ÇOK DEĞİŞKENLİ İKİ FAKTÖR VARYANS ANALİZİ

P: Bağımlı değişken sayısı $(P \geq 2)$ olmak üzere, bu değişkenlerin iki bağımsız değişken (faktör) tarafından açıklanmaya çalışıldığını (etkilendiğini) kabul edelim. Bağımlı değişkenler eşit aralıklı veya oranlama düzeyinde ölçülüyorken, faktörler sınıflama ya da sıralama düzeyinde ölçülüyorsa, söz konusu faktörler ile bağımlı değişkenleri açıklamada kullanılan istatistiksel analize çok değişkenli iki faktör varyans analizi (MANOVA) adı verilir.

Kabul edelim ki I. faktörün (F_1) g-tane ve II. faktörün (F_2) b-tane düzeyi olsun. Bu takdirde deneme kombinasyonlarının (bağımsız grupların) sayısı "$g \times b$" tane olup, her bir deneme kombinasyonu bir çok değişkenli kitle ile temsil edilebilir.

Örneğin; (k,l) deneme kombinasyonu için değişkenler vektörü $\underline{X}_{kl}$ ile gösterilirse, bu deneme kombinasyonuna karşılık gelen çok değişkenli kitle

$$\underline{X}_{kl} = \begin{bmatrix} x_{1kl} \\ x_{2kl} \\ \vdots \\ x_{pkl} \end{bmatrix}, \quad \begin{matrix} k = \overline{1,g} \\ l = \overline{1,b} \end{matrix}$$

değişkenler vektörü ile gösterilir.

Her bir deneme kombinasyonunda n-tane denemenin yapıldığı ya da her bir çok değişkenli kitleden rastgele ve birbirinden bağımsız olarak n-birimlik örnekler çekildiğinde örnek veri düzeni aşağıdaki gibi olur.

I. FAKTÖR

II. FAKTÖR		1	2	...	g	$\underline{\mu}_{.\ell}$
1	x_1 x_2 $\vdots$ x_p	$\underline{X}_{111}$ $\underline{X}_{112} \dots \underline{X}_{11n}$ $\underline{X}_{11}(\underline{\mu}_{11})$	$\underline{X}_{21}(\underline{\mu}_{21})$	...	$\underline{X}_{g1}(\underline{\mu}_{g1})$	$\underline{\mu}_{.1}$
2	x_1 x_2 $\vdots$ x_p	$\underline{X}_{12}(\underline{\mu}_{12})$	$\underline{X}_{22}(\underline{\mu}_{22})$	...	$\underline{X}_{g2}(\underline{\mu}_{g2})$	$\underline{\mu}_{.2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
b	x_1 x_2 $\vdots$ x_p	$\underline{X}_{1b}(\underline{\mu}_{1b})$	$\underline{X}_{2b}(\underline{\mu}_{2b})$	...	$\underline{X}_{gb}(\underline{\mu}_{gb})$	$\underline{\mu}_{.b}$
$\underline{\mu}_{k.}$		$\underline{\mu}_{1.}$	$\underline{\mu}_{2.}$	...	$\underline{\mu}_{g.}$	$\underline{\mu}_{..}$

Kabul edelim ki bağımsız grupları temsil etmek için kullanılan değişkenler vektörünün dağılımı;

$\underline{X}_{k\ell} \sim N_p(\underline{\mu}_{k\ell}, \Sigma_{k\ell})$ ve gruplar homojen kovaryans matrisli olsun. Yani;

$$\Sigma_{11} = \Sigma_{21} = \dots = \Sigma_{g1} = \Sigma_{21} = \dots = \Sigma_{gb} = \Sigma$$

Burada; $\underline{\mu}_{k\ell} = \begin{bmatrix} \mu_{1k\ell} \\ \mu_{2k\ell} \\ \vdots \\ \mu_{pk\ell} \end{bmatrix}$ ve $\Sigma = [\sigma_{jt}]_{p\times p}$, $j, t = \overline{1, p}$ şeklindedir.

Bu deney planına uygun model denklemi ise;

$$\underline{X}_{kli} = \underline{\mu}_{..} + \underline{\alpha}_k + \underline{\beta}_l + (\underline{\alpha\beta})_{kl} + \underline{\varepsilon}_{kli} \quad , \quad k=\overline{1,g} \quad \ldots (1)$$
$$l=\overline{1,b}$$
$$i=\overline{1,n}$$

olup, bu denkleme çok değişkenli 2 faktör varyans analizinin matematiksel modeli denir. Burada;

$\underline{X}_{kli}$: (kl) deneme kombinasyonunda i-inci gözlem vektörü (örnek birimi)

$$\underline{X}_{kli} = \begin{bmatrix} x_{1kli} \\ x_{2kli} \\ \vdots \\ x_{pkli} \end{bmatrix}$$

$\underline{\mu}_{..}$: Genel kitle ortalama vektörü

$\underline{\alpha}_k$: I-inci faktörün k-ıncı düzeyinin bağımlı değişkenler üzerindeki etkisi

$$\underline{\alpha}_k = \underline{\mu}_{k.} - \underline{\mu}_{..} \quad , \quad \sum_{k=1}^{g} \underline{\alpha}_k = \underline{0}$$

$$\underline{\beta}_l = \underline{\mu}_{.l} - \underline{\mu}_{..} \quad , \quad \sum_{l=1}^{b} \underline{\beta}_l = \underline{0}$$

$(\underline{\alpha\beta})_{kl}$: (kl) deneme kombinasyonunun etkisi (etkileşim)

$$(\underline{\alpha\beta})_{kl} = \underline{\mu}_{kl} - \underline{\mu}_{k.} - \underline{\mu}_{.l} + \underline{\mu}_{..}$$

$\underline{\varepsilon}_{kli}$: Hata terimi, $\underline{\varepsilon}_{kli} \sim N_p(\underline{0}, \Sigma)$ ve bağımsız.

PARAMETRE TAHMİNİ

Eşitlik (1) ile verilen modele ait parametrelerin tahmin edilmesinde En Küçük Kareler veya En çok olabilirlik tahmin yöntemleri kullanılabilir. Burada en küçük kareler yöntemi ile tahmin edicilerin elde edilmesi verilecektir. Çünkü her iki yöntemle aynı tahmin ediciler elde edilmektedir.

En küçük kareler yönteminde hata kareler toplamı en küçük yapılacak şekilde parametreler tahmin edilmektedir. (1) modeli ile ilgili HKT'mı;

$$HKT = \sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n} \underline{\varepsilon}_{kli}\,\underline{\varepsilon}_{kli}'$$

$$= \sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n} (\underline{x}_{kli} - \underline{\mu}_{..} - \underline{\alpha}_k - \underline{\beta}_l - (\underline{\alpha\beta})_{kl})(\underline{x}_{kli} - \underline{\mu}_{..} - \underline{\alpha}_k - \underline{\beta}_l - (\underline{\alpha\beta})_{kl})'$$

$\Rightarrow$ $\underline{\mu}_{..}$ parametresinin tahmini;

$$\left.\frac{\partial HKT}{\partial \underline{\mu}_{..}}\right|_{\underline{\mu}_{..}=\hat{\underline{\mu}}_{..}} = -2\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n} (\underline{x}_{kli} - \hat{\underline{\mu}}_{..} - \underline{\alpha}_k - \underline{\beta}_l - (\underline{\alpha\beta})_{kl}) = \underline{0} \Rightarrow$$

$$\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n} \underline{x}_{kli} - g\times b\times n\times\hat{\underline{\mu}}_{..} - b\times n \underbrace{\cancel{\sum_{k=1}^{g}\underline{\alpha}_k}}_{(0)} - g\times n \underbrace{\cancel{\sum_{l=1}^{b}\underline{\beta}_l}}_{(0)} - n \underbrace{\cancel{\sum_{k=1}^{g}\sum_{l=1}^{b}(\underline{\alpha\beta})_{kl}}}_{(0)}$$

$$\Rightarrow g\times b\times n\times\hat{\underline{\mu}}_{..} = \sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n} \underline{x}_{kli} \Rightarrow N = g\times b\times n \text{ olduğundan}$$

$$\hat{\underline{\mu}}_{..} = \frac{1}{N}\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n} \underline{x}_{kli} = \bar{\underline{x}}_{..}$$ Genel örnek ortalama vektörü olarak bulunu

$\underline{\mu_{k.},\ (k=1,2,\dots,g)}$ parametresinin tahmini $(\alpha_k = \mu_{k.} - \mu_{..})$

$$\left.\frac{\partial HKT}{\partial \mu_{k.}}\right|_{\mu_{k.}=\hat{\mu}_{k.}} = \frac{\partial}{\partial \mu_{k.}}\left[\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli}-\mu_{..}-\mu_{k.}+\mu_{..}-\beta_l-(\alpha\beta)_{kl})\cdot(\underline{x}_{kli}-\mu_{..}-\mu_{k.}+\mu_{..}-\beta_l-(\alpha\beta)_{kl})'\right]_{\mu_{k.}=\hat{\mu}_{k.}}$$

$$=\frac{\partial}{\partial \mu_{k.}}\left[\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli}-\mu_{k.}-\beta_l-(\alpha\beta)_{kl})(\underline{x}_{kli}-\mu_{k.}-\beta_l-(\alpha\beta)_{kl})'\right]_{\mu_{k.}=\hat{\mu}_{k.}}$$

$$=-2\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli}-\hat{\mu}_{k.}-\beta_l-(\alpha\beta)_{kl})=\underline{0}\Rightarrow$$

$$\sum_{l=1}^{b}\sum_{i=1}^{n}\underline{x}_{kli}-b\times n\,\hat{\mu}_{k}-n\underbrace{\sum_{l=1}^{b}\beta_l}_{(0)}-n\underbrace{\sum_{l=1}^{b}(\alpha\beta)_{kl}}_{(0)}=\underline{0}$$

$$\Rightarrow \hat{\mu}_{k.}=\frac{1}{bn}\sum_{l=1}^{b}\sum_{i=1}^{n}\underline{x}_{kli}=\underline{\bar{x}}_{k.},\ k=1,2,\dots,g \rightarrow$$ I. faktörün k.ncı düzeyine ait örnek ortalama vektörüdür.

$\underline{\alpha_k=\mu_{k.}-\mu_{..},\ (k=1,2,\dots,g)}$ parametresinin tahmini

$$\left.\frac{\partial HKT}{\partial \alpha_k}\right|_{\alpha_k=\hat{\alpha}_k} = \frac{\partial}{\partial \alpha_k}\left[\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli}-\mu_{..}-\alpha_k-\beta_l-(\alpha\beta)_{kl})\cdot(\underline{x}_{kli}-\mu_{..}-\alpha_k-\beta_l-(\alpha\beta)_{kl})'\right]_{\substack{\alpha_k=\hat{\alpha}_k\\ \mu_{..}=\hat{\mu}_{..}}}$$

$$=\frac{\partial}{\partial \alpha_k}\left[\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli}-\mu_{..}-\alpha_k-\beta_l-(\alpha\beta)_{kl})(\underline{x}_{kli}-\mu_{..}-\alpha_k-\beta_l-(\alpha\beta)_{kl})'\right]=\underline{0}\Rightarrow$$

$$=-2\left[\sum_{l=1}^{b}\sum_{i=1}^{n}\underline{x}_{kli}-b\times n\times\hat{\mu}_{..}-b\times n\,\hat{\alpha}_k-n\underbrace{\sum_{l=1}^{b}\beta_l}_{(0)}-n\underbrace{\sum_{l=1}^{b}(\alpha\beta)_{kl}}_{(0)}\right]=\underline{0}\Rightarrow$$

$$b\times n\times\hat{\alpha}_k=\sum_{l=1}^{b}\sum_{i=1}^{n}\underline{x}_{kli}-b\times n\times\hat{\mu}_{..},\quad (\hat{\mu}_{..}=\underline{\bar{x}}_{..}\ \text{old.})$$

$$\hat{\alpha}_k=\frac{1}{b\times n}\sum_{l=1}^{b}\sum_{i=1}^{n}\underline{x}_{kli}-\underline{\bar{x}}_{..}\Rightarrow\boxed{\hat{\alpha}_k=\underline{\bar{x}}_{k.}-\underline{\bar{x}}_{..}},\ k=1,2,\dots,g.$$

$\mu_{.l}$, $(l=1,2,\ldots,b)$ parametresinin tahmini $(\beta_l = \mu_{.l} - \mu_{..})$

$$\cancel{\mu_{..}} - \alpha_k - \mu_{.l} + \cancel{\mu_{..}} - (\alpha\beta)_{kl}.$$

$$\left.\frac{\partial HKT}{\partial \mu_{.l}}\right|_{\mu_{.l}=\hat{\mu}_{.l}} = \frac{\partial}{\partial \mu_{.l}}\left[\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli} - \cancel{\mu_{..}} - \alpha_k - \mu_{.l} + \cancel{\mu_{..}} - (\alpha\beta)_{kl})' \right.$$

$$\left.(\underline{x}_{kli} - \cancel{\mu_{..}} - \alpha_k - \mu_{.l} + \cancel{\mu_{..}} - (\alpha\beta)_{kl})\right]$$

$$= \frac{\partial}{\partial \mu_{.l}}\left\{\sum_{k=1}^{g}\sum_{i=1}^{n}(\underline{x}_{kli} - \alpha_k - \mu_{.l} - (\alpha\beta)_{kl})'(\underline{x}_{kli} - \alpha_k - \mu_{.l} - (\alpha\beta)_{kl})\right]$$

$$= -2\left[\sum_{k=1}^{g}\sum_{i=1}^{n}\underline{x}_{kli} - n\underbrace{\sum_{k=1}^{g}\underline{\alpha}_k}_{(0)} - g\times n\hat{\mu}_{.l} - n\underbrace{\cancel{\sum_{k=1}^{g}(\alpha\beta)_{kl}}}_{(0)}\right] = 0$$

$$\Rightarrow gn\hat{\mu}_{.l} = \sum_{k=1}^{g}\sum_{i=1}^{n}\underline{x}_{kli} \Rightarrow \hat{\mu}_{.l} = \frac{1}{g\times n}\sum_{k=1}^{g}\sum_{i=1}^{n}\underline{x}_{kli} = \bar{\underline{x}}_{.l}, \quad l=1,2,\ldots,b$$

↓

II. Faktörün l-nci düzeyine ait örnek ortalama vektörüdür.

β_l, $(l=1,2,\ldots,b)$ parametresinin tahmini

$$\mu_{..} - \alpha_k - \beta_l - (\alpha\beta)_{kl}.$$

$$\left.\frac{\partial HKT}{\partial \beta_l}\right|_{\beta_l=\hat{\beta}_l} = \frac{\partial}{\partial \beta_l}\left[\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli} - \mu_{..} - \alpha_k - \beta_l - (\alpha\beta)_{kl})' \right.$$

$$\left.(\underline{x}_{kli} - \mu_{..} - \alpha_k - \beta_l - (\alpha\beta)_{kl})\right]$$

$$= \frac{\partial}{\partial \beta_l}\left[\sum_{k=1}^{g}\sum_{i=1}^{n}(\underline{x}_{kli} - \mu_{..} - \alpha_k - \beta_l - (\alpha\beta)_{kl})(\underline{x}_{kli} - \mu_{..} - \alpha_k - \beta_l - (\alpha\beta)_{kl})'\right]$$

$$= -2\left[\sum_{k=1}^{g}\sum_{i=1}^{n}\underline{x}_{kli} - g\times n\,\hat{\mu}_{..} - n\underbrace{\cancel{\sum_{k=1}^{g}\alpha_k}}_{(0)} - g\times n\,\hat{\beta}_l - n\underbrace{\cancel{\sum_{k=1}^{g}(\alpha\beta)_{kl}}}_{(0)}\right] = 0$$

$$\Rightarrow \hat{\beta}_l = \frac{1}{g\times n}\sum_{k=1}^{g}\sum_{i=1}^{n}\underline{x}_{kli} - \hat{\mu}_{..}, \quad (\hat{\mu}_{..} = \bar{\underline{x}}_{..} \text{ olduğundan})$$

$$\boxed{\hat{\beta}_l = \bar{\underline{x}}_{.l} - \bar{\underline{x}}_{..}}, \quad l=1,2,\ldots,b$$

$\underline{\mu}_{kl}$, $(k=1,2,\dots,g;\ l=1,2,\dots,b)$ Tahmini

$$\left.\frac{\partial HKT}{\partial \underline{\mu}_{kl}}\right|_{\underline{\mu}_{kl}=\hat{\underline{\mu}}_{kl}} = \frac{\partial}{\partial \underline{\mu}_{kl}}\left[\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli}-\underline{\mu}_{..}-\underline{\alpha}_k-\underline{\beta}_l-(\underline{\alpha\beta})_{kl})\cdot(\underline{x}_{kli}-\underline{\mu}_{..}-\underline{\alpha}_k-\underline{\beta}_l-(\underline{\alpha\beta})_{kl})'\right]=\underline{0} \Rightarrow$$

$\underline{\alpha}_k=\underline{\mu}_{k.}-\underline{\mu}_{..}$, $\underline{\beta}_l=\underline{\mu}_{.l}-\underline{\mu}_{..}$ ve $(\underline{\alpha\beta})_{kl}=\underline{\mu}_{kl}-\underline{\mu}_{k.}-\underline{\mu}_{.l}+\underline{\mu}_{..}$

olduğu dikkate alınırsa;

$$\frac{\partial}{\partial \underline{\mu}_{kl}}\left[\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli}-\cancel{\underline{\mu}_{..}}-\cancel{\underline{\mu}_{k.}}+\cancel{\underline{\mu}_{..}}-\cancel{\underline{\mu}_{.l}}+\cancel{\underline{\mu}_{..}}-\underline{\mu}_{kl}+\cancel{\underline{\mu}_{k.}}+\cancel{\underline{\mu}_{.l}}-\cancel{\underline{\mu}_{..}})(\underline{x}_{kli}-\cancel{\underline{\mu}_{..}}-\cancel{\underline{\mu}_{k.}}+\cancel{\underline{\mu}_{..}}-\cancel{\underline{\mu}_{.l}}+\cancel{\underline{\mu}_{..}}-\underline{\mu}_{kl}+\cancel{\underline{\mu}_{k.}}+\cancel{\underline{\mu}_{.l}}-\cancel{\underline{\mu}_{..}})'\right]=\underline{0}$$

$$=\frac{\partial}{\partial \underline{\mu}_{kl}}\left[\sum_{i=1}^{n}(\underline{x}_{kli}-\underline{\mu}_{kl})(\underline{x}_{kli}-\underline{\mu}_{kl})'\right]_{\underline{\mu}_{kl}=\hat{\underline{\mu}}_{kl}}$$

$$-2\left[\sum_{i=1}^{n}\underline{x}_{kli}-n\hat{\underline{\mu}}_{kl}\right]=\underline{0} \Rightarrow \boxed{\hat{\underline{\mu}}_{kl}=\frac{1}{n}\sum_{i=1}^{n}\underline{x}_{kli}=\bar{\underline{x}}_{kl.}}$$

(k,l) deneme kombinasyonuna ait
örnek ortalama vektörü.

$(\underline{\alpha\beta})_{kl}=(\underline{\gamma}_{kl})$, $(k=1,2,\dots,g;\ l=1,2,\dots,b)$ Tahmini

$$\left.\frac{\partial HKT}{\partial (\underline{\alpha\beta})_{kl}}\right|_{(\underline{\alpha\beta})_{kl}=(\hat{\underline{\alpha\beta}})_{kl}} = \frac{\partial}{\partial (\underline{\alpha\beta})_{kl}}\left[\sum_{k=1}^{g}\sum_{l=1}^{b}\sum_{i=1}^{n}(\underline{x}_{kli}-\underline{\mu}-\underline{\alpha}_k-\underline{\beta}_l-(\underline{\alpha\beta})_{kl})\cdot(\underline{x}_{kli}-\underline{\mu}_{..}-\underline{\alpha}_k-\underline{\beta}_l-(\underline{\alpha\beta})_{kl})'\right]$$

$$=\frac{\partial}{\partial (\underline{\alpha\beta})_{kl}}\left[\sum_{i=1}^{n}(\underline{x}_{kli}-\underline{\mu}_{..}-\underline{\alpha}_k-\underline{\beta}_l-(\underline{\alpha\beta})_{kl})(\underline{x}_{kli}-\underline{\mu}_{..}-\underline{\alpha}_k-\underline{\beta}_l-(\underline{\alpha\beta})_{kl})'\right]=\underline{0} \Rightarrow$$

$$=-2\left[\sum_{i=1}^{n}\underline{x}_{kli}-n\hat{\underline{\mu}}_{..}-n\hat{\underline{\alpha}}_k-n\hat{\underline{\beta}}_l-n(\hat{\underline{\alpha\beta}})_{kl}\right]$$

$$n(\hat{\underline{\alpha\beta}})_{kl}=\sum_{i=1}^{n}\underline{x}_{kli}-n\hat{\underline{\mu}}_{..}-n(\bar{\underline{x}}_{k.}-\bar{\underline{x}}_{..})-n(\bar{\underline{x}}_{.l}-\bar{\underline{x}}_{..})$$

$$\Rightarrow (\hat{\alpha\beta})_{kl} = \frac{1}{n}\sum_{i=1}^{n} x_{kli} - \cancel{\bar{x}_{..}} - \bar{x}_{k.} + \cancel{\bar{x}_{..}} - \bar{x}_{.l} + \bar{x}_{..}$$

$$\boxed{(\hat{\alpha\beta})_{kl} = \bar{x}_{kl} - \bar{x}_{k.} - \bar{x}_{.l} + \bar{x}_{..} \,, \quad \begin{array}{l} k=1,2,\dots,8 \\ l=1,2,\dots,8 \end{array}}$$

bulunur.

Hata teriminin en küçük kareler kestirimi ise;

$$\varepsilon_{kli} = x_{kli} - \mu_{..} - \alpha_k - \beta_l - (\alpha\beta)_{kl}$$

$$= x_{kli} - \cancel{\mu_{..}} - \cancel{\mu_{k.}} + \cancel{\mu_{..}} - \cancel{\mu_{.l}} + \cancel{\mu_{..}} - \mu_{kl} + \cancel{\mu_{k.}} + \cancel{\mu_{.l}} - \cancel{\mu_{..}}$$

$$= x_{kli} - \mu_{kl}$$

olması nedeniyle;

$$\hat{\varepsilon}_{kli} = x_{kli} - \hat{\mu}_{kl} = x_{kli} - \bar{x}_{kl} \,, \quad \begin{array}{l} k=1,2,\dots,8 \\ l=1,2,\dots,b \\ i=1,2,\dots,n \end{array}$$

olarak elde edilir.