

Sosyolojide Nicel Araştırma Teknikleri

Lecture 9

Ali Mertcan KOSE Ph.D.

amertcankose@ticaret.edu.tr

İstanbul Ticaret Üniversitesi



İSTANBUL TİCARET
ÜNİVERSİTESİ

Bağımsız İkiden fazla Gruplara İlişkin Hipotez Testleri

Amaç

2'den fazla bağımsız grubun sayısal bir değişken bakımından karşılaştırılması

- *Parametrik test:* Tek Yönlü Varyans Analizi
- *Parametrik olmayan test:* Kruskal Wallis testi

Tek Yönlü Varyans Analizi

Parametrik test varsayımları sağlandığında ölçümle belirtilen bir değişken yönünden ikiden fazla bağımsız grubun ortalamaları arasında fark olup olmadığını test etmek için kullanılır. İki ortalama arasındaki farkın anlamlılık testi için gerekli varsayımlar varyans analizi için de geçerlidir.

Tek Yönlü Varyans Analizi

Varsayımlar

- Karşılaştırılacak gruplar normal dağılım göstermeli
- Grupların varyansları homojen olmalı
- Gruplar birbirinden bağımsız olmalı

Hipotezler:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

H_a : En az bir μ_i farklıdır.

Tek Yönlü Varyans Analizi

1	2	3	...	k
x ₁₁	x ₁₂	x ₁₃	...	x _{1k}
x ₂₁	x ₂₂	x ₂₃	...	x _{2k}
x ₃₁	x ₃₂	x ₃₃	...	x _{3k}
...
x _{n11}	x _{n22}	x _{n33}	...	x _{nk}
T ₁	T ₂	T ₃		T _k
x _{ort1}	x _{ort2}	X _{ort3}		X _{ortk}

Tek Yönlü Varyans Analizi

$T_j = \sum_{i=1}^{n_j} x_{ij}$ = j. sütunun toplamı

$\bar{x}_{..j} = \frac{T_{..j}}{n_j}$ = j.sütunun ortalaması

$T_{..} = \sum_{i=1}^k T_{..j} = \sum_{j=1}^k \sum_{i=1}^k x_{ij}$ = Bütün gözlemlerin toplamı

$\bar{x}_{..} = \frac{T_{..}}{N}$

$N = \sum_{j=1}^k n_j$

Tek Yönlü Varyans Analizi

Genel Kareler Toplamı

$$GNKT = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_{..})^2 = \sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij} - \frac{T^2}{N}$$

Grup İçi Kareler Toplamı

$$GIKT = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 = \sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij}^2 - \frac{T_j^2}{n_j}$$

Gruplar Arası Kareler Toplamı

$$GAKT = \sum_{j=1}^k \sum_{i=1}^{n_j} n_j (x_{ij} - \bar{x}_j)^2 = \sum_{j=1}^k \sum_{i=1}^{n_j} \frac{T_j}{n_j} - \frac{T^2}{N}$$

Tek Yönlü Varyans Analizi

Gruplar Arası Kareler Ortalaması

$$\text{GAKO} = \frac{\text{GAKT}}{k-1}$$

Gruplar İçi Kareler Ortalaması

$$\text{GIKO} = \frac{\text{GIKT}}{n-k}$$

$$F_h = \frac{\text{GAKO}}{\text{GIKO}}$$

Tek Yönlü Varyans Analizi

DK	KT	SD	KO	F
Gruplar Arası	GAKT	k-1	GAKO	GAKO/GIKO
Grup İçi	GIKT	N-k	GIKO	
Genel	GNKT	N-1		

Varyans Analizi Sonucu Anlamlı Olduğunda Farklı Grupların Belirlenmesi

Varyans analizi sonucunda gruplar arasında fark yoksa işlemler sona erer. Ancak, gruplar arasında fark varsa, farklılığın hangi grup ya da gruplar arasında olduğu farklı yöntemlerle araştırılabilir. Bu yöntemlere post-hoc testleri denir. Bu yöntemlerde en çok kullanılanlar;

- LSD
- Tukey
- Benferroni
- Sidak
- Dunnett's C
- Dunnett's T3

Tek Yönlü Varyans Analizi

LSD Testi Örneklem genişlikleri eşit olduğunda ($n_1 = n_2 = n_3 = \dots = n_k = n$)

$$|\bar{x}_i - \bar{x}_j| \geq t \sqrt{\frac{2(GIKO)}{n}} \quad p < 0.05$$

Örneklem genişlikleri eşit olmadığındada ($n_1 \neq n_2 \neq n_3 \neq \dots \neq n_k$)

$$|\bar{x}_i - \bar{x}_j| \geq t \sqrt{GIKO \frac{1}{n_i} + \frac{1}{n_j}} \quad p < 0.05$$

Örnek 1

Adolesan dönemindeki 90 kız, yaş gruplarına göre (11-14,15-18,19-24) 3 gruba ayrılmıştır. Günlük kilo başına tüketikleri kaloriler hesaplanmıştır. Yaş gruplarına göre tüketilen kaloriler bakımından farklılık var mıdır?

Örnek 1

	11-14	15-18	19-24
	42.45	39.98	43.30
	46.81	45.29	42.85
	45.62	33.08	32.43
	53.82	38.60	46.81

	50.68	37.57	35.18
Toplam	1380.76	1193.82	1105.72
Ortalama	46.02	39.79	36.86

Örnek 1

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

H_a : En az bir μ_i farklıdır.

$$\begin{aligned} \text{GNKT} &= \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_{..})^2 = \sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij} - \frac{T_{..}^2}{N} = 42.45^2 + \\ &46.81^2 + \dots + 35.18^2 - \frac{3680.30^2}{90} \\ &= 154138.01 - 150495.65 = 3642.36 \end{aligned}$$

$$\begin{aligned} \text{GIKT} &= \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_{.j})^2 = \sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij}^2 - \frac{T_{.j}^2}{n_j} = 42.45^2 + \\ &46.81^2 + \dots + 35.18^2 - \left[\frac{1380.76^2}{30} + \frac{1193.82^2}{30} + \frac{1105.72^2}{30} \right] \\ &= 2327.31 \end{aligned}$$

$$\text{GAKT} = \text{GNKT} - \text{GIKT} = 3642.36 - 2327.31 = 1315.05$$

Örnek 1

$$GIKO = GIKT / (90-3) = 2327.31 / 87 = 26.75$$

$$GAKO = GAKT / (3-1) = 1315.05 / 2 = 657.53$$

$$F = GAKO / GIKO = 657.53 / 26.75 = 24.58$$

DK	KT	SD	KO	F
Gruplar Arası	1315.05	2	657.53	24.58
Grup İçi	2327.31	87	26.75	
Genel	3642.36	89		

Grup ortalamaları arasında anlamlı bir farklılık vardır.

Örnek 1

Gruplardaki kişi sayıları birbirine eşit olduğu için

$$|\bar{x}_i - \bar{x}_j| \geq t \sqrt{\frac{2(GIKO)}{n}}$$

$$|46.02 - 39.79|$$

$$H_0 : \mu_1 = \mu_2 \quad 1.98 \sqrt{\frac{2(26.75)}{30}} = 2.64 \rightarrow 6.23 > 2.64 \quad H_0 \text{ red}$$

$$H_0 : \mu_1 = \mu_3 \quad 1.98 \sqrt{\frac{2(26.75)}{30}} = 2.64 \rightarrow 9.16 > 2.64 \quad H_0 \text{ red}$$

$$H_0 : \mu_2 = \mu_3 \quad 1.98 \sqrt{\frac{2(26.75)}{30}} = 2.64 \rightarrow 2.93 > 2.64 \quad H_0 \text{ red}$$

Kruskal Wallis Testi

Tek yönlü varyans analizinin parametrik olmayan karşılığıdır. Veriler ölçümlle belirtildiği halde parametrik test varsayımları sağlanmıyorsa (gözlem sayısı az ya da gruplar normal dağılmıyor ise) Kruskal-Wallis testi kullanılır.

Kruskal Wallis Testi

Bu testte ve parametrik olmayan diğer testlerde, grplara ait ölçüm-lerin karşılaştırılmasında aritmetik ortalama yerine ortanca değer esas alınır.

Mann-Whitney U testinin 3 veya daha çok grup olduğu duruma genişletilmesidir.

Kruskal Wallis Testi

Testin aşamaları şu şekilde gerçekleşir:

- ① k grubun n_1, n_2, \dots, n_k gözlemleri tek bir değişken altında küçükten büyüğe sıralanır. Tüm gözlemlere sıra numarası verilir.
- ② k grubun sıra numaraları ayrı ayrı toplanır(R_j).
- ③ Test istatistiği

$$KW = \frac{12}{n(n+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(n+1)$$

şeklinde hesaplanır.

R_j = j.gruptaki sıra sayıları toplamı

n_j = j.gruptaki gözlem sayısı

Kruskal Wallis Testi

- ④ Üç grup olduğunda ve her bir grupta beş ve daha az gözlem olduğunda hesaplanan KW istatistiği, özel tablolar kullanılarak karşılaştırılır. Bir ya da daha fazla grupta beşten fazla gözlem olduğunda ise KW, $k-1$ serbestlik dereceli χ^2 tablo değeriyile karşılaştırılır.

Kruskal Wallis Testi

Test Sonucu Anlamlı olduğunda farklı grupların belirlenmesi

ANOVA'da olduğu gibi bu Kruskal-Wallis testi de tüm gruplar arasında anlamlı bir farklılık olduğunu vermez. Bunun için çoklu karşılaştırma yapmak gereklidir.

$$|\bar{R}_i - \bar{R}_j| > t \sqrt{\frac{n(n+1)}{12} \frac{n-1-KW}{n-k} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \rightarrow p < 0.05$$

Örnek 2

Üniversite öğrencilerinin çay içme miktarına göre hemoglobin düzeylerinin değişip değişmediği incelenmek istenmektedir. Bu amaçla 13 kişi “yemekten 1 saat önce veya sonra çay içenler”, “yemekten 30 dakika önce ya da sonra çay içenler” ve “yemekle birlikte çay içenler” olmak üzere üç gruba ayrılmışlardır ve hemoglobin düzeyleri ölçülmüştür. Buna göre hemoglabin düzeyi çayın içilme zamanına göre değişmekte midir?

Hipotezler:

H_0 : Kitle dağılımları benzerdir.

H_1 : En az kitle dağılımı diğerinden farklıdır.

Örnek 2

I	II	III
13.5(9)	12.9(6.5)	10.9(1)
13.8(10)	12.5(5)	11.5(4)
15.5(13)	13(8)	11.2(3)
14(11)	12.9(6.5)	11(2)
14.7(12)		
Rj 55	26	10

I.Yemekten 1 saat

II.Yemekten 30 dakika önce ya da sonra çay içenler

III.Yemekle birlikte çay içenler

Örnek 2

$$KW = \frac{12}{n(n+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(n+1)$$

$$\frac{12}{13(13+1)} \left[\frac{55^2}{5} + \frac{26^2}{4} + \frac{10^2}{4} \right] - 3(13+1) = 10.68$$

$$KW_{(5,4,4;0.05)} = 5.657 < KW = 10.68$$

p<0.05, H_0 red

Örnek 2

Critical Values of the Kruskal-Wallis H Distribution
Taken from Zar, 1984 Table B.12

n_1	n_2	n_3	$H = 5.12$	5.05	5.02	5.01	5.005	5.008	5.001	n_1	n_2	n_3	$H = 5.12$	5.05	5.02	5.01	5.005	5.008	5.001		
2	2	2	4.575							7	7	7	4.598	5.019	5.332	5.378	5.473	10.536	11.210		
3	2	2		4.590	4.714					8	8	8	4.599	5.005	5.355	5.465	5.495	10.505	11.170		
3	3	1		4.575	5.153					2	2	2	4.597	5.023	5.378	5.473	5.495	10.536	11.210		
3	3	2		4.596	5.361	4.850				2	2	2	4.607	5.187	5.507	5.657	5.737	10.536	11.210		
3	3	3		4.625	5.690	4.849	4.700			3	3	1	4.640	5.114	5.434	5.584	5.654	10.536	11.210		
4	2	2		4.495	5.000	5.000				3	2	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
4	2	3		4.433	5.233	5.030				3	2	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
4	3	1		4.504	5.298	5.000				3	2	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
4	3	2		4.520	5.304	5.164	5.000			3	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
4	3	3		4.700	5.791	5.004	5.795	5.710	5.010	4	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
4	4	1		4.157	5.967	5.657	5.657	5.657	5.010	4	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
4	4	2		4.555	5.455	6.020	7.037	7.032	7.055	3	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
4	4	3		4.645	5.595	6.712	7.154	7.098	6.27	8.999	3	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
4	4	4		4.668	5.640	7.054	7.054	8.200	8.200	8.200	3	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	2	1		4.575	5.930					4	4	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
5	2	2		4.375	5.150	6.020	6.523			4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
5	3	1		4.315	4.960	5.044	5.044			4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
5	3	2		4.605	5.295	6.150	6.099	7.152		4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
5	3	3		4.625	5.302	6.077	7.077	8.040	8.040	8.040	4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	4	1		3.997	4.985	6.420	6.955	7.364		4	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
5	4	2		4.561	5.273	7.205	7.573	8.114	8.091	8.091	4	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	4	3		4.549	5.659	6.675	7.445	7.927	8.041	8.041	4	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	4	4		4.569	5.127	7.205	7.573	8.114	8.091	8.091	4	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	5	1		4.139	5.127	7.205	7.573	8.114	8.091	8.091	4	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	5	2		4.623	5.338	6.485	7.338	8.131	6.446	7.338	4	4	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	5	3		4.593	5.686	7.054	7.054	8.049	8.049	8.049	4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	5	4		4.523	5.686	7.054	7.054	8.049	8.049	8.049	4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	5	5		4.543	5.780	7.227	8.020	8.760	8.600	8.600	4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	6	1		4.203	4.802					4	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
5	6	2		4.665	5.315	6.182	6.982			4	4	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
5	6	3		3.999	4.405	4.226	5.000			4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210		
5	6	4		4.548	5.307	6.970	7.070	7.515	8.182	8.182	4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	6	5		4.538	5.615	6.910	7.410	7.072	8.020	8.040	4	4	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	7	1		4.338	4.947	6.174	7.174	7.014	7.014	7.014	4	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	7	2		4.494	5.340	6.571	7.240	7.014	8.048	8.048	8.048	4	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210
5	7	3		4.549	5.615	6.910	7.410	7.072	8.020	8.040	4	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	7	4		4.567	5.729	7.110	8.020	8.059	9.771	10.471	4	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	7	5		4.597	5.881	7.110	8.020	8.059	9.771	10.471	4	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	7	6		4.120	4.990	6.130	7.120	8.165	9.277	9.602	4	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	8	2		4.590	5.338	6.565	7.376	8.195	8.967	9.109	5	1	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	8	3		4.535	5.692	6.829	7.590	8.314	9.150	9.609	5	2	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	8	4		4.567	5.729	7.110	8.020	8.059	9.771	10.471	5	2	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	8	5		4.597	5.729	7.110	8.020	8.059	9.771	10.471	5	2	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	8	6		4.000	4.945	6.220	7.121	8.165	9.277	9.602	5	2	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	9	2		4.430	5.410	6.667	7.407	8.210	9.219	9.702	5	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	9	3		4.568	5.425	6.800	7.725	8.456	9.490	10.150	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	9	4		4.598	5.729	7.110	8.020	8.059	9.771	10.471	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	9	5		4.597	5.729	7.110	8.020	8.059	9.771	10.471	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	9	6		4.643	5.801	7.240	8.020	9.170	10.187	10.409	5	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	10	1		4.643	5.410	6.667	7.407	8.210	9.219	9.702	5	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	10	2		4.568	5.425	6.800	7.725	8.456	9.490	10.150	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	10	3		4.598	5.729	7.110	8.020	8.059	9.771	10.471	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	10	4		4.597	5.729	7.110	8.020	8.059	9.771	10.471	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	10	5		4.643	5.801	7.240	8.020	9.170	10.187	10.409	5	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	10	6		4.643	5.410	6.667	7.407	8.210	9.219	9.702	5	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	1		4.568	5.425	6.800	7.725	8.456	9.490	10.150	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	2		4.598	5.729	7.110	8.020	8.059	9.771	10.471	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	3		4.597	5.729	7.110	8.020	8.059	9.771	10.471	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	4		4.643	5.801	7.240	8.020	9.170	10.187	10.409	5	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	5		4.643	5.410	6.667	7.407	8.210	9.219	9.702	5	3	1	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	6		4.568	5.425	6.800	7.725	8.456	9.490	10.150	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	7		4.598	5.729	7.110	8.020	8.059	9.771	10.471	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	8		4.597	5.729	7.110	8.020	8.059	9.771	10.471	5	3	2	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	9		4.643	5.801	7.240	8.020	9.170	10.187	10.409	5	3	3	4.535	5.033	5.378	5.473	5.495	10.536	11.210	
5	11	10		4.643	5.410	6.667															

Örnek 2

Gruplar	$ \bar{R}_i - \bar{R}_j $	KW	istatistiksel karar
1-2	4.5	2.115	p<0.05
1-3	8.5	2.115	p<0.05
2-3	4	2.229	p<0.05

ÖDEV 1

Egzersiz süreleri farklı gruplar arasında anksiyete skorları bakımından anlamlı farklılık var mıdır?(Anksiyete skorlarının normal dağılıma sahip olduğu bilinmektedir.)

grup1	grup2	grup3
5 saat	10 saat	15 saat
48	55	51
50	52	52
53	53	50
52	55	53
50	53	50

ÖDEV 2

Tıp eğitiminde kullanılan 4 farklı yöntemden sonra öğrencilerin başarı notları aşağıdaki gibidir. Başarı notları bakımından teknikler arasında anlamlı farklılık var mıdır?(Normal dağılıma uygun değildir.)

1	2	3	4
65	75	59	94
87	69	78	89
73	83	67	80
79	81	62	88