

T.C.
GEBZE YKSEK TEKNOLOJİ ENSTİTS
MHENDİSLİK VE FEN BİLİMLERİ ENSTİTS

BİLEŞEN YAKLAŞIMIYLA ÇOK BOYUTLU
DURUM İÇİN UMBRELLA TEST

160047

Ahmet NASIRLI
YKSEK LİSANS TEZİ
MATEMATİK ANABİLİM DALI

TEZ DANIŞMANI
Yrd.Doç.Dr. Mustafa NADAR

GEBZE
2003

..Ahmet...NASIRLI...’ in tez çalışması, G.Y.T.E Mühendislik ve
Fen Bilimleri Enstitüsü Yönetim Kurulu’nun ..16.07.2009..... tarih ve
..2009/22..... sayılı kararıyla oluşturulan jüri tarafından
.....Matematik..... Anabilim Dalında YÜKSEK LİSANS tezi
olarak kabul edilmiştir.



160047

JÜRİ

ÜYE

(Tez Danışmanı)

: Yrd. Doç. Dr. Mustafa Nader

M. Nader

ÜYE

: Prof. Dr. Tahir Aliyev

A

ÜYE

: Yrd. Doç. Dr. Bulut Akbaş

Bulut Akbaş

ONAY

G.Y.T.E Mühendislik ve Fen Bilimleri Enstitüsü Yönetim Kurulu’nun
..06.11.2009..... tarih ve ..2009/76..... sayılı kararı.

İmza / Mühür
Prof. Dr. Adem KILIÇ
GYTE Mühendislik ve Fen Bil.
Enstitüsü Müdürü

ÖZET

Bu çalışmada, Gregory A.Mack ve Douglas A.Wolfe [Bkz 5] tarafından tek boyutlu durum için geliştirilmiş olan Umbrella testini çok boyutlu durum için bileşen yaklaşımıyla genelledik. Öncelikle çok boyutlu durum için bir A_g test istatistiği oluşturuldu.

$$A_g = \sum_{1 \leq u < v \leq \ell} U_{uv} + \sum_{\ell \leq u < v \leq k} U_{vu} \quad g=1, \dots, p$$

Daha sonra A_g test istatistiğimizin null hipotezi altında asimptotik dağılımının beklenen değeri 0 ve kovaryans matrisinin

$$\Sigma_A = \left[\left(1 - \sum_{j=1}^{\ell} \lambda_{j1}^3 \right) + \left(1 - \sum_{j=\ell}^k \lambda_{j2}^3 \right) + 2 \lambda_{\ell} \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^k \lambda_{j1} \lambda_{j2} \right] \rho_{gh} / 36$$

olan p boyutlu normal dağılım olduğunu gösterdik.

Nispi etkinlik değerlerini bulabilmek için yerel alternatif altında A_g test istatistiğimizin asimptotik dağılımını inceledik. Ayrıca, normal teori yöntemleri kullanılarak geliştirilmiş tek boyutlu umbrella test istatistiğini çok boyutlu durum için inceledik. Son olarak, önceden seçilen α değerleri için sonlu örnek büyüklükleri aldığımızda yaklaşık kritik değerlerini elde ettik.

SUMMARY

In this study, Umbrella Test for univariate case developed by Gregory A. Mack and Douglas A. Wolfe has been generalized to multivariate case with component –wise approach. Firstly, an A_g test statistics for multivariate case has been constructed.

$$A_g = \sum_{1 \leq u < v \leq \ell} U_{uv} + \sum_{\ell \leq u < v \leq k} U_{vu} \quad g=1, \dots, p$$

Secondly, it is shown that under null hypothesis the asymptotic distribution is p variate normal with the expected value $\mathbf{0}$ and the covariance matrix is

$$\Sigma_A = \left[\left(1 - \sum_{j=1}^{\ell} \lambda_{j1}^3 \right) + \left(1 - \sum_{j=\ell}^k \lambda_{j2}^3 \right) + 2 \lambda_{\ell} \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^k \lambda_{j1} \lambda_{j2} \right] \rho_{gh} / 36$$

Moreover, to be able to get the relative efficiency values, the asymptotic asymptotic distribution of A_g test statistics had been developed under the local alternative. Furthermore, the univariate Umbrella test developed by using the normal theory approach has been generalized for the multivariate case. Finally, a critical values for a finite sample size under predetermined α values has been calculated.

TEŞEKKÜR

Değerli hocam ve tez danışmanım Yrd. Doç. Dr. Mustafa Nadar 'a gerek tez konusu seçiminde gerekse tezin hazırlanmasındaki esirgemediği yardım, katkı ve kıymetli zamanlarından dolayı, Matematik Bölümünün diğer kıymetli hocalarına ve Gebze Yüksek Teknoloji Enstitüsünün akademik ve idari personeline teşekkürü bir borç bilirim.

Ayrıca yüksek lisans eğitimim boyunca beni destekleyen Sayın Okul Müdürüm R. Yılmaz Özek'e ve değerli öğretmen arkadaşlara teşekkürler.

Ve bugünlere ulaşmamda en büyük destekçim sevgili aileme her şey için sonsuz teşekkürler.



İÇİNDEKİLER

SAYFA

ÖZET	IV
SUMMARY	V
TEŞEKKÜR	VI
İÇİNDEKİLER DİZİNİ	VII
1. GİRİŞ	1
1.1 Giriş	1
1.2 Temel Bilgiler	2
2. TEK BOYUTLU DURUMDA UMBRELLA TEST	4
2.1 Tepe Noktası Biliniyorken Umbrella Testi	4
3.ÇOK BOYUTLU DURUMDA UMBRELLA TEST	8
3.1 Tanımlar	8
3.2 Bileşen Yaklaşımıyla Umbrella Test	9
3.3 Null Hipotezi Altında Koşullu Dağılım	9
3.4 Null Hipotezi Altında Asimptotik Dağılım	15
3.5 Alternatif Hipotez Altında Asimptotik Dağılım	17
3.6 Bileşen Yaklaşımıyla Normal Teori İstatistiği	20
3.7 Null Hipotezi Altında	20
3.8 Alternatif Hipotez Altında Dağılım	21
3.9 Çok Boyutlu Test İstatistikleri	22
3.10 Asimptotik Nispi Etkinlik	22
4.SİMÜLASYON YÖNTEMİYLE YAKLAŞIK KRİTİK	25
DEĞERLERİN BULUNMASI	
5.SONUÇLAR ve GELECEĞE YÖNELİK ÇALIŞMALAR	27
KAYNAKLAR	28
ÖZGEÇMİŞ	29

1 GİRİŞ

1.1 Giriş

Bilindiği gibi istatistiksel teknikler, parametrel ve parametresiz olmak üzere iki ana grupta toplanmaktadır. Parametrik teknikler; yoğunluk fonksiyonunun bilindiği, parametresiz teknikler ise yoğunluk fonksiyonunun bilinmediği tekniklerdir. Parametrik teknikler kesin sonuçlar vermesine rağmen yaygın olarak karşılaşılan durumlarda yoğunluk fonksiyonunu kesin olarak bilemeyiz. Parametresiz tekniklerde gözlemlerin gerçek değerleri değil; işaretleri veya sıraları (rank) gibi rastgele değişkenler kullanılır.

Bu çalışmada, parametrik olmayan teknikler kullanılarak geliştirilen tek boyutlu umbrella testini,

$$A_\ell = \sum_{1 \leq i < j \leq \ell} U_{ij} + \sum_{\ell \leq i < j \leq k} U_{ji}$$

bileşen yaklaşımıyla çok boyutlu umbrella testine genelledik. Burada ℓ 'ye tepe noktası denilmektedir. Tek boyutlu umbrella testte, gözlemler tek bir değişken hakkında bilgi içermektedir. Geliştirdiğimiz çok boyutlu umbrella testte ise gözlemlerin vektör olup birkaç değişken hakkında bilgi içermesini sağladık. Bütün grupların sıralı olması durumunda çok boyutlu yapısı E.J Dietz'in[Bkz 1] tarafından gösterilmiştir. Çalışmada ℓ tepe noktasının bilindiği durum incelenmiştir.

İkinci bölümde literatürde mevcut olan tek boyutlu umbrella test ele alındı. Üçüncü bölümde, çok boyutlu umbrella testi için tanımlar ve geliştirilen testin

$$H_0 : F_1^{(1)}(x) = \dots = F_k^{(p)}(x) \text{ bütün } x \text{ 'ler için} \quad (1)$$

H_0 altındaki koşullu beklenen değeri ve varyans kovaryans matrisi oluşturuldu. Testin alternatif hipotez altındaki asimptotik dağılımı gösterildi. Dördüncü bölümde, eğer yoğunluk fonksiyonu biliniyor ise umbrella test için asimptotik dağılımı oluşturuldu. Son bölümde ise, geliştirdiğimiz A_g testini simülasyon yöntemi kullanarak yaklaşık kritik değerleri bulundu.

1.2 Temel Bilgiler

Tez içinde geçen bazı terimlerin tanımları burada verilecektir.

Tanım 1 Olasılık Uzayı:

Ω bir küme, F , Ω 'nın alt kümelerinden oluşmuş bir σ - cebir ve P , Kolmogrov

aksiyomlarını sağlamak üzere, F üzerinde tanımlı olasılık ölçüsünden oluşan (Ω, F, P) üçlüsüne olasılık uzayı denir.

Tanım 2 Ω 'nın alt kümelerinden oluşmuş F aşağıdaki koşulları sağlıyor ise bir σ – cebirdir denir.

- 1) $\emptyset \in F$
- 2) Eğer $A \in F$ ise $A^c \in F$
- 3) Eğer $\{A_n\}_{n \geq 1}$, F 'in elemanlarından oluşmuş bir dizi ise $\cup A_n \in F$ dir

Tanım 3 *Rastgele Değişken:*

$X : \Omega \longrightarrow R \cup \{\pm\infty\}$ fonksiyonu $\forall \lambda \in R$ için $\{w : X(w) < \lambda\} \in F$ koşulunu sağlıyorsa ona (Ω, F, P) olasılık uzayında rastgele değişken veya (Ω, F, P) üzerinde F -ölçülebilir fonksiyon denir.

Tanım 4 *Olasılık Yoğunluk Fonksiyonu:*

X rastgele değişkeni sürekli olsun ve bir $[a, b]$ aralığında değerler alsın. $f_x(x)$ fonksiyonu aşağıdaki koşulları sağlıyorsa ona olasılık yoğunluk fonksiyonu denir.

- 1) $f_x(x) \geq 0$
- 2) $\int_a^b f_x(x) dx = 1$

Tanım 5 *Beklenen Değer:*

Sürekli bir X rastgele bir değişkeninin olasılık yoğunluk fonksiyonu $f(x)$ olsun ve $g(x)$ 'de X 'in tek değerlikli bir fonksiyonu olsun. $g(x)$ 'in beklenen değeri

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

biçiminde tanımlanır.

$g(x) = (x - a)^r$ biçiminde seçilirse

$$E[g(x)] = E[(x - a)^r] = \int_{-\infty}^{\infty} (x - a)^r f(x)dx$$

ifadesi, X rastgele değişkeninin a üzerinden hesaplanan r .inci momenti olarak isimlendirilir. Özel olarak $r=2$ aldığımızda

$$E[(x - a)^2] = \int_{-\infty}^{\infty} (x - a)^2 f(x)dx$$

ifadesine X 'in varyansı denir ve $Var(X)$ biçiminde gösterilir.

$Var(X) = E[X^2] - (E[X])^2$ veya $Var(X) = E[X - E[X]]^2$ biçimlerinde de gösterilebilir.

Tanım 6 Kovaryans:

$E(X) = \mu_x$ ve $E(Y) = \mu_y$ olsun. O halde

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

şeklinde tanımlanan eşitliğe kovaryans denir.

Tanım 7 Mann-Whitney U_{ij} Fonksiyonu:

i. ve j. grupta yer alan gözlemlerden, j. gruptaki gözlemlerin i. gruptaki gözlemlerden büyük olanlarının sayısını veren fonksiyondur.

Tanım 8 Umbrella Test:

k tane grup için öyle bir ℓ , $1 < \ell < k$ grubu vardır ki ℓ 'den önce gruplar arası bir artışın ve ℓ 'den sonrası için ise azalışın olup olmadığını ortaya koymayı amaçlayan test istatistiğidir.

Tanım 9 k-Grup U-Istatistiği:

Derecesi $(r_1, \dots, r_k)$ ve simetrik kernel fonksiyonu $h(\cdot)$ olan γ parametresi için $n_i \geq r_i$ olduğunda, aşağıdaki şekilde gösterilen form k-grup U-istatistiğidir. Burada $\alpha_i = (\alpha_{i1}, \dots, \alpha_{ir_i})$ ve A_i 'ler $\{1, \dots, n_i\}$ ($i = 1, \dots, k$) sayılarından tekrarsız seçilen r_i 'li bütün alt kümelerin bir koleksiyonudur.

$$U(X_{11}, \dots, X_{1n_1}; \dots; X_{k1}, \dots, X_{kn_k}) = \frac{1}{\binom{n_1}{r_1} \dots \binom{n_k}{r_k}} \sum_{\alpha_1 \in A_1} \dots \sum_{\alpha_k \in A_k} h(X_{1\alpha_{11}}, \dots, X_{1\alpha_{1r_1}}; \dots; X_{k\alpha_{k1}}, \dots, X_{k\alpha_{kr_k}}) \quad (2)$$

Tanım 13 Merkezi Limit Teoremi:

$X_1, X_2, \dots$ ler aynı dağılımdan gelen beklenen değeri $E(X_i) = \mu$ ve varyansı $Var(X_i) = \sigma^2$ $\sigma^2 < \infty$, olan rastgele değişkenler olsun. Bu takdirde,

$$\sqrt{n} \frac{(\bar{X} - \mu)}{\sigma} \xrightarrow[n \rightarrow \infty]{d} Z \sim n(0, 1) \text{ dir.}$$

2 TEK BOYUTLU DURUMDA UMBRELLA TEST

Varsayalım $X_{ij}, j=1, \dots, n_i$ değişkenleri $F_i(x)$ sürekli dağılım fonksiyonu olup $X_{i1}, X_{i2}, \dots, X_{in_i}, i=1, \dots, k$; k tane bağımsız rastgele örnekler olsun. Null hipotezimizi aşağıdaki şekilde yazalım.

$$H_0 : F_1(x) = \dots = F_\ell(x) = \dots = F_k(x) \quad (3)$$

Parametresiz tekniklerde bu durum için geliştirilmiş bir çok test yöntemi literatürde mevcuttur. Ancak H_0 'a alternatif olarak önceden bir bilgi mevcut ise parametresiz tekniklerde daha isabetli seçim yapabilmekteyiz. Örneğin; k 'nın en az ikisinin farklı dağılımdan geldiği araştırılıyorsa yaygın olarak kullanılan yöntem *Kruskal Wallies* [Bkz 2] testidir. Eğer k grup arasında bir artışın olup olmadığı araştırılıyorsa alternatif hipotezimiz,

$$H_a : F_1(x) \geq \dots \geq F_\ell(x) \geq \dots \geq F_k(x) \quad (4)$$

şeklinde yazılır ve en az bir noktada mutlak eşitsizlik mevcuttur. Bu H_a için geliştirilen test ise *Jonkheere* [Bkz 3] ve *Tempstra* [Bkz 4] testidir. Alternatif hipotezimiz,

$$H_a : F_1(x) \geq \dots \geq F_\ell(x) \leq \dots \leq F_k(x) \quad (5)$$

için *A. Mack ve A. Wolfe* [Bkz 5] parametrik olmayan teknikler kullanarak (5)'de tanımlanan umbrella test istatistiğini geliştirmişlerdir. Literatürde (5) durumu *umbrella alternatives* olarakda adlandırılmaktadır. Burada ℓ 'ye tepe (peak) noktası denmektedir. Aşağıdaki yapılanlarda ℓ 'nin bilindiği durumlar gözöntüne alınmıştır. Eğer ℓ bilinmiyorsa, önce elimizdeki verilerden yararlanarak mevcut metodlarla grup saptanır.

2.1 Tepe Noktası Biliniyorken Umbrella Testi

Varsayalım ℓ ile işaretlenen tepe noktası biliniyor olsun. Bu durumda (5)'de ifade edilmiş H_a 'nın lehinde H_0 hipotezimizi test etmek için,

$$A_\ell = \sum_{1 \leq i < j \leq \ell} U_{ij} + \sum_{\ell \leq i < j \leq k} U_{ji} \quad (6)$$

test istatistiğini kullanırız. A_ℓ 'nin yeteri kadar büyük değerleri için H_0 reddedilir. Burada U_{ij} , *Mann-Whitney* istatistiğidir [Bkz 6]. ℓ 'nin bilindiği durumda ki asimtotik özellikleri *Archambav, Mack ve Wolfe* [Bkz 7] tarafından gösterilmiştir. Örneğin; eğer

her $n_i \rightarrow \infty$, $N = \sum_{i=1}^k n_i \rightarrow \infty$ benzer şekilde $\lim_{n \rightarrow \infty} \frac{n_i}{N} = \lambda_i$, $0 < \lambda_i < 1$, $i=1, \dots, k$ olup

$$\frac{A_\ell - \mu_0(A_\ell)}{\sqrt{\nu_0(A_\ell)}} \xrightarrow{MLT} N(0, 1) \quad (7)$$

olur.

Theorem 1 H_0 hipotezi altında A_ℓ istatistiğinin beklenen değeri ve varyansı

$$E_0(A_\ell) = \left[N_1^2 + N_2^2 - \sum_{i=1}^k n_i^2 - n_\ell^2 \right] / 4 \quad (8)$$

$$\nu_0^2(A_\ell) = \frac{1}{72} \left[2(N_1^3 + N_2^3) + 3(N_1^2 + N_2^2) - \sum_{i=1}^k n_i^2(2n_i + 3) - n_\ell^2(2n_\ell + 3) + 12n_\ell N_1 N_2 - 12n_\ell^2 N \right] \quad (9)$$

dir. Burada; $N_1 = \sum_{i=1}^{\ell} n_i$, $N_2 = \sum_{i=\ell}^k n_i$ dir.

İspat. A_ℓ istatistiğini, $A_\ell = J_{\text{artan}} + J_{\text{azalan}}$, şeklinde 1. ile ℓ ve ℓ ile k grupları arasındaki Jonckheere–Terpstra istatistikleri olarak yazalım. Öyleyse null hipotezi altında A_ℓ 'nin beklenen değeri;

$$\begin{aligned} E_0(A_\ell) &= E_0(J_{\text{artan}}) + E_0(J_{\text{azalan}}) \\ &= \frac{1}{4} \left[N_1^2 - \sum_{i=1}^{\ell} n_i^2 \right] + \frac{1}{4} \left[N_2^2 - \sum_{i=\ell}^k n_i^2 \right] \\ &= \frac{1}{4} \left[N_1^2 + N_2^2 - \sum_{i=1}^k n_i^2 - n_\ell^2 \right] \end{aligned} \quad (10)$$

Benzer şekilde A_ℓ 'nin varyans değeri,

$$\begin{aligned} \nu_0(A_\ell) &= \nu_0(J_{\text{artan}} + J_{\text{azalan}}) \\ &= \nu_0(J_{\text{artan}}) + \nu_0(J_{\text{azalan}}) + 2\text{Cov}_0(J_{\text{artan}}, J_{\text{azalan}}) \end{aligned} \quad (11)$$

Öncelikle (11)'daki son terime bakalım,

$$\begin{aligned}
 Cov_0(J_{artan}, J_{azalan}) &= Cov_0\left(\sum_{u=1}^{v-1} \sum_{v=2}^{\ell} U_{uv}, \sum_{s=\ell}^{t-1} \sum_{t=\ell+1}^k U_{ts}\right) \\
 &= Cov_0\left(\sum_{u=1}^{v-1} \sum_{v=2}^{\ell-1} U_{uv} + \sum_{u=1}^{\ell-1} U_{u\ell}, \sum_{s=\ell+1}^{t-1} \sum_{t=\ell+2}^k U_{ts} + \sum_{t=\ell+1}^k U_{t\ell}\right) \\
 &= Cov_0\left(\sum_{u=1}^{v-1} \sum_{v=2}^{\ell-1} U_{uv}, \sum_{s=\ell+1}^{t-1} \sum_{t=\ell+2}^k U_{ts}\right) + Cov_0\left(\sum_{u=1}^{v-1} \sum_{v=2}^{\ell-1} U_{uv}, \sum_{t=\ell+1}^k U_{t\ell}\right) \\
 &\quad + Cov_0\left(\sum_{u=1}^{\ell-1} U_{u\ell}, \sum_{s=\ell+1}^{t-1} \sum_{t=\ell+2}^k U_{ts}\right) + Cov_0\left(\sum_{u=1}^{\ell-1} U_{u\ell}, \sum_{t=\ell+1}^k U_{t\ell}\right)
 \end{aligned} \tag{12}$$

Açıktırki;

$$Cov_0\left(\sum_{u=1}^{v-1} \sum_{v=2}^{\ell-1} U_{uv}, \sum_{s=\ell+1}^{t-1} \sum_{t=\ell+2}^k U_{ts}\right) = Cov_0\left(\sum_{u=1}^{v-1} \sum_{v=2}^{\ell-1} U_{uv}, \sum_{t=\ell+1}^k U_{t\ell}\right) = 0$$

Benzer şekilde, $\sum_{u=1}^{\ell-1} U_{u\ell}$ ℓ .gruptan gelen gözlemleri, $\sum_{s=\ell+1}^{t-1} \sum_{t=\ell+2}^k U_{ts}$ ise $\ell+1, \dots, k$ gruplarından gelen gözlemlerini içerdiğinden dolayı,

$$Cov_0\left(\sum_{u=1}^{\ell-1} U_{u\ell}, \sum_{s=\ell+1}^{t-1} \sum_{t=\ell+2}^k U_{ts}\right) = 0$$

Sonuç olarak,

$$\begin{aligned}
 Cov_0(J_{artan}, J_{azalan}) &= Cov_0\left(\sum_{u=1}^{\ell-1} U_{u\ell}, \sum_{t=\ell+1}^k U_{t\ell}\right) \\
 &= \sum_{u=1}^{\ell-1} \sum_{t=\ell+1}^k Cov_0(U_{u\ell}, U_{t\ell})
 \end{aligned} \tag{13}$$

$$Cov(U_{u\ell}, U_{t\ell}) = n_u n_v n_t / 12 \quad [Bkz 8] \tag{14}$$

yerine koyulduğunda

$$\begin{aligned}
 Cov_0(J_{artan}, J_{azalan}) &= \frac{n_\ell}{12} \sum_{u=1}^{\ell-1} \sum_{t=\ell+1}^k n_u n_t = \frac{n_\ell}{12} \left(\sum_{u=1}^{\ell-1} n_u\right) \left(\sum_{t=\ell+1}^k n_t\right) \\
 &= \frac{n_\ell (N_1 - n_\ell) (N_2 - n_\ell)}{12}
 \end{aligned} \tag{15}$$



ve

$$\begin{aligned} v_0(J_{\text{artan}}) &= \frac{1}{72} \left[N_1^2(2N_1 + 3) - \sum_{j=1}^{\ell} n_j^2(2n_j + 3) \right] \\ v_0(J_{\text{azalan}}) &= \frac{1}{72} \left[N_2^2(2N_2 + 3) - \sum_{j=\ell}^k n_j^2(2n_j + 3) \right] [\text{Bkz } 8] \end{aligned}$$

Gerekli cebirsel işlemlerden sonra;

$$\nu_0^2(A_\ell) = \frac{1}{72} \left[\begin{aligned} &2(N_1^3 + N_2^3) + 3(N_1^2 + N_2^2) - \sum_{i=1}^k n_i^2(2n_i + 3) \\ &- n_\ell^2(2n_\ell + 3) + 12n_\ell N_1 N_2 - 12n_\ell^2 N \end{aligned} \right] \quad (16)$$

elde edilir. ■



3 ÇOK BOYUTLU DURUMDA UMBRELLA TEST

Önceki bölümde, literatürde olan tek boyutlu durumda umbrella testi için yapılan çalışmaları gösterdik. Bu bölümde ise çok boyutlu durumda umbrella testini geliştireceğiz.

3.1 Tanımlar

Varsayalım $\mathbf{X}_{ij} = (X_{ij1}, \dots, X_{ijp})'$ $p \times 1$ boyutlu vektör ve j . grubun i . gözlemini göstereyin. ($j=1, \dots, k$ $i=1, \dots, n_j$). $\mathbf{X}_{ij}$ 'lerin her biri bağımsız olup $F_j(\mathbf{X})$ sürekli dağılım fonksiyonu ve $F_j^{(1)}(x), \dots, F_j^{(p)}(x)$ $j=1, \dots, k$ marjinal dağılım fonksiyonuna sahip olsun. Null hipotezimiz;

$$H_0 : F_1^{(1)}(x) = \dots = F_k^{(p)}(x) \text{ bütün } x \text{ 'ler için} \quad (17)$$

Alternatif hipotezimiz;

$$H_a : F_1^{(g)}(x) \geq \dots, F_{\ell-1}^{(g)}(x) \geq F_{\ell}^{(g)}(x) \leq F_{\ell+1}^{(g)}(x) \leq \dots \leq F_k^{(g)}(x) \quad (18)$$

en az bir g için mutlak eşitsizlik sağlanır ($g = 1, \dots, p$).

H_a durumu bize bileşenlerden en az birinin ℓ . gruba kadar istatistiksel olarak bir artışın daha sonrası için ise bir azalışın olduğunu göstermektedir. Tanımlayacağımız çok boyutlu umbrella istatistiği sadece gözlemlerin ranklarına bağlı olacaktır. Varsayalım $\mathbf{R}_l = (R_{l1}, \dots, R_{lp})'$ ($l=1, \dots, N$) $\mathbf{X}_{ij}$ 'lerin bileşen bazında p boyutlu rankını göstereyin. $\mathbf{X}_{ij}$ 'ler rastgele vektör ise rankları da rastgele bir vektör oluşturacaktır. Burada $N = \sum_{j=1}^k n_j$ dir.

Aşağıdaki tanımlara daha sonraki bölümlerde ihtiyacımız olacaktır.

1. g ve h bileşenleri arasındaki Kendall korelasyon katsayısı

$$\tau_{gh} = 2 \sum_{i < j}^N \text{sign} [(R_{ig} - R_{jh})(R_{jh} - R_{ih})] / N(N-1) \quad (19)$$

$$\text{sign}(a) = \begin{cases} 1 & \text{eğer } a > 0 \\ 0 & \text{eğer } a = 0 \\ -1 & \text{eğer } a < 0 \end{cases} \quad (20)$$

2. g ve h bileşenleri arasındaki Sperman korelasyon katsayısı [Bkz 9]

$$r_{gh} = 3 \sum_{i,j,k}^N \text{sign} [(R_{jg} - R_{ig})(R_{jh} - R_{kh})] / (N^3 - N) \quad (21)$$

3. r_{gh} istatistliği olasılıkta

$$\rho_{gh} = \lim_{N \rightarrow \infty} E(r_{gh}) = 3(2\pi_{gh} - 1) \quad (22)$$

$$\pi_{gh} = P[(R_{jg} - R_{ig})(R_{jh} - R_{kh}) > 0] \quad (23)$$

[Bkz 10] Bu tanımları verdikten sonra çok boyutlu durum için bileşen yardımıyla umbrellla test istatistliğini tanımlayalım.

3.2 Bileşen Yaklaşımıyla Umbrella Testi

Her g ($g=1, \dots, p$) bileşeni için test istatistliğimizi aşağıdaki gibi tanımlayalım.

$$A_g = \sum_{1 \leq u < v \leq \ell}^{\ell} (U_{uv}^{(g)} - n_u n_v / 2) + \sum_{\ell \leq u < v \leq k}^k (U_{vu}^{(g)} - n_u n_v / 2); \quad g=1, \dots, p \quad (24)$$

$$U_{uv}^{(g)} = \sum_{i=1}^{n_u} \sum_{i'=1}^{n_v} \phi(X_{iug}, X_{i'vg}) \quad (25)$$

u ve v gruplarında, $u < v$, g bileşenine göre Mann-Whitney [Bkz 6] istatistliği olsun.

$$\phi(a, b) = \begin{cases} 1 & \text{eğer } a < b \\ 1/2 & \text{eğer } a = b \\ 1 & \text{eğer } a > b \end{cases} \quad (26)$$

Null hipotezi altında A_g , g bileşenine göre merkezleştirilmiş olur, yani $E(A_g) = 0$ her g için.

3.3 Null Hipotezi Altında Koşullu Dağılım

H_0 altında, N tane R_i vektörünün k gruba aynı şansla atanması $N!/(n_1! \dots n_k!)$ farklı şekilde olur. Bu permütasyon dağılımı altında A_g nin beklenen değeri,

$$E_0(A_g) = 0 \quad (27)$$

ve varyansı teorem 1'de ispatlandığı gibi,

$$\nu_0^2(A_g) = \frac{1}{72} \left[\begin{aligned} &2(N_1^3 + N_2^3) + 3(N_1^2 + N_2^2) - \sum_{i=1}^k n_i^2(2n_i + 3) \\ &- n_\ell^2(2n_\ell + 3) + 12n_\ell N_1 N_2 - 12n_\ell^2 N \end{aligned} \right] \quad (28)$$

dır. Burada; $N_1 = \sum_{i=1}^{\ell} n_i$, $N_2 = \sum_{i=\ell}^k n_i$, ve $N = N_1 + N_2 - n_\ell$ dir.

Theorem 2 H_0 altında A_g ve A_h 'in koşullu kovaryansı,

$$\begin{aligned} Cov_0(A_g, A_h) = & (N+1) \left[\left(N_1^3 - \sum_{j=1}^{\ell} n_j^3 \right) - 3 \left(N_1^2 - \sum_{j=1}^{\ell} n_j^2 \right) \right] r_{gh}/36(N-2) \\ & + \left[3N \left(N_1^2 - \sum_{j=1}^{\ell} n_j^2 \right) - 2 \left(N_1^3 - \sum_{j=1}^{\ell} n_j^3 \right) \right] \tau_{gh}/24(N-2) \\ & + n_\ell(N_1 - n_\ell)(N_2 - n_\ell)[(N+1)r_{gh} - 3\tau_{gh}]/6(N-2) \\ & + (N+1) \left[\left(N_2^3 - \sum_{j=\ell}^k n_j^3 \right) - 3 \left(N_2^2 - \sum_{j=\ell}^k n_j^2 \right) \right] r_{gh}/36(N-2) \\ & + \left[3N \left(N_2^2 - \sum_{j=\ell}^k n_j^2 \right) - 2 \left(N_2^3 - \sum_{j=\ell}^k n_j^3 \right) \right] \tau_{gh}/24(N-2) \quad (29) \end{aligned}$$

Burada ki τ_{gh} ve r_{gh} sırasıyla (19) ve (21) de tanımlanmıştır.

İspat:. İspata geçmeden önce kovaryansın bize yardımcı olacak özelliklerinden birini hatırlayalım.

$$Cov(ax_1 + bx_2; cy_1 + dy_2) = acCov(x_1, y_1) + adCov(x_1, y_2) + bcCov(x_2, y_1) + bdCov(x_2, y_2)$$

Öyleyse şimdi her bir gözlemin bileşenleri arasındaki kovaryansı ve bileşenlerin birbirlerine karşı olan kovaryanslarını bulalım. Her bir gözlemin kendi bileşenleri arasındaki kovaryansını (15)'de göstermiştik. Bileşenler arasındaki kovaryansı hesaplarken A_g test istatistiğimizi $A_g = A_{g_1} + A_{g_2}$ şeklinde ikiye ayırıp *Jonkheere-Tempstra* trend testinden yararlanacağız. Notasyonu basite indirgeyebilmek için X_{ij} ve Y_{ij} herhangi iki gözlemimizin sırasıyla g. ve h. bileşenlerine karşı gelen $j=1, \dots, k$; $i=1, \dots, n_j$ gözlemleri olsun. Bu durumda $R_i(S_i)$ ile $i=1, \dots, N$; X_{ij} (Y_{ij}) gözlemlerinin kendi aralarındaki

ranklarını belirtsin. Bu durumda,

$$\begin{aligned}
 A_X &= \sum_{1 \leq u < v \leq \ell}^{\ell} \left(U_{uv}^{(g)} - n_u n_v / 2 \right) + \sum_{\ell \leq u < v \leq k}^k \left(U_{vu}^{(g)} - n_u n_v / 2 \right) \\
 &= \sum_{1 \leq u < v \leq \ell}^{\ell} \left[\sum_{i=1}^{n_u} \sum_{i'=1}^{n_v} \phi(X_{iu}, X_{i'v}) - n_u n_v / 2 \right] \\
 &\quad + \sum_{\ell \leq u < v \leq k}^k \left[\sum_{j=1}^{n_{u'}} \sum_{j'=1}^{n_{v'}} \phi(X_{j'v}, X_{ju}) - n_{u'} n_{v'} / 2 \right] \\
 &= \left[\sum_{1 \leq u < v \leq \ell}^{\ell} \sum_{i=1}^{n_u} \sum_{i'=1}^{n_v} \text{sign}(X_{i'v} - X_{iu}) / 2 + \sum_{\ell \leq u < v \leq k}^k \sum_{j=1}^{n_{u'}} \sum_{j'=1}^{n_{v'}} \text{sign}(X_{ju} - X_{j'v}) / 2 \right]
 \end{aligned} \tag{30}$$

A_Y 'da benzer şekilde tanımlanır.

$E_0(A_X) = E_0(A_Y) = 0$ olduğundan dolayı A_X ve A_Y 'lerin koşullu kovaryansı ise,

$$Cov_0(A_X, A_Y) = E(A_X A_Y \mid (R_i, S_i), i = 1, \dots, N) \tag{31}$$

dir. (31) dört parçanın toplamı olarak incelenebilir.

$$\begin{aligned}
 Cov_0(A_X, A_Y) &= Cov_0(A_{g1}, A_{h1}) + Cov_0(A_{g1}, A_{h2}) \\
 &\quad + Cov_0(A_{g2}, A_{h1}) + Cov_0(A_{g2}, A_{h2})
 \end{aligned} \tag{32}$$

Burada,

$$\begin{aligned}
 Cov_0(A_{g1}, A_{h1}) &= E \left\{ \sum_{u < v}^{\ell} \sum_{i_1=1}^{n_u} \sum_{i_2=1}^{n_v} \text{sign}(X_{i_2v} - X_{i_1u}) / 2 \times \sum_{u' < v'}^{\ell} \sum_{i'_1=1}^{n_{u'}} \sum_{i'_2=1}^{n_{v'}} \text{sign}(Y_{i'_2v} - Y_{i'_1u}) \right\} \\
 Cov_0(A_{g1}, A_{h2}) &= E \left\{ \sum_{u < \ell}^{\ell} \sum_{i_1=1}^{n_u} \sum_{j=1}^{n_{\ell}} \sum_{v > \ell}^k \sum_{i_2=1}^{n_v} \text{sign}(X_{j\ell} - X_{i_1u}) / 2 \times \text{sign}(Y_{j\ell} - Y_{i_2v}) / 2 \right\} \\
 Cov_0(A_{g2}, A_{h1}) &= E \left\{ \sum_{\ell < k}^k \sum_{i_1=1}^{n_u} \sum_{j=1}^{n_{\ell}} \sum_{u < \ell}^{\ell} \sum_{i_2=1}^{n_v} \text{sign}(X_{j\ell} - X_{i_1u}) / 2 \times \text{sign}(Y_{j\ell} - Y_{i_2v}) / 2 \right\} \\
 Cov_0(A_{g2}, A_{h2}) &= E \left\{ \sum_{\ell \leq u < v}^k \sum_{i_1=1}^{n_u} \sum_{i_2=1}^{n_v} \text{sign}(X_{i_2u} - X_{i_1v}) / 2 \times \sum_{\ell \leq u' < v'}^k \sum_{i'_1=1}^{n_{u'}} \sum_{i'_2=1}^{n_{v'}} \text{sign}(Y_{i'_2u} - Y_{i'_1v}) \right\}
 \end{aligned}$$

dır. Şimdi bu değerleri tek tek hesaplayalım.

$$\begin{aligned}
 4Cov_0(A_{g_1}, A_{h_1}) &= E \left[\sum_{u < v}^{\ell} \sum_{i_1=1}^{n_u} \sum_{i_2=1}^{n_v} \text{sign}(X_{i_2v} - X_{i_1u}) \times \sum_{u' < v'}^{\ell} \sum_{i'_1=1}^{n_u} \sum_{i'_2=1}^{n_v} \text{sign}(Y_{i'_2v'} - Y_{i'_1u'}) \right] \\
 &= \sum_{u < v}^{\ell} \sum_{i_1=1}^{n_u} \sum_{i_2=1}^{n_v} \sum_{u' < v'}^{\ell} \sum_{i'_1=1}^{n_u} \sum_{i'_2=1}^{n_v} E \left[\text{sign}(X_{i_2v} - X_{i_1u}) \text{sign}(Y_{i'_2v'} - Y_{i'_1u'}) \right] \quad (33)
 \end{aligned}$$

Burada indisler gözönüne alındığında dört durum gözlenir.

(1) Eğer i_2v, i_1u, i'_2v, i'_1u dört farklı gözlemi gösteriyorsa (33) 'ün beklenen değeri sıfırdır;

(2) i_2v, i_1u, i'_2v ve i'_1u indisleri iki farklı gözlemi gösterebilir (toplamdaki 1 terim);

(3) Bu indisler iki farklı düzeyden üç farklı gözlemi gösterebilir (toplamdaki 2. ve 3. terimler);

(4) Son olarakta üç farklı düzeyden üç farklı gözlemi gösterebilirler.

Öyleyse,

$$\begin{aligned}
 4Cov_0(A_{g_1}, A_{h_1}) &= \\
 &\sum_{u < v}^{\ell} \sum_{i_1=1}^{n_u} \sum_{i_2=1}^{n_v} E [\text{sign}(X_{i_2v} - X_{i_1u}) \text{sign}(Y_{i_2v} - Y_{i_1u})] \\
 &+ \sum_{\substack{j=1 \\ j \neq i_2}}^{n_v} E [\text{sign}(X_{i_2v} - X_{i_1u}) \text{sign}(Y_{jv} - Y_{i_1u})] \\
 &+ \sum_{\substack{j=1 \\ j \neq i_1}}^{n_u} E [\text{sign}(X_{i_2v} - X_{i_1u}) \text{sign}(Y_{i_2v} - Y_{ju})] \\
 &+ \sum_{\substack{w < v \\ w \neq u}}^{\ell} \sum_{j=1}^{n_w} E [\text{sign}(X_{i_2v} - X_{i_1u}) \text{sign}(Y_{i_2v} - Y_{jw})] \\
 &+ \sum_{w > v}^{\ell} \sum_{j=1}^{n_w} E [\text{sign}(X_{i_2v} - X_{i_1u}) \text{sign}(Y_{jw} - Y_{i_2v})] \\
 &+ \sum_{w < u}^{\ell} \sum_{j=1}^{n_w} E [\text{sign}(X_{i_2v} - X_{i_1u}) \text{sign}(Y_{i_1u} - Y_{jw})] \\
 &+ \sum_{\substack{w > v \\ w \neq u}}^{\ell} \sum_{j=1}^{n_w} E [\text{sign}(X_{i_2v} - X_{i_1u}) \text{sign}(Y_{jw} - Y_{i_1u})]
 \end{aligned}$$

$$\begin{aligned}
s &= E[\text{sign}(R_j - R_i)(S_j - S_i)] = \tau_{XY} \\
t &= E[\text{sign}(R_j - R_i)(S_j - S_k)] \\
t &= \sum_{j,i \neq k} \text{sign}(R_j - R_i)(S_j - S_k) / N(N-1)(N-2) \\
t &= \sum_{i,j,k} \text{sign}(R_j - R_i)(S_j - S_k) / N(N-1)(N-2) \\
&\quad - 2 \sum_{i < j} \text{sign}(R_j - R_i)(S_j - S_i) / N(N-1)(N-2) \\
t &= (N+1)r_{XY}/3(N-2) - \tau_{XY}/(N-2)
\end{aligned} \tag{34}$$

olsun. $\text{Sign}(a-b) = -\text{Sign}(b-a)$ özelliğinden yararlanarak

$$4\text{Cov}_0(A_{g_1}, A_{h_1}) = \sum_{u < v}^{\ell} \sum_{i_1=1}^{n_u} \sum_{i_2=1}^{n_v} \left\{ s + \sum_{\substack{j=1 \\ j \neq i_2}}^{n_v} t + \sum_{\substack{j=1 \\ j \neq i_1}}^{n_u} t + \sum_{\substack{w < v \\ w \neq u}}^{\ell} \sum_{j=1}^{n_w} t - \sum_{w > v}^{\ell} \sum_{j=1}^{n_w} t - \sum_{w < u}^{\ell} \sum_{j=1}^{n_w} t + \sum_{\substack{w > v \\ w \neq v}}^{\ell} \sum_{j=1}^{n_w} t \right\} \tag{35}$$

(35) 'teki eşitliği derlediğimizde,

$$\begin{aligned}
4\text{Cov}_0(A_{g_1}, A_{h_1}) &= s \sum_{u < v}^{\ell} n_u n_v + t \sum_{u < v}^{\ell} n_u n_v (n_u + n_v - 2) + 2t \sum_{u < w < v}^{\ell} n_u n_v n_w \\
&= (N+1) \left[\left(N_1^3 - \sum_{j=1}^{\ell} n_j^3 \right) - 3 \left(N_1^2 - \sum_{j=1}^{\ell} n_j^2 \right) \right] r_{gh}/36(N-2) \\
&\quad \left[3N \left(N_1^2 - \sum_{j=1}^{\ell} n_j^2 \right) - 2 \left(N_1^3 - \sum_{j=1}^{\ell} n_j^3 \right) \right] \tau_{gh}/24(N-2) \tag{36}
\end{aligned}$$

elde edilir. Burada $N_1 = \sum_{j=1}^{\ell} n_j$ dir.

Şimdi $\text{Cov}_0(A_{g_1}, A_{h_2})$ hesaplayalım. Burada X_{ij} ler tepe noktasından önce ve Y_{ij} ler ise tepe noktasından sonra gelen gözlemlerdir. Dolayısıyla indisler dört farklı gözlemi gösterdiğinde kovaryansın sıfır olacağı açıktır. Bunun için X_{ij} ve Y_{ij} gözlemlerin kovaryansı sadece tepe noktasının bulunduğu gruptaki gözlemlerde sıfırdan farklı ola-

caktır.

$$\begin{aligned}
 4Cov_0(A_{g_1}, A_{h_2}) &= \sum_{u < \ell}^{\ell} \sum_{i_1=1}^{n_u} \sum_{j=1}^{n_\ell} \sum_{v > \ell}^k \sum_{i_2=1}^{n_v} E[\text{sign}(X_{j\ell} - X_{i_1u}) \text{sign}(Y_{j\ell} - Y_{i_2v})] \\
 &= \sum_{u < \ell}^{\ell} \sum_{i_1=1}^{n_u} tn_\ell \left(\sum_{v > \ell}^k n_v \right) \\
 &= tn_\ell \left(\sum_{u < \ell}^{\ell} n_u \right) \left(\sum_{v > \ell}^k n_v \right) \\
 &= tn_\ell (N_1 - n_\ell) (N_2 - n_\ell)
 \end{aligned} \tag{37}$$

Benzer şekilde $Cov_0(A_{g_2}, A_{h_1})$ hesaplarsak,

$$\begin{aligned}
 4Cov_0(A_{g_2}, A_{h_1}) &= \sum_{\ell < k}^k \sum_{i_1=1}^{n_u} \sum_{j=1}^{n_\ell} \sum_{u < \ell}^{\ell} \sum_{i_2=1}^{n_v} E[\text{sign}(X_{j\ell} - X_{i_1u}) \text{sign}(Y_{j\ell} - Y_{i_2v})] \\
 &= tn_\ell (N_1 - n_\ell) (N_2 - n_\ell)
 \end{aligned} \tag{38}$$

olur.

$Cov_0(A_{g_2}, A_{h_2})$ 'yi hesap edelim. $\text{Sign}(a - b) = -\text{Sign}(b - a)$ olduğundan dolayı $Cov_0(A_{g_1}, A_{h_1})$ için yaptığımız işlemleri bu sefer $\ell \leq u < v \leq k$ koşulu için tekrar edeceğiz. Bulacağımız sonuç sadece tepe noktasından sonraki gözlemlerle ilgili olup N_1 yerine $N_2 = \sum_{j=\ell}^k n_j$ yazılmasıyla hesap edilir.

$$4Cov_0(A_{g_2}, A_{h_2}) = s(N_2^2 - \sum_{j=\ell}^k n_j^2)/2 + t \left[(N_2^3 - \sum_{j=\ell}^k n_j^3)/3 - (N_2^2 - \sum_{j=\ell}^k n_j^2) \right] \tag{39}$$

Sonuç olarak

$$\begin{aligned}
 4Cov_0(A_X, A_Y) &= s(N_1^2 - \sum_{j=1}^{\ell} n_j^2)/2 + t \left[(N_1^3 - \sum_{j=1}^{\ell} n_j^3)/3 - (N_1^2 - \sum_{j=1}^{\ell} n_j^2) \right] \\
 &\quad + 2tn_\ell (N_1 - n_\ell) (N_2 - n_\ell) \\
 &\quad + s(N_2^2 - \sum_{j=\ell}^k n_j^2)/2 + t \left[(N_2^3 - \sum_{j=\ell}^k n_j^3)/3 - (N_2^2 - \sum_{j=\ell}^k n_j^2) \right]
 \end{aligned} \tag{40}$$

olur. s ve t değerlerini yerlerine yazdığımızda teorem ispatlanmış olur.

$N_1 = N_2 = M$ alındığında, $M^{-3/2} \mathbf{A} = M^{-3/2} (A_1, \dots, A_p)$ istatistiğimizin koşullu beklenen değeri 0 ve koşullu kovaryans matrisi ise Σ_0 dır. Burada Σ_0 matrisinin

köşegenleri $M^{-3}Var(A_g)$ ve diğer yerler ise $M^{-3}Cov(A_g, A_h)$ terimleriyle oluşmuş bir matristir. ■

3.4 Null Hipotezi Altında Asimptotik Dağılım ($M=N_1=N_2$)

$M^{-3/2}A'$ 'nın asimptotik dağılımını aşağıdaki kısıtlar altında inceleyeceğiz. Varsayalım

$$\begin{aligned} \lim_{M \rightarrow \infty} \frac{n_{j1}}{M} &= \lambda_{j1}, j_1=1, \dots, \ell-1; 0 < \lambda_{j1} < 1, \\ \lim_{M \rightarrow \infty} \frac{n_{j2}}{M} &= \lambda_{j2}, j_2=\ell+1, \dots, k; 0 < \lambda_{j2} < 1, \\ \lim_{M \rightarrow \infty} \frac{n_\ell}{M} &= \lambda_\ell, 0 < \lambda_\ell < 1 \text{ dir.} \end{aligned}$$

Temelde burada söylenen M sonsuza giderken gruptaki gözlem sayılarının M 'e oranının sabit olduğudur. Şimdi aşağıdaki teoremi yazabiliriz.

Theorem 3 H_0 altında $M^{-3/2}A'$ 'nın asimptotik dağılımı p boyutlu normal dağılım olup beklenen değeri 0 ve kovaryans matrisi Σ_A dır.

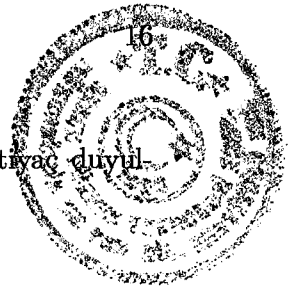
$$\Sigma_A = \left[\left(1 - \sum_{j=1}^{\ell} \lambda_{j1}^3 \right) + \left(1 - \sum_{j=\ell}^k \lambda_{j2}^3 \right) + 2\lambda_\ell \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^k \lambda_{i1} \lambda_{j2} \right] \rho_{gh}/36 \quad (41)$$

ρ_{gh} eşitlik (22)'de tanımlanmıştır.

İspat:. Bu teoremden $M^{1/2}(U_1 - \gamma_1), \dots, M^{1/2}(U_t - \gamma_t)$ ve $M^{1/2}(U_1 - \gamma_1), \dots, M^{1/2}(U_s - \gamma_s)$ 'lerin bileşik limit dağılımları gösterilecektir. Burada $U_1, \dots, U_t$ 'ler $\gamma_1, \dots, \gamma_t$ parametreleriyle, $U_1, \dots, U_s$ 'ler $\gamma_1, \dots, \gamma_s$ parametreleriyle k -grup için U istatistikleridir. Biz Randles ve Wolfe'un [Bkz 11] teorem 3.6.9 'unu g . bileşen için $t=\ell(\ell-1)p/2$, $s=(k-\ell)(k-\ell+1)p/2$ ve U istatistikleri $U_{uv}^{(g)}/n_u n_v$ ve $U_{vu}^{(g)}/n_u n_v$ olan sırasıyla $1 \leq u < v \leq \ell$ ve $\ell \leq u < v \leq k$ durumları için uygulayacağız. $U_{uv}^{(g)}/n_u n_v$ ve $U_{vu}^{(g)}/n_u n_v$ istatistikleri için parametreleri ise $\gamma_{uv}^{(g)} = P(X_{1vg} > X_{1ug})$ ve $\gamma_{vu}^{(g)} = P(X_{1ug} > X_{1vg})$ olarak seçtik. Teorem 3.6.9'a göre

$$\Omega_{uv}^{(g)} = M^{1/2} \left(U_{uv}^{(g)}/n_u n_v - \gamma_{uv}^{(g)} \right) \text{ ve } \Omega_{vu}^{(g)} = M^{1/2} \left(U_{vu}^{(g)}/n_u n_v - \gamma_{vu}^{(g)} \right) \quad (42)$$

istatistiğinin bileşik limit dağılımı $[\ell(\ell-1) + (k-\ell)(k-\ell+1)]p/2$ boyutlu normal dağılıma sahip olup beklenen değeri 0 ve varyansı ise aynı teoremin 3.6.10 eşitliği kullanılarak gösterilecektir.



Σ_A kovaryans matrisini hesaplayabilmek için aşağıdaki eşitliklere ihtiyaç duyulmaktadır. [(43) - (46)]

$$Cov \left(\Omega_{uv}^{(g)}, \Omega_{st}^{(h)} \right) = \begin{cases} 0 & \text{eğer } u, v, s, t \text{ hepsi farklı ise} \\ \rho_{gh}/12\lambda_{u_1} + \rho_{gh}/12\lambda_{v_1} & \text{eğer } u = s, v = t, g \neq h \\ 1/12\lambda_{u_1} & \text{eğer } u = s, v \neq t, g = h \\ \rho_{gh}/12\lambda_{u_1} & \text{eğer } u = s, v \neq t, g \neq h \\ 1/12\lambda_{v_1} & \text{eğer } u \neq s, v = t, g = h \\ \rho_{gh}/12\lambda_{v_1} & \text{eğer } u \neq s, v = t, g \neq h \\ -1/12\lambda_{v_1} & \text{eğer } v = s, g = h \\ -\rho_{gh}/12\lambda_{v_1} & \text{eğer } v = s, g \neq h \end{cases} \quad (43)$$

$$Cov \left(\Omega_{uv}^{(g)}, \Omega_{ts}^{(h)} \right) = \begin{cases} 0 & \text{eğer } v \neq \ell \text{ veya } t \neq \ell \\ 1/12\lambda_{\ell} & \text{eğer } \ell = v = t, g = h \\ \rho_{gh}/12\lambda_{\ell} & \text{eğer } \ell = v = t, g \neq h \end{cases} \quad (44)$$

$$Cov \left(\Omega_{vu}^{(g)}, \Omega_{st}^{(h)} \right) = \begin{cases} 0 & \text{eğer } u \neq \ell \text{ veya } s \neq \ell \\ 1/12\lambda_{\ell} & \text{eğer } \ell = u = s, g = h \\ \rho_{gh}/12\lambda_{\ell} & \text{eğer } \ell = u = s, g \neq h \end{cases} \quad (45)$$

$$Cov \left(\Omega_{vu}^{(g)}, \Omega_{ts}^{(h)} \right) = \begin{cases} 0 & \text{eğer } u, v, s, t \text{ hepsi farklı ise} \\ \rho_{gh}/12\lambda_{u_2} + \rho_{gh}/12\lambda_{v_2} & \text{eğer } u = s, v = t, g \neq h \\ 1/12\lambda_{v_2} & \text{eğer } v = t, u \neq s, g = h \\ \rho_{gh}/12\lambda_{v_2} & \text{eğer } v = t, u \neq s, g \neq h \\ 1/12\lambda_{u_2} & \text{eğer } v \neq t, u = s, g = h \\ \rho_{gh}/12\lambda_{u_2} & \text{eğer } v \neq t, u = s, g \neq h \\ -1/12\lambda_{u_2} & \text{eğer } u = t, g = h \\ -\rho_{gh}/12\lambda_{u_2} & \text{eğer } u = t, g \neq h \end{cases} \quad (46)$$

Açıkça, $M^{-3/2}A$ 'nın beklenen değeri 0 ve asimptotik kovaryansı ise (43)-(46) kullanılarak

$$\Sigma_A = \left[\left(1 - \sum_{j=1}^{\ell} \lambda_{j1}^3 \right) + \left(1 - \sum_{j=\ell}^k \lambda_{j2}^3 \right) + 2\lambda_{\ell} \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^k \lambda_{i1} \lambda_{j2} \right] \rho_{gh}/36 \quad (47)$$

dir. H_0 altında $\gamma_{uv}^{(g)} = \gamma_{vu}^{(g)} = 1/2$ olduğundan,

$$M^{-3/2}A_g = M^{-3/2} \left[\sum_{1 \leq u < v \leq \ell} n_u n_v \left(\Omega_{uv}^{(g)} M^{-1/2} + \gamma_{uv}^{(g)} - 1/2 \right) + \sum_{\ell \leq u < v \leq k} n_u n_v \left(\Omega_{vu}^{(g)} M^{-1/2} + \gamma_{vu}^{(g)} - 1/2 \right) \right] \quad (48)$$

$$M^{-3/2}A_g = \left(\sum_{1 \leq u < v \leq \ell} n_u n_v \Omega_{uv}^{(g)} + \sum_{\ell \leq u < v \leq k} n_u n_v \Omega_{vu}^{(g)} \right) / M^2, \quad g=1, \dots, p \quad (49)$$

dir. $\Omega_{uv}^{(g)}$ ve $\Omega_{vu}^{(g)}$ asimptotik normal dağıldığını göstermiştik. $M^{-3/2}A$ 'da bunların lineer kombinasyonu olmasından dolayı asimptotik normal dağılır. $M^{-3/2}A$ 'ların asimptotik kovaryansları [(43) - (46)] eşitliklerinden kolayca hesap edilir. ■

3.5 Alternatif Hipotez Altında Asimptotik Dağılım ($M=N_1=N_2$)

Nisbi etkinlik değerini elde edebilmek için yerel alternatif dizisi tanımlanmalıdır. $F_{jN}(\mathbf{X}) = F(\mathbf{X} - \theta_j M^{-1/2})$ olsun. Burada F ortalaması $\mathbf{0}$, kesin sürekli dağılım fonksiyonu ve f_g , F 'in g . bileşeninin marjinal dağılım fonksiyonu olsun. Varsayalım $\int f_g^2(x) dx < \infty$, $g=1, \dots, p$ için.

Theorem 4 F_{jN} yerel alternatif altında, $M^{-3/2}A$ asimptotik dağılımı p boyutlu normal dağılım olup ortalaması,

$$\mu_A = \left[\sum_{1 \leq u < v \leq \ell} \lambda_{u1} \lambda_{v1} (\theta_{vg} - \theta_{ug}) + \sum_{\ell \leq u < v \leq k} \lambda_{u2} \lambda_{v2} (\theta_{ug} - \theta_{vg}) \right] \int f_g^2(x) dx \quad (50)$$

ve kovaryans matrisi (41) ile verilen Σ_A dir.

İspat:. İlk önce $M^{-3/2}A$ 'nın asimptotik ortalamasını yerel alternatif altında bulalım. F_g ($g=1, \dots, p$); F 'in g . bileşeninin dağılım fonksiyonu, F_{gh} ise F 'in g . ve h . bileşenlerinin bileşik dağılım fonksiyonu olsun. $E(\Omega_{uv}^{(g)}) = 0$ ve $E(\Omega_{vu}^{(g)}) = 0$ olduğundan dolayı eşitlik (48) kullanılarak;

$$E(M^{-3/2}A_g) = M^{-3/2} \left[\sum_{1 \leq u < v \leq \ell} n_u n_v (\gamma_{uv}^{(g)} - 1/2) + \sum_{\ell \leq u < v \leq k} n_u n_v (\gamma_{vu}^{(g)} - 1/2) \right] \quad (51)$$

$$\begin{aligned}
&= M^{-3/2} \sum_{1 \leq u < v \leq \ell} n_u n_v \left[\int F_g(x - \theta_{ug} M^{-1/2}) f_g(x - \theta_{vg} M^{-1/2}) dx - \int F_g(x) f_g(x) dx \right] \\
&\quad + M^{-3/2} \sum_{\ell \leq u < v \leq k} n_u n_v \left[\int F_g(x - \theta_{vg} M^{-1/2}) f_g(x - \theta_{ug} M^{-1/2}) dx - \int F_g(x) f_g(x) dx \right] \\
&= M^{-2} \sum_{1 \leq u < v \leq \ell} n_u n_v (\theta_{vg} - \theta_{ug}) \left[\int \frac{F_g(x + (\theta_{vg} - \theta_{ug}) M^{-1/2}) - F_g(x)}{(\theta_{vg} - \theta_{ug}) M^{-1/2}} f_g(x) dx \right] \\
&\quad + M^{-2} \sum_{\ell \leq u < v \leq k} n_u n_v (\theta_{ug} - \theta_{vg}) \left[\int \frac{F_g(x + (\theta_{ug} - \theta_{vg}) M^{-1/2}) - F_g(x)}{(\theta_{ug} - \theta_{vg}) M^{-1/2}} f_g(x) dx \right]
\end{aligned}$$

elde edilir. $M \rightarrow \infty$ için,

$$\mu_A = \left[\sum_{1 \leq u < v \leq \ell} \lambda_{u1} \lambda_{v1} (\theta_{vg} - \theta_{ug}) + \sum_{\ell \leq u < v \leq k} \lambda_{u2} \lambda_{v2} (\theta_{ug} - \theta_{vg}) \right] \int f_g^2(x) dx \quad (52)$$

dır[Bkz 12]. Teoremin geri kalan kısmı Lehmann'ın [Bkz 13] (iii.) sonucu kullanılarak ispatlandı. Dolayısıyla önce Lehmann'ın (iii.) sonucunu görelim:

Varsayalım $F, G, H, \dots$ N 'e bağlı dağılımlar olsun. $\theta_N^{(i)} = E[U^{(i)}]$ ve Σ_N ise $(F_N, G_N, H_N, \dots)$ için hesaplanan Σ kovaryans matrisi olsun. Eğer Σ_N , $N \rightarrow \infty$ için singular olmayan Σ kovaryans matrisine yakınsıyor ise

$$N^{1/2}(U^{(1)} - \theta_N^{(1)}), \dots, N^{1/2}(U^{(r)} - \theta_N^{(r)})$$

de beklenen değeri 0 ve kovaryans matrisi Σ olan normal dağılıma yakınsar.

Şimdi $M^{1/2}(U_1 - \gamma_{1M}), \dots, M^{1/2}(U_t - \gamma_{tM})$ ve $M^{1/2}(U_1 - \gamma_{1M}), \dots, M^{1/2}(U_s - \gamma_{sM})$ 'lerin bileşik limit dağılımlarını bulacağız. Burada $U_1, \dots, U_t, \dots, U_{t+s}$ 'ler k -grup için U istatistikleridir. Biz bu sonucu $t=\ell(\ell-1)p/2$, $s=(k-\ell)(k-\ell+1)p/2$ değerleri için teorem 3 deki gibi $\Omega_{uv}^{(g)}$ ve $\Omega_{vu}^{(g)}$ istatistiklerine uygulayacağız. Yukarıda belirtilen Lehmann'ın sonucuna göre eğer, yerel alternatif F_{jN} altında, $\Omega_{uv}^{(g)}$ ve $\Omega_{vu}^{(g)}$ 'ların varyans ve kovaryans limit değerleri null hipotezi altında hesaplanan varyans ve kovaryans limit değerlerine yakınsıyor ise $\Omega_{uv}^{(g)}$ ve $\Omega_{vu}^{(g)}$ 'ların bileşik limit dağılımları null hipotezi altındakiyle aynıdır. Randles ve Wolfe'un 3.6.10 denkleminde göre $\Omega_{uv}^{(g)}$ ve $\Omega_{vu}^{(g)}$ 'ların

varyanslarının ve kovaryanslarının limit değerleri sırasıyla (53-55) 'de verilmiştir.

$$Var(\Omega_{uv}^{(g)}) = (p_2 - p_1^2) / \lambda_{u_1} + (p_3 - p_1^2) / \lambda_{v_1} \quad (53)$$

$$Var(\Omega_{vu}^{(g)}) = (p_5 - p_4^2) / \lambda_{u_2} + (p_6 - p_4^2) / \lambda_{v_2} \quad (54)$$

$$Cov(\Omega_{uv}^{(g)}, \Omega_{vu}^{(g)}) = -2(p_3 - p_1^2) \lambda_\ell \quad (55)$$

Burada,

$$\begin{aligned} p_1 &= P(X_{1vg} > X_{1ug}), \\ p_2 &= P(X_{1vg} > X_{1ug} \text{ ve } X_{2vg} > X_{1ug}), \\ p_3 &= P(X_{1vg} > X_{1ug} \text{ ve } X_{1vg} > X_{2ug}). \end{aligned} \quad (56)$$

olup tepe noktasından önceki γ_{uv} parametrelerini,

$$\begin{aligned} p_4 &= P(X_{1ug} > X_{1vg}), \\ p_5 &= P(X_{1ug} > X_{1vg} \text{ ve } X_{2ug} > X_{1vg}), \\ p_6 &= P(X_{1ug} > X_{1vg} \text{ ve } X_{1ug} > X_{2vg}). \end{aligned} \quad (57)$$

ise tepe noktasından sonraki γ_{vu} parametrelerini tanımlar.

F_{jN} altında; $j=1, \dots, k$

$$\begin{aligned} p_1 &= \int F_g \left(x + (\theta_{vg} - \theta_{ug}) M^{-1/2} \right) f_g(x) dx, \\ p_2 &= \int \left[1 - F_g \left(x - (\theta_{vg} - \theta_{ug}) M^{-1/2} \right) \right]^2 f_g(x) dx, \\ p_3 &= \int F_g^2 \left(x + (\theta_{vg} - \theta_{ug}) M^{-1/2} \right) f_g(x) dx, \\ p_4 &= \int F_g \left(x + (\theta_{ug} - \theta_{vg}) M^{-1/2} \right) f_g(x) dx, \\ p_5 &= \int \left[1 - F_g \left(x + (\theta_{ug} - \theta_{vg}) M^{-1/2} \right) \right]^2 f_g(x) dx, \\ p_6 &= \int F_g^2 \left(x + (\theta_{ug} - \theta_{vg}) M^{-1/2} \right) f_g(x) dx \end{aligned} \quad (58)$$

olur. "Dominated Convergence teoremine" göre $(p_1, p_4); (p_2, p_5); (p_3, p_6)$, sırasıyla $1/2$; $1/3$; ve $1/3$; 'e yakınsar. Kovaryansın limitini tepe noktasından önce $u=s$, $v \neq t$ ve $g \neq h$ durumu için bulalım.

$$Cov(\Omega_{uv}^{(g)}, \Omega_{ut}^{(h)}) = (p_7 - p_1 p_8) / \lambda_{u_1} \quad (59)$$

Burada $p_8 = P(X_{1th} > X_{1uh})$, ve $p_7 = P(X_{1vg} > X_{1ug} \text{ ve } X_{1th} > X_{1uh})$ dir. F_{jN} altında; $j=1, \dots, k$

$$p_7 = \int \int \left[1 - F_g(x - \theta_{vg} M^{-1/2}) \right] \left[1 - F_h(y - \theta_{tg} M^{-1/2}) \right] \times dF_{gh}(x - \theta_{vg} M^{-1/2}, y - \theta_{tg} M^{-1/2}) \quad (60)$$

değişken dönüşümünden sonra "Dominated Convergence teoremini" uygularsak $M \rightarrow \infty$ için,

$$\int \int [1 - F_g(x)] [1 - F_h(y)] * dF_{gh}(x, y) = \pi_{gh}/2 \quad (61)$$

olur. Öyleyse, p_8 ve p_1 'in $1/2$ yakınsadığından, (23) kullanılarak $Cov(\Omega_{uv}^{(g)}, \Omega_{ut}^{(h)}) = \rho_{gh}/12\lambda_{u1}$ olduğu görülür. Diğer durumlarda aynı şekilde gösterilebilir. ■

3.6 Bileşen Yaklaşımıyla Normal Teori İstatistiği

Puri [Bkz 14] Jonckheere test istatistiğine karşı tek boyutlu durum için normal teori yöntemleri kullanarak yeni bir istatistik tanımladı. Dietz E.J [Bkz 1] bu istatistiği çok boyutlu duruma bilişen yöntemi kullanarak genelledi. Douglas A.W ve Gregory A.Mack [Bkz 5] umbrella testi tek boyutlu durum için normal teori yöntemlerini kullandı. Bizde burada çok boyutlu umbrella testi için normal teori yöntemlerini kullanarak aşağıdaki şekilde tanımladık.

$$T_g = \sum_{1 \leq u < v \leq \ell} n_u n_v (\bar{x}_{vg} - \bar{x}_{ug}) + \sum_{\ell \leq u < v \leq k} n_u n_v (\bar{x}_{ug} - \bar{x}_{vg}) \quad (62)$$

Burada $\bar{x}_{jg} = \sum_{i=1}^{n_j} X_{ijg}/n_j$ $j=1, \dots, k$ dir.

3.7 Null Hipotezi

H_0 altında X_{ij} 'ler hepsi bağımsız ve aynı dağılımdan gelsinler. Varsayalım X_{ij} 'lerin kovaryans matrisi Σ olsun. Kolayca gösterilebilir ki $M^{-3/2}T = M^{-3/2}(T_1, \dots, T_p)$ 'nin ortalaması 0 ve kovaryans matrisi Σ_T dir.

$$\Sigma_T = \left[\left(1 - \sum_{j=1}^{\ell} \left(\frac{n_{j1}}{M} \right)^3 \right) + \left(1 - \sum_{j=\ell}^k \left(\frac{n_{j2}}{M} \right)^3 \right) + 2 \left(\frac{n_{\ell}}{M} \right) \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^k \left(\frac{n_{j1}}{M} \right) \left(\frac{n_{j2}}{M} \right) \right] \Sigma/3 \quad (63)$$

Theorem 5 H_0 altında $M^{-3/2}T$ 'nin asimtotik dağılımı beklenen değeri 0 ve kovaryans

matriksi

$$\Sigma_T = \left[\left(1 - \sum_{j=1}^{\ell} \lambda_{j1}^3 \right) + \left(1 - \sum_{j=\ell}^k \lambda_{j2}^3 \right) + 2\lambda_{\ell} \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^k \lambda_{i1} \lambda_{j2} \right] \Sigma/3 \quad (64)$$

olan p boyutlu normal dağılımdır. Burada,

$$\begin{aligned} \lambda_{j1} &= \lim_{M \rightarrow \infty} (n_{j1}/M), \quad j_1 = 1, \dots, \ell-1 \\ \lambda_{j2} &= \lim_{M \rightarrow \infty} (n_{j2}/M), \quad j_2 = \ell+1, \dots, k \\ \lambda_{\ell} &= \lim_{M \rightarrow \infty} n_{\ell}/M \end{aligned}$$

dir.

İspat:. Randles ve Wolfe 'un [Bkz 11] teorem 3.6.9'da kullandığı U istatistiğini $\bar{X}_{jg}$ ($j=1, \dots, k$; $g=1, \dots, p$) olarak seçelim. Daha sonra teorem 3 'deki yöntem izlendiğinde istenen sonuca ulaşılacaktır. ■

3.8 Alternatif Hipotez Altında Dağılım

Nisbi etkinlik değerini elde edebilmek için yerel alternatif dizisi tanımlanmalıdır. $F_{jN}(\mathbf{X}) = F(\mathbf{X} - \theta_j M^{-1/2})$ olsun. Burada F ortalaması 0, kesin sürekli dağılım fonksiyonu ve f_g , F 'in g . bileşeninin marjinal dağılım fonksiyonu olsun. Varsayalım $\int f_g^2(x) dx < \infty$, $g=1, \dots, p$ için.

Theorem 6 *Yerel alternatif altında $M^{-3/2}\mathbf{T}$ asimtotik dağılımı p boyutlu normal dağılım olup, ortalaması*

$$\mu_T = \left[\sum_{1 \leq u < v \leq \ell} \lambda_{u1} \lambda_{v1} (\theta_{vg} - \theta_{ug}) + \sum_{\ell \leq u < v \leq k} \lambda_{u2} \lambda_{v2} (\theta_{ug} - \theta_{vg}) \right] \quad (65)$$

ve kovaryansı matriksi (64) ile verilen Σ_T dir.

İspat:. Randles ve Wolfe 'un [Bkz 11] teorem 3.6.9'da kullandığı U istatistiğini $\bar{X}_{jg}$ ($j=1, \dots, k$; $g=1, \dots, p$) olarak seçelim. Daha sonra teorem 4 'deki yöntem izlendiğinde istenen sonuca ulaşılacaktır. ■

3.9 Çok Boyutlu Test İstatistikleri

Toplam İstatistikler

H_a 'nın lehinde H_0 hipotezimizi test etmek için, $1'A = \sum_{g=1}^p A_g$ ve $1'T = \sum_{g=1}^p T_g$ istatistikleri hipotez testinde kullanılabilir. Burada $1'$, $p \times 1$ boyutlu 1lerden oluşan bir vektördür. Teorem 3' den H_0 altında Σ_0 'ın Σ_A 'ya yakınsadığı biliyoruz. Öyleyse;

$$A^{(1)} = \frac{M^{-3/2} 1'A}{(1'\Sigma_0 1)^{1/2}} \quad (66)$$

asimptotik dağılımı standart normaldir. $1'T$ 'ye bağlı bir test kurabilmek için öncelikle Σ_T 'yi tahmin etmeliyiz. Σ_T 'nin tahmin edicisi

$$\hat{\Sigma}_T = \left[\left(1 - \sum_{j=1}^{\ell} \lambda_{j1}^3 \right) + \left(1 - \sum_{j=\ell}^k \lambda_{j2}^3 \right) + 2\lambda_{\ell} \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^k \lambda_{i1} \lambda_{j2} \right] \times \sum_{j=1}^k \left[\sum_{i=1}^{n_j} X_{ij} X'_{ij} - n_j \bar{X}_j \bar{X}'_j \right] / 3(N - k) \quad (67)$$

olsun. Burada $\bar{X}_j = \sum_{i=1}^{n_j} X_{ij} / n_j$ olup $\hat{\Sigma}_T$, Σ_T 'nin tutarlı bir tahmin edicisidir [Bkz 15]. Bu takdirde;

$$T^{(1)} = \frac{M^{-3/2} 1'T}{(1'\hat{\Sigma}_T 1)^{1/2}} \quad (68)$$

H_0 altında asimptotik dağılımı standart normaldir.

3.10 Asimptotik Nisbi Etkinlik

Bu bölümde $A^{(1)}$ ve $T^{(1)}$ istatistiklerinin F_{jN} alternatif dizisi için Pitman asimptotik nisbi etkinlik değerlerini kıyaslayacağız.

Theorem 7 Asimptotik olarak $A^{(1)}$ ve $T^{(1)}$ 'nin asimptotik nisbi etkinlik değeri,

$$e_{A,T} = \frac{(1'\mu_A)^2 (1'\Sigma 1)}{(1'\mu_T)^2 (1'P 1/12)} \quad (69)$$

dır. Burada $P = (\rho_{gh})$ dir.

Eğer $p=1$ ise,

$$e_{A,T} = 12\sigma^2 \left[\int f^2(x) dx \right]^2 \text{ olur.} \quad (70)$$

Bu Mann-Whitney'in t-testine karşı asimptotik nisbi etkinliğidir [Bkz 5].

İspat:. $A^{(1)}$ ve $T^{(1)}$ 'nin asimtotik nisbi etkinliği, bunların alternatif altındaki dağılımlarının α düzeyinde örnek büyüklüklerinin limit oranlarıdır.

F_{jN} altında $A_N = M^{-3/2}1'A$, ortalaması $1'\mu_A$ ve varyansı $1'\Sigma_A1$ olan asimtotik normal dağılımdır. $T_N = M^{-3/2}1'T$ de ortalaması $1'\mu_T$ ve varyansı $1'\Sigma_T1$ olan asimtotik normal dağılıma sahiptir. Varsayalım $\{N_l\}$ ve $\{N_l^*\}$ α düzeyinde $\{A_{N_l}\}$ ve $\{T_{N_l^*}\}$ test etmek için artan bir dizi olsun. Öyleki $\{A_{N_l}\}$ ve $\{T_{N_l^*}\}$ alternatif altında ki sırasıyla $\theta_j N_l^{-1/2}$ ve $\theta_j^* N_l^{*-1/2}$ $j=1,...,k$ limit kuvvetleri birbirine eşittir. Eldeki bu iki dizinin testleri aynı limit kuvvetlerine eşit olması için

$$(1'\mu_A)(1'\Sigma_A1)^{-1/2} = (1'\mu_T)(1'\Sigma_T1)^{-1/2} \quad (71)$$

sağlanmalıdır. Eğer bu iki test alternatifte aynı ise,

$$\theta_j^* = \theta_j (N_l/N_l^*)^{1/2} \quad (72)$$

olur. ■

Theorem 8 *Eğer F ; p boyutlu normal dağılıma sahip, kovaryans matrisi*

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & & \rho \\ & 1 & \\ \rho & & 1 \end{bmatrix} \quad (73)$$

ise

$$e_{A,T} = \frac{3[1 + (p-1)\rho]}{\pi[1 + (p-1)\rho^*]} \quad (74)$$

olur. Burada $\rho^* = (6/\pi) \sin^{-1}(\rho/2)$ dir.

İspat:. Teoremdeki varsayımlar altında,

$$\int f_g^2(x)dx = (2\pi^{1/2}\sigma)^{-1} \quad g=1,...,p \quad (75)$$

$$\rho_{gh} = \rho^* \quad g,h=1,...,p \quad (76)$$

olur[Bkz 16].

Eğer $p > 2$ ise; ρ , $-(p-1)^{-1}$ minimum değerlerine yakınsarken (74)'de tanımlı $e_{A,T}$ sifıra yakınsar. Aşağıda bazı ρ değerleri için etkinlik değerleri gösterilmiştir.

p	ρ	Nisbi Etkinlik
3	-0,47	0,610
4	-0,31	0,617
5	-0,23	0,639

(77)

$p=2$ için ρ -1 değerine yaklaşırken asimtotik nisbi etkinliğide minimum değerine yani $(0,75)^{1/2}$ değerine yakınsar. $e_{A,T}$, $-(p-1)^{-1} < \rho < 1$ değerleri için birden küçüktür. $e_{A,T}$ $2 \leq p \leq 5$ için aşağıda belirtilen maksimum değerleri almaktadır. [Bkz 1]

p	ρ	Nisbi Etkinlik
2	0,52	0,966
3	0,47	0,972
4	0,44	0,976
5	0,41	0,978

(78)

4 SİMÜLASYON YÖNTEMİYLE YAKLAŞIK KRİTİK DEĞERLERİN BULUNMASI

Yapılan teorik çalışmaların uygulamada ki sonuçlarını görmek için simülasyon çalışması gerçekleştirdik. A_g 'nin null dağılımı altında değişik örnek büyüklükleri için $\alpha=0,1$; $\alpha=0,05$ ve $\alpha=0,01$ 1.tip hata oranlarına ait kritik değerler bulundu. Tablo 4.1 'de, grup sayısı $k=5$, tepe noktası $\ell=3$, $n_1=\dots=n_5$ koşulları altında örnek büyüklükleri $n=10, 20, 30$ ve 50 için üç boyutlu normal dağılımdan türetilen 3000 rastgele örnek üzerindeki sonuçları gösterilmektedir. Burada görülmüştür; küçük örnek büyüklükleri için dahi bulduğumuz kritik değerler seçilen α değerlerine çok yakın değerler vermiştir.

Tablo 4.1 Örnek büyüklükleri eşit gruplarda α hata paylarına ait kritik değerler

$k=5$ ve $\ell=3$			
Grup Büyüklüğü	$\alpha = 0, 10$	$\alpha = 0, 05$	$\alpha = 0, 01$
$n_1 = \dots = n_5$	$P_0(A_g \geq 1, 453)$	$P_0(A_g \geq 1, 812)$	$P_0(A_g \geq 2, 652)$
10	0,107	0,055	0,010
20	0,102	0,057	0,010
30	0,103	0,061	0,010
50	0,095	0,047	0,010

Ayrıca A_g 'nin uç gözlemlerdeki davranışıda; Tablo 4.2 de grup sayısı $k=5$, tepe noktası $\ell=3$, $n_1=\dots=n_5$ koşulları altında örnek büyüklükleri $n=10, 20, 30$ ve 50 için üç boyutlu normal dağılımdan türetilen 3000 rastgele örnek üzerindeki sonuçları gösterilmektedir. Burada her gruba örnek büyüklüğünün %10 'u kadar uç gözlemler yerleştirilmiştir. Uç gözlemler; her bir bileşene, standart sapmasının 10 katı kadar kendisine ilave edilmesi suretiyle elde edildi. Uç gözlemler, A_g test istatistiğinin yapısını ve kritik

değerlerimizi değiştirmedeği görüldü.

Tablo 4.2 Uç gözlemlerle gerçekleştirilen α hata paylarına ait kritik değerler

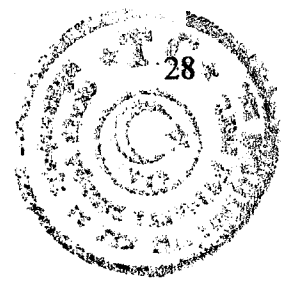
k=5 ve $\ell=3$			
Grup Büyüklüğü	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
$n_1 = \dots = n_5$	$P_0(A_g \geq 1,453)$	$P_0(A_g \geq 1,812)$	$P_0(A_g \geq 2,652)$
10	0,093	0,056	0,010
20	0,099	0,057	0,009
30	0,102	0,050	0,010
50	0,094	0,052	0,010

(79)

5 SONUÇLAR ve GELECEĞE YÖNELİK ÇALIŞMALAR

Çok boyutlu durum için umbrella testinin koşullu varyans ve kovaryans matriksi bulundu. Null hipotezi ve yerel alternatif altında asimptotik dağılımları incelendi. Yoğunluk fonksiyonun bilindiği durumda normal teori yöntemleri, çok boyutlu durum için umbrella testine genellendi.

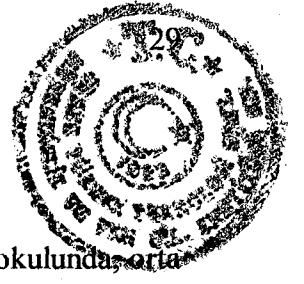
Çok boyutlu umbrella testi için geliştirdiğimiz A_g test istatistiğinin asimptotik nisbi etkinliğine bakıldığında bu değerlerin bire olup parametrik olmayan bu yöntemin iyi sonuçlar verdiğinin bir göstergesi olarak algılanabilir. Yapılan simülasyon çalışmaları değişik grup sayılarında (k) ve farklı boyutlarda (p) geliştirilebilir. Ayrıca biz bu çalışmada $M=N_1=N_2$ durumunu gruptaki gözlem sayılarını eşit alarak gerçekleştirdik. Pekala M sayısı tepe noktasından önceki ve sonraki grupların toplamalarını eşit kılacak şekilde de seçilebilir. Son olarak, tezin başında da belirttiğimiz gibi A_g test istatistiğini tepe noktasının bilindiği durumda inceledik. Ancak tepe noktasının bilinmediği durum içinde bu test genelleştirilebilir.



KAYNAKLAR

- [1] DIETZ E. JACQUELIN Commun. Statistical Theory Meth. “ Multivariate Generalizations of Jonckheere’s Test For Ordered Alternatives ”. (1989). 3763-3783
- [2] KRUSKAL WILLIAM H., and WALLIS W.ALLEN (1952) Journal of American Statistical Association ,47, 583-621
- [3] JONCKHEERE A.R (1954) “A Distribution-Free K-Sample Test Against Ordered Alternatives,” Biometrika, 41, 133-145
- [4] TERPSTRA T.J (1952) Indagationes Mathematicae 14, 327-333
- [5] GREGORY A. MACK and DOUGLES A.WOLFE (1981) Journal of American Statistical Association 76V., 373, 175-181
- [6] MANN, H.B and WHITNEY, D.R (1947) Ann. Math. Statistic, 18, 50-60
- [7] ARCHAMBAULT W.A.T, JR, MACK, GREGORY A. MACK and DOUGLES A.WOLFE (1978) Canadian Journal of Statistics, 5, 195-207
- [8] DOUGLES A.WOLFE (2002) Nonparametric Statistical Methods, 2nd Edi
- [9] LEHMANN, E.L (1975) Nonparametrics :Statistical Methods Based on Ranks, p. 370
- [10] GIBBONS, J.D (1985) Nonparametric Statistical Inference (2nd ed.) 238-240
- [11] RANGLES R.II, and WOLFE D.A, (1979) Introduction to The Theory of Nonparametric Statistic p.107
- [12] MEHRA K.L and SARANGIL J (1967) “Asymptotic Efficiency of Certain Rank Tests For Comparative Experiments” Ann. Math Stat, 38, 90-107
- [13] LEHMANN, E.L (1963) “Robust Estimation in Analysis of Variance” Ann. Math Stat, 34, 957-966
- [14] PURI, MADAN L. (1965) “Some Distribution Free K Sample Rank Test of Homogeneity Against Ordered Alternatives, “ Communication on Pure and Applied Mathematics, 18, 51-63
- [15] MORISON D.F (1967), Multivariate Statistical Methods New York, McGraw-Hill
- [16] KENDALL, M. G (1975). Rank Correlation Methods (4th ed.) London

ÖZGEÇMİŞ



Ahmet NASIRLI 1976 İstanbul’ da doğdu. İlkokulu İskender Paşa İlkokulunda, orta okulu ise Oruç Gazi Ortaokulu’nda tamamladı. Lise eğitimini ise 1993 yılında Pertevniyal Lisesi’nde bitirdi. 1993 yılında girdiği İstanbul Üniversitesi’nden 1995 yılında ayrılıp, aynı yıl Mimar Sinan Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümüne başlayıp, 1999 yılında mezun oldu. 1999 güz döneminde Gebze Yüksek Teknoloji Enstitüsü Matematik Bölümünde yüksek lisans eğitimine başladı. 1999 yılında Milli Eğitim Bakanlığına yapmış olduğu öğretmenlik başvurusu sonucu Gülsuyu İ.Ö. Okulu’na Matematik öğretmeni olarak atandı. Halen Bnb. Necati Bey İ.Ö. Okulu’nda Matematik öğretmeni olarak çalışmaktadır.

