



## Bağıntı Analizi Sonuçlarının Yorumlanması

Selim Kılıç<sup>1</sup>

### ÖZET:

Bağıntı analizi sonuçlarının yorumlanması

Bağıntı analizi iki değişken arasındaki doğrusal ilişkinin boyutunu belirlemede kullanılan analizdir. Bağıntı katsayısı "r" simgesi ile gösterilir ve (-)1 ile (+)1 arasında değişen değerler alabilir. Bağıntı katsayısının başındaki "-" ya da "+" işareti, bağıntının yönünü belirtir. İlişkinin yönü, ilişkinin gücünü etkilemez. İyi derecede ilişki için r katsayısı 0.70 ve üzerinde olmalıdır. Bağıntı katsayısı, yalnızca doğrusal ilişkinin derecesini belirtir. Bağıntı katsayısının büyük bulunması, tek başına neden-sonuç ilişkisini açıklayamaz. Bağıntı analizinde de istatistiksel önemlilik örneklem büyüklüğünden çok etkilenir. Buna karşılık, bağıntı katsayısı değişkenlerin ölçüldüğü birimlerden etkilenmez. Bağıntı katsayısının karesi (R<sup>2</sup>) açıklayıcılık katsayısı olarak tanımlanır. Bağıntı katsayısı da, standart sapma gibi uç değerlerden çok etkilenir. Bağıntı analizinde incelenen değişkenin tipine ve dağılım özelliğine (sıralı, sürekli, normal dağılıma uyuyor, uyumuyor gibi) göre uygun olan test seçimi değişebilir. Pearson, Spearman ve Kendall'ın tau-B testleri en yaygın kullanılan testlerdir. Kısmi bağıntı ise üçüncü bir değişkenin etkisi kontrol altına alındığında, iki değişken arasındaki ilişkiyi açıklamayı sağlar. Bağıntı analizi sonuçlarını yorumlarken öncelikle ilişkinin gücünü değerlendirmeli ve bulunan bu sonucun istatistiksel olarak da önemli olup olmadığı vurgulanmalıdır.

**Anahtar sözcükler:** bağıntı, doğrusal ilişki, ilişkinin gücü

### ABSTRACT:

Interpretation of correlation analysis results

Correlation analysis is used to quantify the degree of linear association between two variables. Correlation coefficient is showed as "r" and it may have values between (-) 1 and (+)1. The symbols (-) or (+) in front of "r coefficient" show the direction of correlation. The direction of association does not affect the strength of association. A "r coefficient" which is equal or greater than 0.70 is accepted as a good association. Correlation coefficient only remarks the strength of association. Finding a good correlation coefficient between two variables does not explain a causal relationship between these two variables. Statistical significance is much affected from sample size in also correlation analysis. Square of correlation coefficient is described as coefficient of determination. Correlation coefficient is also much affected from outliers as standard deviation does. The proper test selection may vary depend on the type of variable and their distribution characteristics (ordinal, continuous, normally distributed or not). Pearson, Spearman and Kendall's tau-B tests are the most common used tests. Partial correlation provides the explanation of the association between two variables when the third variable was controlled. When we interpret the correlation analysis results we firstly evaluate the strength of association and then indicate whether it is also statistically significant or not.

**Key words:** correlation, linear association, strength of association

Journal of Mood Disorders 2012;2(4):191-3

<sup>1</sup>MD, Gülhane Askeri Tıp Fakültesi, Ankara-Türkiye

Yazışma Adresi / Address reprint requests to:  
Selim Kılıç, Gülhane Askeri Tıp Fakültesi,  
Ankara-Türkiye

Elektronik posta adresi / E-mail address:  
drselimkili@gmail.com

Kabul tarihi / Date of acceptance:  
20 Eylül 2012 / September 20, 2012

### Bağıntı beyanı:

S.K.: Yazar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemiştir.

### Declaration of interest:

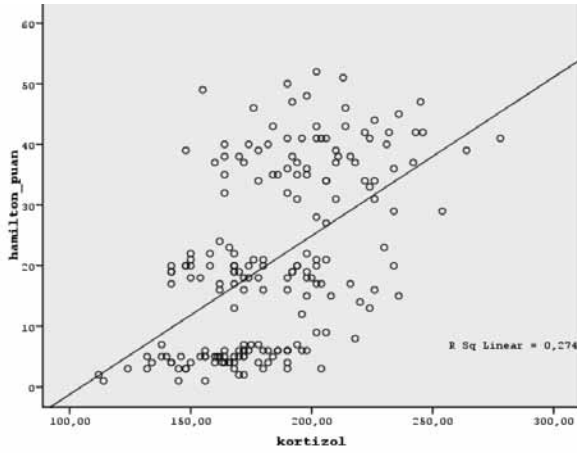
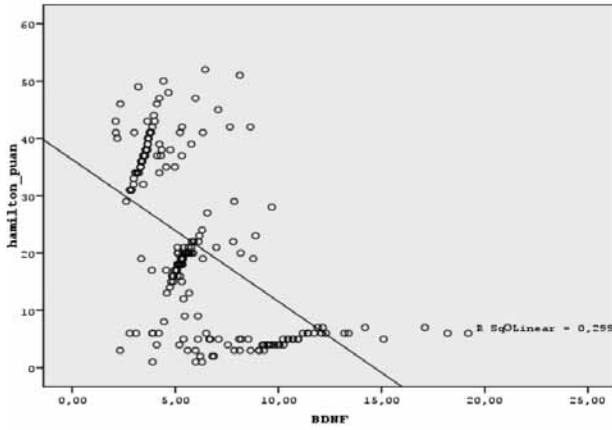
S.K.: The author declare that they have no conflict of interests regarding the content of this article.

## GİRİŞ

Bağıntı analizi iki değişken arasındaki doğrusal ilişkinin boyutunu belirlemede kullanılan analizdir (1). Değişkenlerin birlikte değişiminin derecesi olarak da tanımlanmaktadır (2). Bağıntı analizi, iki sayısal (sürekli) veya sıralı (ordinal) değişken arasında doğrusal ilişki olup olmadığını, ilişki varsa bunun yönünü ve gücünü araştıran bir tekniktir. Değişkenler arasındaki bağıntı incelenirken, değişkenlerin ölçüm skalalarına dikkat etmek gerekir. Değişkenin özelliğine yani sürekli veya sıralı olmasına

göre analizde kullanılacak olan uygun test değişir.

Bağıntı katsayısı "r" simgesi ile gösterilir ve (-)1 ile (+)1 arasında değişen değerler alabilir (3-6). Bağıntı katsayısının başındaki "-" ya da "+" işareti, bağıntının yönünü belirtir. "+" katsayısı her iki değişkenin birlikte arttığı veya azaldığını gösterirken, "-" katsayı bir değişkenin artarken diğerinin azaldığını gösterir. İlişkinin yönü, ilişkinin gücünü etkilemez. Yani (-)1'de (+)1'de ilişkinin gücü anlamında eşittir, sadece ilişkinin yönleri farklıdır. Aşağıda pozitif ve negatif doğrusal ilişkiye örnek iki şekil sunulmuştur (Şekil1, 2).

Şekil 1: Pozitif bağıntı ( $r=0.523$ ,  $p<0.001$ )Şekil 2: Negatif bağıntı ( $r= -0.547$ ,  $p<0.001$ )

### Bağıntı Katsayısının Yorumlanması

Katsayının (-) veya (+) olması ilişkinin gücü anlamında fark oluşturmamaktadır. Mutlak değer olarak aldığımızda;

- 0-0.249 Hiç ilişki yok ya da çok zayıf ilişki
- 0.25- 0.499 Zayıf – orta derecede ilişki
- 0.50–0.699 Orta derecede ilişki
- 0.70-0.899 İyi derece ilişki
- 0.90– 1.00 Çok iyi derecede ilişki olarak kabul edilir (4).

Örneğin  $r=0$  ise iki değişken arasında doğrusal bir ilişki yoktur. Ancak bu sonuç değişkenler arasında herhangi bir ilişki olmadığı anlamına gelmez. Bağıntı katsayısı, yalnızca doğrusal ilişkinin derecesini belirtir. Bağıntı katsa-

yısının büyük bulunması, tek başına neden-sonuç ilişkisini açıklamaz (7). Altta yatan mantıksal bir neden olmadıkça, büyük bulunan bir bağıntı katsayısı ile bir değişkenin diğer değişkeni etkilediği ya da ondaki değişimin nedeni olduğu yorumuna gidilmemelidir.

İstatistiksel önemliliğin örneklem büyüklüğünden çok etkilendiğini daha önceki makalelerde belirtmiştik (8-9). Sözgelimi örneklem büyüklüğü 10 olan bir grupta incelenen bağıntıda,  $p<0.05$  değeri için  $r=0.60$ 'ın üzerinde orta derece bir ilişkiye ihtiyaç var iken, örneklem büyüklüğü 100 kişi olduğunda  $r=0.20$  gibi zayıf bir ilişki derecesi istatistiksel önemlilik için yeterli olabilir.

Bağıntı katsayısı, değişkenlerin ölçüldüğü birimlerden etkilenmez. Yani birimin “metre” veya “santimetre” olması katsayıyı etkilemez. Bağıntı katsayısı, simetrik bir ölçüttür. Yani nokta grafikte hangi değişkenin yatay (x), hangi değişkenin dikey (y) ekseninde olduğu önemli değildir.

Bağıntı katsayısının karesi ( $R^2$ ) açıklayıcılık katsayısı olarak tanımlanır iken yabancı literatürde de “coefficient of determination” olarak ifade edilir. Açıklayıcılık katsayısı 1'e ne kadar yakınsa, iki değişken arasındaki doğrusal ilişki o kadar güçlüdür. Bulunan katsayı değerinin yorumu ise, incelenen iki değişkenden birindeki değişimin (varyasyonun) ne kadarının diğer değişkence açıklanabileceğidir (4,6,10). Bir değişkendeki değişimin kabaca yarısı ve daha fazlasının (%50 ve üstü) diğer değişkence açıklanabilmesi için iki değişken arasındaki bağıntı katsayısı “r” değerinin en az 0.70 olması gerekir.

Bağıntı katsayısı da, standart sapma gibi uç değerlerden çok etkilenir. Sürekli değişkenlerin birbirleriyle doğrusal ilişkisinin derecesi eğer her iki değişkende normal dağılım gösteriyorsa parametrik bir yöntem olan Pearson bağıntı katsayısı ile belirlenir. Örneğin Hamilton depresyon puanı ve kortizol değeri gibi iki sürekli değişken arasındaki bağıntının incelemesinde eğer bu iki değişkende normal dağılıma uygun dağılmışlarsa Pearson bağıntı testi uygun seçenektir. Pearson bağıntı analizi uç değerlere çok duyarlı olduğundan, normal dağılıma uymayan sürekli değişkenler arası doğrusal ilişki incelemesinde Spearman bağıntı analizi kullanılmalıdır (1,4). Aralarındaki bağıntının araştırıldığı iki değişkenden en az birinin simetrik olmayan dağılım yani normal dağılıma uymayan sürekli değişken olması ya da sıralı durumunda Spearman rho sıra bağıntıya uygun olan testtir. Sözgelimi Hamilton depresyon puanı veya kortizol değerinden biri

normal dağılıma uygun dağılmıyorsa veya Hamilton depresyon puanı veya kortizol değerlerinden bir veya ikisi sıralı değişken olarak analize alınmışsa bu durumda uygun seçenek Spearman rho testi olacaktır. Diğer yandan incelenen değişkenlerin biri veya her ikisi de sıralı değişkense Kendall'in tau-b katsayısı kullanılması uygun seçenek olacaktır (4,10). Spearman bağıntı katsayısı ile Kendall'in tau-b katsayısı arasındaki fark oldukça küçüktür. Bu testlerden Kendall tau testinin küçük örneklemelerde Spearman testine göre daha üstün olduğu belirtilmektedir (6,11).

Kısmi (partial) bağıntı ise üçüncü bir değişkenin etkisi kontrol altına alındığında, iki değişken (eğer çoklu bağıntı dikkate alınırsa, daha fazla değişken) arasındaki ilişkiyi açıklamayı sağlar. Kısmi bağıntı  $r_{12.3}$  şeklinde yazılabilir. Gizli ilişkilerin ortaya çıkmasında kullanışlı bir yöntem olan kısmi bağıntı analizi, çoklu regresyon ile yakından ilişkili olup, bir ya da daha fazla ilave değişkenin doğrusal

etkisini sabit tutmak suretiyle, iki değişken arasındaki doğrusal ilişkinin belirlenmesinin sağlar (4). Örnek olarak aylara ait dondurma tüketimi ile aylara ait boğulmalar arasındaki bağıntı incelenirken üçüncü değişken olarak aylara ait ortalama hava sıcaklığı kısmi bağıntı analizine kontrol edilen değişken olarak konulursa boğulma sayıları ile dondurma tüketimi arasında doğrusal bir ilişki olmadığı bulunur.

Sonuç olarak; değişkenler arasında özellikle sürekli ve sıralı değişkenler arasında doğrusal bir ilişkinin varlığını incelediğimiz bağıntı analizinde bulunan katsayı (-)1 veya (+)1'e ne kadar yakınsa doğrusal ilişki o kadar güçlüdür. Buna karşılık ilişkinin istatistiksel önemliliği ise ilişkinin gücünün yanı sıra özellikle örnek büyüklüğünden etkilenir. Dolayısı ile sonuçları yorumlarken öncelikle ilişkinin gücünü gösteren "r" katsayısı dikkate alınarak sonuçlar yorumlanmalı ve bulunan katsayının istatistiksel olarak da önemli olup olmadığı vurgulanmalıdır.

## Kaynaklar:

1. Pagano M, Gauvreau. Principles of Biostatistics. Duxbury Press, 363-372, 1993.
2. Last JM. A Dictionary of Epidemiology. Oxford University Press. 4<sup>th</sup> Edition, 41, 2001.
3. Morton RF, Hebel JR, McCarter RJ. A Study Guide to epidemiology and Biostatistics, Aspen Publication, 92, 1996.
4. Alpar R. Spor, Sağlık ve Eğitim Bilimlerinden Örneklerle Uygulamalı İstatistik ve Geçerlik-Güvenirlilik, Detay Yayıncılık, 261-182, 2010.
5. Kirkwood BR, Sterne JAC. Medical Statistics. Blackwell Science Publishing. 2<sup>nd</sup> Edition, 94, 2003.
6. Jekel JF, Elmore JG, Katz DL. Epidemiology Biostatistics and Preventive Medicine. WB Saunders Company, 140-144. 1996.
7. Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. Medical Epidemiology. Lange Medical Books, 3<sup>rd</sup> Edition. 38-40, 2001.
8. Kılıç S. Neyin Peşindeyiz? Kutsal p değerinin mi (istatistiksel önemlilik) yoksa klinik önemliliğin mi? JMOOD. 2011;1:46-8.
9. Kılıç S. Örnek büyüklüğü, güç kavramları ve örnek büyüklüğü hesaplaması. JMOOD. 2012;2:140-2.
10. Dawson B, Trapp RG. Basic & Clinical Biostatistics. Lange Medical Books, 3<sup>rd</sup> Edition, 184-94, 2001.
11. Arndt S, Turvey C, Andreasen NC. Correlating and predicting psychiatric symptom ratings: Spearman's r versus Kendall's tau correlation. J Psychiatr Res. 1999;33:97-104.