

Birden Çok Değişken Arasındaki İlişkilerin Analizi

Çeviri: Harika Gözde GÖZÜKARA BAĞ

BÖLÜM İÇERİĞİ

1. Çoklu İstatistiklere Genel Bakış

2. Çoklu Yöntemlerin Altında Yatan Varsayımlar

- 2.1 Çoklu Analiz İçin Denklemlerin Kavramsal Olarak Anlaşılması
- 2.2 En İyi Kestirimler
- 2.3 Genel Doğrusal Model
- 2.4 Çoklu İstatistiklerin Kullanımları

3. Çoklu Analiz İçin İşlemler

3.1 Varyans Analizi (ANOVA)

3.1.a. Tek-Yönlü ANOVA (F-Testi)

3.1.b. N-Yönlü ANOVA

3.2 Kovaryans Analizi (ANCOVA)

3.3 Çoklu Doğrusal Regresyon

3.4 Çoklu Analiz İçin Diğer İşlemler

4. Özet

Değerlendirme Soruları, Cevaplar ve Açıklamalar

Doğru probleme yaklaşık bir cevap, yaklaşık bir probleme kesin bir cevaptan çok daha değerlidir.

John Tukey

1. ÇOKLU İSTATİSTİKLERE GENEL BAKIŞ

Bir tane sonuç değişkeni, fakat birden fazla bağımsız değişkeni olan istatistiksel modeller genel olarak çoklu modeller (ya da birçok istatistikçinin bu terimi birden çok bağımlı değişkeni olan modeller için kullanmasına rağmen çok değişkenli modeller) olarak adlandırılırlar. Çoklu modeller, araştırmacılara sezgisel olarak çekici gelmektedir, çünkü sadece bir bağımsız değişkenli modellere göre daha “gerçek hayattaki” gibi gözükmektedirler. Bölüm 10’da tanımlanan, iki değişkenli bir analiz, basitçe X ’deki harekete bağlı olarak Y ’de anlamlı bir hareket olup olmadığını gösterir. Çoklu analiz, diğer faktörlerin (örneğin A, B ve C) etkisi göz önüne alındığında X ’deki değişimin etkisinin ve Y ’deki değişimin değerlendirilmesine olanak verir.

Çoklu analiz, bağımlı (sonuç) bir değişkendeki (y) değişimin açıklanmasında farklı bağımsız değişkenlerin tek başlarına ve birlikte çalıştıklarında (etkileşim) görece önemlerinin anlaşılmasına yardım eder. Farklı bağımsız değişkenlerin, bağımlı bir değişkeni açıklama becerisinde dikkate değer çakışmalar olabilir. Örneğin, hayatın ilk yirmi yılında, yaş ve boy uzunluğu vücut ağırlığını kestirir, ancak yaş ve boy uzunluğu genellikle ilişkilidir. Büyüme yılları boyunca, boy uzunluğu ve vücut ağırlığı yaşla birlikte artar, böylece yaş vücut ağırlığını etkileyen alta yatan açıklayıcı değişken gibi dikkate alınabilir ve boy uzunluğu *aracı değişken* olarak görülebilir. Çocuklar farklı hızda büyürler, bu ne-

denle boy uzunluğu, yaşa ek açıklayıcı güç katacaktır. Ortalamadan yaşına göre daha uzun olan çocuklar, aynı zamanda aynı yaşta olup yaşlarına göre kısa olan çocuklardan daha ağır olacaklardır. Her bir bağımsız değişken diğer bağımsız değişkenler ile açıklayıcı gücü paylaşabilir ve diğer herhangi bir değişkenin açıkladığının ötesinde y ’deki bazı değişimleri açıklayabilir.

Tüm istatistiksel denklemler gerçekte var olanı modellemeyi dener. Bununla birlikte, tek bir değişkenin (örneğin bir besin) başka bir değişken (örneğin bir bebeğin büyüme hızı) üzerindeki etkisi gibi gerçeğin sadece bir boyutunu temsil edebilirler. Bilimsel değeri olan basit bir model için, araştırma tasarımı, çalışılan bağımsız ve bağımlı değişkenlerden başka tüm faktörleri dengelemeyi denemek zorundadır. Hayvan çalışmalarında bu durum genetik olarak özdeş hayvanların kullanılmasıyla sağlanabilir. Tek yumurta ikizlerinin bazı gözlemsel çalışmaları dışında, bu durum insanlar için yapılamaz.

İnsanları içeren deneysel araştırmalar için ilk adım, çalışma katılımcılarının çalışma gruplarına atamalarını rasgele yaparak deney ve kontrol gruplarını benzer yapmaktır. Bununla birlikte, bazen randomizasyon mümkün değildir, veya önemli faktörler bu strateji ile yeterince kontrol edilemeyebilir. İstenmeyen bu faktörlerin etkisini kaldırmanın bir yolu, **çoklu istatistiksel analiz** kullanarak onları kontrol altına almaktır. Yaygın kullanımın içerdikleri:

1. Tıbbi veya halk sağlığı girişimlerinin etkilerini çalışırken, araştırma gruplarını eşitlemek (örneğin onları olabildiğince karşılaştırılabilir yapmak).
2. Toplumda farklı hastalıkların riskini hangi faktörlerin etkilediğini, araştırmacıların anlamasına yardım eden *gözlemsel çalışmalardan* model kurmak (klinik ve halk sağlığı çabalarına, sağlı-

kullanışlı bir lojistik regresyon modelinin geliştirilmesi için esastır. Bu yöntem, bakteriyel menenjitli hastaların iyi bir sonuç (tam iyileşme) veya kötü bir sonuç (kalıntı nörolojik sekel veya ölüm) yaşayacağını hastaneye kabul edildiğinde kestirmek için geliştirilmiş bir modeli test etmek için kullanılmıştır.¹⁶

Sağkalım verileri için yaygın olarak kullanılan bir regresyon modeli, gruplar arasındaki anlık ölüm veya başka bir olay riskindeki farklılıkları test etmek için kullanılan **Cox orantılı riskler (hazard) modelidir**. Aynı zamanda, hangi değişkenlerin daha iyi sağkalım ile ilişkili olduğunu belirlemek için de kullanılır. Örneğin, rektum kanserli iki hasta grubunun: radyasyon artı uzun süreli florourasil infüzyonu ile tedavi edilen hastalar ve bolus florourasil enjeksiyonu verilen hastaların nüksüz sağkalımını karşılaştırmak için Cox modelleri kullanılmıştır.¹⁷ Bu tür yöntemlerin mekanığı bu metnin kapsamı dışındadır, ancak tıbbi literatürü yorumlamada bunların ne zaman garanti edildiğini bilmek önemlidir.

4. ÖZET

Çoklu analiz, birkaç bağımsız (muhtemelen nedensel) değişkenin, ayrı ayrı ve birlikte, tek bir bağımlı (sonuç) değişkenindeki varyasyonu ne kadar iyi açıkladığını belirlemek için istatistiksel yöntemler içerir. Tıbbi araştırmalarda, çoklu analiz için üç yaygın kullanım vardır: (1) sonuç değişkeni üzerindeki diğer değişkenlerin etkilerini kontrol ederek bir klinik denemede bir müdahalenin test edilmesini iyileştirmek; (2) bir veya daha fazla bağımsız değişkenin hastalık veya ölüm riski üzerindeki göreceli etkisini kestirerek, gözlemsel çalışmalarda bir hastalığın etiyolojisi veya prognozuna ışık tutmak; ve (3) bir tanı veya prognostik skorlama sisteminde kullanılan farklı değişkenler için ağırlıklar geliştirmek. Tablo 11.1'de gösterildiği gibi, çoklu analiz için kullanılacak uygun yöntemin seçimi, bağımlı ve bağımsız değişkenlerin sürekli, iki durumlu, nominal, ordinal veya bunların bir kombinasyonu olmasına bağlıdır. Klinik araştırmalarda çoklu tekniklerin kullanılmasının birçok potansiyel sorunu ve sınırlaması olduğundan, bu prosedürler anlayarak ve özenle kullanılmalı ve yorumlanmalıdır.

KAYNAKLAR

1. Kleinbaum DG. *Applied Regression Analysis and Other Multivariable Methods*, 4th ed. Belmont, CA: Brooks/Cole; 2008.
2. Horton NJ, Switzer SS. Statistical methods in the journal. *N Engl J Med*. 2005;353:1977-1979.
3. Arbogast PG, Ray WA. Performance of disease risk scores, propensity scores, and traditional multivariable

outcome regression in the presence of multiple confounders. *Am J Epidemiol*. 2011;174:613-620.

4. Holford TR. *Multivariate Methods in Epidemiology*. New York, NY: Oxford University Press; 2002.
5. Freeman Jr DH, D'Atri DA, Hellenbrand K, et al. The prevalence distribution of hypertension: Connecticut adults 1978-1979. *J Chronic Dis*. 1983;36:171-181.
6. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA*. 2006;295:801-808.
7. Yoo HH, De Paiva SA, Silveira LV, Queluz TT. Logistic regression analysis of potential prognostic factors for pulmonary thromboembolism. *Chest*. 2003;123:813-821.
8. Ayanian JZ, Landrum MB, Guadagnoli E, Gaccione P. Specialty of ambulatory care physicians and mortality among elderly patients after myocardial infarction. *N Engl J Med*. 2002;347: 1678-1686.
9. D'Agostino Jr RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*. 1998;17:2265-2281.
10. Nakayama M, Osaki S, Shimokawa H. Validation of mortality risk stratification models for cardiovascular disease. *Am J Cardiol*. 2011;108:391-396.
11. van den Hoogen PC, Feskens EJ, Nagelkerke NJ, Menotti A, Nissinen A, Kromhout D. The relation between blood pressure and mortality due to coronary heart disease among men in different parts of the world. *N Engl J Med*. 2000;342: 1-8.
12. Dawson B, Trapp RG. *Basic & Clinical Biostatistics*. 4th ed. New York, NY: Lange Medical Books/McGraw-Hill; 2004.
13. Pearlman DS, Chervinsky P, LaForce C, et al. A comparison of salmeterol with albuterol in the treatment of mild-to-moderate asthma. *N Engl J Med*. 1992;327:1420-1425.
14. Finkelstein JS, Klibanski A, Schaefer EH, Hornstein MD, Schiff I, Neer RM. Parathyroid hormone for the prevention of bone loss induced by estrogen deficiency. *N Engl J Med*. 1994;331: 1618-1623.
15. Stacpoole PW, Wright EC, Baumgartner TG, et al. A controlled clinical trial of dichloroacetate for treatment of lactic acidosis in adults. *N Engl J Med*. 1992;327:1564-1569.
16. Aronin SI, Peduzzi P, Quagliarello VJ. Community-acquired bacterial meningitis: risk stratification for adverse clinical outcome and effect of antibiotic timing. *Ann Intern Med*. 1998; 129:862-869.
17. O'Connell MJ, Martenson JA, Wieand HS, et al. Improving adjuvant therapy for rectal cancer by combining protracted infusion fluorouracil with radiation therapy after curative surgery. *N Engl J Med*. 1994;331:502-507

OKUMA ÖNERİLERİ

Feinstein AR. *Multivariable Analysis: An Introduction*. New Haven, CT: Yale University Press; 1996