AMAÇ: HİPOTEZ TESTİ

$g \times b$ - tane bağımsız alt kitleyi (grubu), ortalama vektörleri ($\underline{\mu}_{kl}$) yönünden karşılaştırmak ve faktörlerin bağımlı değişkenler üzerindeki etkisini incelemektir.

Bu amaçla test edilecek olan hipotezler şu şekilde oluşturulur.

i) I-inci faktörün etkisinin önemliliği için;

$$H_0: \underline{\mu}_{1.} = \underline{\mu}_{2.} = \dots = \underline{\mu}_{g.}$$
$$H_1: \exists\, \underline{\mu}_{k.} \text{ diğerlerinden farklıdır.} \qquad \dots (2)$$

ii) II-inci faktörün etkisinin önemliliği için;

$$H_0: \underline{\mu}_{.1} = \underline{\mu}_{.2} = \dots = \underline{\mu}_{.b}$$
$$H_1: \exists\, \underline{\mu}_{.l} \text{ diğerlerinden farklıdır} \qquad \dots (3)$$

iii) Etkileşimin önemliliği için;

$$H_0: \underline{\mu}_{11} = \underline{\mu}_{21} = \dots = \underline{\mu}_{g1} = \underline{\mu}_{12} = \dots = \underline{\mu}_{gb} \quad \dots (*)$$
$$H_1: \exists\, \underline{\mu}_{kl} \text{ diğerlerinden farklıdır} \qquad \dots (4)$$

(*): Etkileşimin önemsiz olduğu anlamına gelir. Eğer etkileşim önemsiz ise I. ve II. faktörler, bağımlı değişkenleri birbirinden bağımsız olarak etkiler.

Çok değişkenli 2 faktör varyans analizi için model denkleminde yer alan parametrelerin tahmin edicilerini bulalım.

i) $\underline{\mu}_{kl}$'nin tahmin edicisi;

$$\underline{\bar{X}}_{kl} = \frac{1}{n}\sum_{i=1}^{n} \underline{X}_{kli} \quad , \quad k=\overline{1,g} \quad , \quad l=\overline{1,b}$$

olup, bu ortalamalardan $g \times b$ tane vardır.

4

ii) $\mu_{k.}$ parametresinin tahmin edicisi;

$$\bar{X}_{k.} = \frac{1}{b.n} \sum_{l=1}^{b} \sum_{i=1}^{n} X_{kli} = \frac{1}{b} \sum_{l=1}^{b} \bar{X}_{kl} \quad , k=\overline{1,g} \quad \text{olup, } g\text{-tanedir}$$

$\mu_{.l}$ parametresinin tahmin edicisi;

$$\bar{X}_{.l} = \frac{1}{gn} \sum_{k=1}^{g} \sum_{i=1}^{n} X_{kli} = \frac{1}{g} \sum_{k=1}^{g} X_{kl} \quad , l=\overline{1,b} \quad \text{olup, b tanedir}$$

$\mu_{..}$ parametresinin tahmin edicisi;

$$\bar{X}_{..} = \frac{1}{g.b.n} \sum_{k=1}^{g} \sum_{l=1}^{b} \sum_{i=1}^{n} X_{kli} = \frac{1}{g.b} \sum_{k=1}^{g} \sum_{l=1}^{b} \bar{X}_{kl}$$

$$= \frac{1}{g} \sum_{k=1}^{b} \bar{X}_{k.}$$

$$= \frac{1}{b} \sum_{l=1}^{b} \bar{X}_{.l}$$

H_0- hipotezlerinin test edilmesinde kullanılacak olan test istatistikleri çok değişkenli 2 faktör varyans analizinin genel denklemi olacak bilinen,

$$T = F_1 + F_2 + F_{12} + H \quad \ldots \quad (5)$$

F_1: I-inci faktöre ait değişim $\quad F_{12}$: etkileşim

F_2: II-inci faktöre ait değişim $\quad H$: hata

T : Toplam değişim

matris denkleminden yararlanılır.

Burada;

T: genel kareler ve çarpımlar toplamı matrisi (pxp)

$$T = \sum_{k=1}^{g} \sum_{l=1}^{b} \sum_{i=1}^{n} (\underline{X}_{kli} - \underline{\overline{X}}_{..})(\underline{X}_{kli} - \underline{\overline{X}}_{..})' \quad \ldots (6)$$

$\underline{\overline{X}}_{..}$: genel ortalama

F_1: I-inci faktör için kareler ve çarpımlar toplamı matrisi

$$F_1 = \sum_{k=1}^{g} bn (\underline{\overline{X}}_{k.} - \underline{\overline{X}}_{..})(\underline{\overline{X}}_{k.} - \underline{\overline{X}}_{..})' \quad \ldots (7)$$

F_2 : II-inci faktör için kareler ve çarpımlar toplamı matrisi

$$F_2 = \sum_{l=1}^{b} gn (\underline{\overline{X}}_{.l} - \underline{\overline{X}}_{..})(\underline{\overline{X}}_{.l} - \underline{\overline{X}}_{..})' \quad \ldots (8)$$

$$F_{12} = \sum_{k=1}^{g} \sum_{l=1}^{b} n (\underline{\overline{X}}_{kl} - \underline{\overline{X}}_{k.} - \underline{\overline{X}}_{.l} + \underline{\overline{X}}_{..})(\underline{\overline{X}}_{kl} - \underline{\overline{X}}_{k.} - \underline{\overline{X}}_{.l} + \underline{\overline{X}}_{..})' \ldots (9)$$

H: hata kareler ve çarpımlar toplamı

$$H = \sum_{k=1}^{g} \sum_{l=1}^{b} \sum_{i=1}^{n} (\underline{X}_{kli} - \underline{\overline{X}}_{kl})(X_{kli} - \underline{\overline{X}}_{kl})' \quad \ldots (10)$$

Test istatistikleri: Wilks'in olabilirlik oran istatistiği (Λ) kullanılarak, tanımlanan hipotezlere göre kurulabilir.

(2) numaralı Ho hipotezi için; test istatistiğini kullanacak olursak,

Ho: $\underline{\mu}_{1.} = \underline{\mu}_{2.} = \ldots = \underline{\mu}_{g.}$

$$\Lambda_1 = \frac{|H|}{|F_1 + H|} , \quad 0 < \Lambda_1 < 1$$

olup, $\Lambda_1 \longrightarrow 0$ iken Ho red edilir.

Örnek hacmi yeterince büyükken Λ_1 istatistiğinin örnekleme dağılımı için:

$$L_1 = -\left[gb(n-1) - \frac{(p+1)-(g-1)}{2} \right] \ln \Lambda_1 \sim \chi^2_{p(g-1)} ,$$

6

Eğer, α önem seviyesinde ; $L_1 > \chi^2_{p(g-1);\alpha} \Rightarrow H_0$ red edilir

$L_1 \leq \chi^2_{p(g-1);\alpha} \Rightarrow H_0$ kabul edilir
(Etkiler önemsiz)

(3) numaralı H_0 hipotezi için; test istatistiğini kullanacak olursak ;

$H_0: \mu_{.1} = \mu_{.2} = \dots = \mu_{.b}$

$$\Lambda_2 = \frac{|H|}{|F_2 + H|} \quad , \quad 0 < \Lambda_2 < 1$$

olup, $\Lambda_2 \longrightarrow 0$ iken H_0 red edilir.

Λ_2 - istatistiğinin örnekleme dağılımı için;

$$L_2 = -\left[gb(n-1) - \frac{p+1-(b-1)}{2} \right] \ln \Lambda_2 \sim \chi^2_{p(b-1)}$$

Eğer α - önem seviyesinde ; $L_2 > \chi^2_{p(b-1);\alpha} \Rightarrow H_0$ red edilir.

$L_2 \leq \chi^2_{p(b-1);\alpha} \Rightarrow H_0$ kabul edilir

H_0 kabul edilirse II - inci faktörün bağımlı değişken üzerine etkisi önemsizdir.

(4) numaralı H_0 hipotezi için; test istatistiğini kullanacak olursak ,

$H_0: \mu_{11} = \mu_{12} = \dots = \mu_{gb}$

$$\Lambda_E = \frac{|H|}{|E_{12} + H|} \quad , \quad 0 < \Lambda_E < 1$$

olup , $\Lambda_E \longrightarrow 0$ iken H_0 red edilir.

Λ_E - istatistiğinin örnekleme dağılımı için;

$$L_E = -\left[gb(n-1) - \frac{p+1-(g-1)(b-1)}{2} \right] \ln \Lambda_E \sim \chi^2_{p(g-1)(b-1)}$$

Eğer α - önem seviyesinde , $L_E > \chi^2_{p(g-1)(b-1);\alpha} \Rightarrow H_0$ red edilir

$L_E \leq \chi^2_{p(g-1)(b-1);\alpha} \Rightarrow H_0$ kabul edilir
(Etkileşim önemsiz)

WILO

Ho kabul edilirse etkileşimin önemsiz olduğu ifade edilir.

Etkileşimin önemsiz olması; I-inci faktör ile II-inci faktörün, bağımlı değişkenler vektörünü birbirinden bağımsız olarak etkilediği anlamına gelir.

Kabul edelim ki (2) ve (3) nolu Ho hipotezleri red edilmiş olsun. Bu durumda farklılığın hangi gruplar arasında ve hangi değişken bakımından ortaya çıktığı araştırılabilir.

Bu amaçla Bonferroni güven aralıklarından yararlanılır.

* I-inci faktörün k-ıncı ve t-inci düzeylerini ($k \neq t$) j-inci değişken bakımından karşılaştıralım:

$$H_0: \mu_{jk.} - \mu_{jt.} = 0 \qquad , \quad k \neq t = \overline{1,g} \quad , \quad j = \overline{1,p}$$
$$H_1: \mu_{jk.} - \mu_{jt.} \neq 0$$

mümkün olan ikili karşılaştırmaların sayısı; $p\binom{g}{2}$'dir.

$$P\left[(\bar{X}_{jk.} - \bar{X}_{jt.}) - t_{gb(n-1);\frac{\alpha}{p \cdot g(g-1)}} \sqrt{\frac{h_{jj}}{gb(n-1)} \cdot \frac{2}{bn}} \leq \mu_{jk.} - \mu_{jt.} \leq (\bar{X}_{jk.} - \bar{X}_{jt.}) + \cdots\right] = 1-\alpha$$

* II-inci faktörün l-inci ve m-inci düzeylerini ($l \neq m$) j-inci değişken bakımından karşılaştıralım:

$$H_0: \mu_{j.l} - \mu_{j.m} = 0 \qquad , \quad l \neq m = \overline{1,b} \quad , \quad j = \overline{1,p}$$
$$H_1: \mu_{j.l} - \mu_{j.m} \neq 0$$

mümkün olan ikili karşılaştırmaların sayısı $p\binom{b}{2}$'dir.

$$P\left[(\bar{X}_{j.l} - \bar{X}_{j.m}) - t_{gb(n-1);\frac{\alpha}{pb(b-1)}} \sqrt{\frac{h_{jj}}{gb(n-1)} \frac{2}{gn}} \leq \mu_{j.l} - \mu_{j.m} \leq (\bar{X}_{j.l} - \bar{X}_{j.m}) + \cdots\right] = 1-\alpha$$

→ h_{jj}: H-matrisinin j-inci köşegen elemanı