \quad \bar{y}_{3.} = \begin{bmatrix}2\\8\end{bmatrix} \quad \bar{y}_{..} = \begin{bmatrix}4\\5\end{bmatrix}$$

$$\mathrm{SSH}_{22} = n\sum_{i=1}^{k}(\bar{y}_{i.2} - \bar{y}_{..2})^2 = \sum_i \frac{y_{i.2}^2}{n} - \frac{y_{..2}^2}{kn},$$

$$\mathrm{SPH}_{12} = n\sum_{i=1}^{k}(\bar{y}_{i.1} - \bar{y}_{..1})(\bar{y}_{i.2} - \bar{y}_{..2}) = \sum_i \frac{y_{i.1}y_{i.2}}{n} - \frac{y_{..1}y_{..2}}{kn}.$$

$$SSH_{11} = \sum_{\mathrm{i=1}}^{3} n_i(\bar{\mathrm{y}}_{\mathrm{i.1}} - \bar{\mathrm{y}}_{..1})^2 = [3(8-4)^2 + 2(1-4)^2 + 3(2-4)^2] = 78$$

$$SSH_{22} = \sum_{\mathrm{i=1}}^{\mathrm{n}} n_i(\bar{\mathrm{y}}_{\mathrm{i.2}} - \bar{\mathrm{y}}_{..2})^2 = 3(4-5)^2 + 2(2-5)^2 + 3(8-5)^2 = 48$$

$$SSH_{12} = \sum_{\mathrm{i=1}}^{\mathrm{n}} n_i(\bar{\mathrm{y}}_{\mathrm{i.1}} - \bar{\mathrm{y}}_{..1})(\bar{\mathrm{y}}_{\mathrm{i.2}} - \bar{\mathrm{y}}_{..2}) = 3(8-4)(4-5) + 2(1-4)(2-5) + 3(2-4)(8-5) = -12$$

$$\begin{pmatrix} \begin{bmatrix}9\\3\end{bmatrix} & \begin{bmatrix}0\\4\end{bmatrix} & \begin{bmatrix}3\\8\end{bmatrix} \\ \begin{bmatrix}6\\2\end{bmatrix} & \begin{bmatrix}2\\0\end{bmatrix} & \begin{bmatrix}1\\9\end{bmatrix} \\ \begin{bmatrix}9\\7\end{bmatrix} & & \begin{bmatrix}2\\7\end{bmatrix} \end{pmatrix} \quad \bar{y}_{1.} = \begin{bmatrix}8\\4\end{bmatrix} \quad \bar{y}_{2.} = \begin{bmatrix}1\\2\end{bmatrix} \quad \bar{y}_{3.} = \begin{bmatrix}2\\8\end{bmatrix} \qquad \bar{y}_{..} = \begin{bmatrix}4\\5\end{bmatrix}$$

$$\mathrm{SSE}_{22} = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij2} - \overline{y}_{i.2})^2 = \sum_{ij} y_{ij2}^2 - \sum_{i} \frac{y_{i.2}^2}{n},$$

$$\mathrm{SPE}_{12} = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij1} - \overline{y}_{i.1})(y_{ij2} - \overline{y}_{i.2}) = \sum_{ij} y_{ij1} y_{ij2} - \sum_{i} \frac{y_{i.1} y_{i.2}}{n}$$

$$SSE_{11} = (9-8)^2 + (6-8)^2 + (9-8)^2 + (0-1)^2 + (2-1)^2 + (3-2)^2 + (1-2)^2 + (2-2)^2 = 10$$

$$SSE_{22} = (3-4)^2 + (2-4)^2 + (7-4)^2 + (4-2)^2 + (0-2)^2 + (8-8)^2 + (9-8)^2 + (7-8)^2 = 24$$

$$SSE_{12} = (9-8)(3-4) + (6-8)(2-4) + (9-8)(7-4) + \cdots + (2-2)(7-8) = 1$$

MANOVA Tablosu

	Çarpımlar ve Kareler Toplamı	Serbestlik Derecesi
İşlem Between	$\begin{bmatrix} 78 & -12 \\ -12 & 48 \end{bmatrix}$	$3-1=2$
Artık Within	$\begin{bmatrix} 10 & 1 \\ 1 & 24 \end{bmatrix}$	$3+2+3-3=5$
Toplam	$\begin{bmatrix} 88 & -11 \\ -11 & 72 \end{bmatrix}$	$8-1=7$

$$\Lambda = \frac{|E|}{|E+H|} = \frac{\begin{vmatrix} 10 & 1 \\ 1 & 24 \end{vmatrix}}{\begin{vmatrix} 88 & -11 \\ -11 & 72 \end{vmatrix}} = \frac{239}{6215} = 0.0385$$

Çok Değişkenli İlişki Ölçüsü

- MANOVA da etki büyüklüğü ya da çok değişkenli ilişkinin ölçüsü olarak

$$\text{MANOVA } \eta^2 = \eta_\Lambda^2 = 1 - \Lambda,$$

- değeri kullanılır.

Kontrastlar

- Tek değişkenli kontrastlar aşağıdaki gibidir

$$\delta = c_1\mu_1 + c_2\mu_2 + \cdots + c_k\mu_k,$$

$$\sum_{i=1}^{k} c_i = 0.$$

- Çok değişkenli kontrastlar da benzerdir.

$$\boldsymbol{\delta} = c_1\boldsymbol{\mu}_1 + c_2\boldsymbol{\mu}_2 + \cdots + c_k\boldsymbol{\mu}_k,$$

$$\sum_{i=1}^{k} c_i = 0.$$