

Hipotezlerin Testi

Çeviri: İsmet DOĞAN

BÖLÜM İÇERİĞİ

1. İstatistiksel Çıkarsama

- 1.1 Tümdengelim ve Tümevarım Arasındaki Farklar
- 1.2 Matematik ve İstatistik Arasındaki Farklar

2. Hipotezlerin Testi

- 2.1. Yanlış-Pozitif ve Yanlış-Negatif Hatalar
 - 2.1.a. Yokluk Hipotezi ve Alternatif Hipotez Geliştirme
 - 2.1.b. Alfa Düzeyinin Belirlenmesi
 - 2.1.c. İstatistiksel Anlamlılık Testi
 - 2.1.d. Alfa ile Elde Edilen p Değerinin Karşılaştırılması
 - 2.1.e. Yokluk Hipotezini Ret veya Kabul Etme,
 - 2.2 Bireysel Gözlemlerde ve Çoklu Örneklemelerde Varyasyon
 - 2.2.a. Standart Sapma ve Standart Hata
 - 2.2.b. Güven Aralıkları

3. İstatistiksel Anlamlılık Testleri

- 3.1. Kritik Oranlar
- 3.2. Serbestlik Derecesi
- 3.3. t-Testlerinin Kullanımı
 - 3.3.a. Örnek Popülasyonlar ve Büyüklükleri
 - 3.3.b. İki Örneklem t-Testi
 - 3.3.c. t Dağılımı
 - 3.3.d. Eşleştirilmiş t-Testi
- 3.4. z -Testlerinin Kullanımı
- 3.5. Diğer Testlerin Kullanımı

4. Özel Durumlar

- 4.1 Gruplar Arası Varyasyona Karşı Grup İçi Varyasyon
- 4.2 Klinik Önem ve Dış Geçerliliğe Karşı İstatistiksel Anlamlılık

5. Özet

Değerlendirme Soruları, Cevaplar ve Açıklamalar

“Olağanüstü iddialar olağanüstü kanıt gerektirir.”

Carl Sagan

Varyasyonun doğası, veri türleri, değişkenler ve Bölüm 8’de bahsedilen veri dağılımının özellikleri hakkındaki birikimlerimizle istatistiksel yöntemler kullanılarak verilerden nasıl çıkarımlarda bulunulacağı ve hipotezlerin nasıl test edileceği incelenecektir.

1. İSTATİSTİKSEL ÇIKARSAMA

Çıkarsama, verilerden sonuçların çıkarılması anlamına gelir. İstatistiksel çıkarsama ise, verilerin tanımlanması ve düzenlenmesi, uygun hipotezlerin test edilmesi ve istatistiksel yöntemleri kullanarak nicel veya nitel bilgilerden sonuçların çıkarılması olarak tanımlanabilir.

1.1 TÜMDENGELİM VE TÜMEVARIM ARASINDAKİ FARKLAR

Veriler kendi yorumlarını sağlamadığından, tümevarım (Latince’de, “yol açmak” anlamına gelir) yoluyla yorumlanmalıdır. Çoğu insan akıl yürütmeye yönelik bu yaklaşıma, matematikten özellikle de geometriden

öğrenilen tümdengelim’e (Latince, “dışarı çıkarmak”) göre daha az aşınadır. Tümdengelim genelden (yani, doğru kabul edilen varsayımlardan, önermelerden veya formüllerden) özele (yani, genel kategoriye ait belirli üyelere) doğru ilerler. Aşağıdaki iki önermeyi göz önünde bulundurun:

- Bütün Amerikalılar demokrasiye inanır.
- Bu kişi bir Amerikalı.

Her iki önerme de doğruysa, “Bu kişi demokrasiye inanıyor” çıkarımı doğru olmalıdır. Tümdengelim, hipotezler oluşturulduktan sonra bilimde özel olarak kullanılır. Tümdengelim’i kullanan bir araştırmacı, “hipotez doğruysa, tahmin veya tahminler de doğru olmalıdır” diyebilir. Bir tahmin deneysel olarak test edilebilirse, bulgulara dayanarak hipotez ret veya kabul edilir. Veriler, hipotezden elde edilen tahminlerle tutarsızsa, hipotez reddedilmeli veya değiştirilmelidir. Veriler hipotezle tutarlı olsa bile, Bölüm 4’te gösterildiği gibi hipotezin doğru olduğunu kanıtlayamazlar (bkz. Şekil 4.2).

Klinisyenler, belirli bir klinik durumda değişkenlerin sahip olması gereken değerleri belirlemek için genellikle doğru olarak kabul edilen formüllerden ve gözlemlenen verilerden yola çıkarlar. Örneğin, bir kilogram vücut ağırlığı başına güvenle verilebilecek ilaç

bir ilaçla kan basıncında ortalama 2 mm Hg'lik bir düşüşün istatistiksel olarak anlamlı olduğunu göstermek mümkündür. Bununla birlikte, kan basıncındaki bu kadar küçük bir düşüş, klinik olarak pek yararlı olmayacak ve bireysel bir hastada klinik olarak önemli olmayacaktır. Bununla birlikte, büyük bir popülasyonda kan basıncında bu miktarda ortalama bir azalma, bazı kalp krizlerini ve felçleri önleyebilir ve bir miktar değeri olabilir, ancak sınırlı faydaların maliyetler ve yan etkilerle karşılaştırılması gerekir. Bazen bir hastayı tedavi eden bir klinisyen ile büyük bir nüfus üzerine etkisi olan bir müdahaleyi düşünen halk sağlığı pratisyeni, araştırma bulgularını farklı yorumlar.

Ek olarak, bir çalışmanın bulguları genel klinik kullanıma sunulmadan önce, çalışmanın dış geçerliliği veya genellenebilirliğinin olup olmadığı konusu ele alınmalıdır. Yeni bir ilacın klinik denemesinde, çalışmadaki hasta örnekleminin, sonunda yeni ilacın kullanılabileceği hasta evrenini temsil edip etmediği sorulmalıdır. Çalışmaların dış geçerliliği olmayabilir, çünkü hasta örneklemindeki hastalık spektrumu, hastaların evrenindeki hastalık spektrumundan farklıdır. Hastalığın türleri, evreleri ve şiddeti değişebilir. Hastaları dahil etme veya hariç tutma kriterleri açısından spektrum açıkça tanımlanmalı ve bu kriterler bulgularla birlikte rapor edilmelidir. Şiddetli hastalığı olan hastalar çalışmadan çıkarılmışsa, çalışmanın sonuçları ciddi hastalığı olan hastalara genellenemeyeceğinden, bu dışlama kriteri bildirilmelidir. (Bölüm 4'teki Doktorların Sağlık Çalışması tartışmasına bakın; sağlıklı, uyumlu doktorlarla yapılan bir çalışmanın sonuçları tüm yetişkinlere genellenemeyebilir.)

Çalışma ayrıca, hasta örneklemindeki bireysel özelliklerin spektrumu, hastaların evrenindeki bireysel özelliklerin spektrumundan farklı olduğu için dış geçerlilikten yoksun olabilir. Örnekleme, diğer özelliklerin yanı sıra yaş, cinsiyet, gelir, eğitim düzeyi veya etnik köken temelinde çoğu hastadan farklılık gösterebilir.

5. ÖZET

İstatistik, gözlemlenen verilerde genelleştirilebilir ilişkiler ve farklılıklar bulma çabası olan tümevarımsal akıl yürütmeye yardımcı olur. Bir sonucu tahmin etmek için bilinen formülleri belirli verilere uygulama girişimi olan matematiğin tersi bir süreçtir. İstatistikler, araştırmacıların gözlemlenen verilerden makul sonuçlara ve tahminlere ulaşmasına yardımcı olur ve verilerden sonuçlar ve tahminler yaparken hatalı olma olasılığına yaklaşık sınırlar sağlar. Anlamlılık testi, bir deney grubundan alınan bir ölçümün ortalaması ile bir kontrol grubunun ortalaması arasında gerçek bir fark olmadığı gibi bir yokluk hipotezi ifadesi ile başlar.

Bir sonraki adım, yokluk hipotezi reddedilirse yanlış olma olasılığını tahmin etmek için bir test istatistiği (genellikle kritik bir oran) ve bununla ilişkili p değerini hesaplamaktır. Anlamlılık testinin sonuçları araştırmacının yokluk hipotezini reddetmesine izin vermesi, gruplar arasında bir fark olduğuna dair kanıttır.

İki örneklem t-testi araştırmacının, ortalamalar arasındaki farkın yalnızca şans eseri beklenenden daha büyük olup olmadığını belirlemek için iki farklı çalışmaya katılımcısı grubundan sürekli bir değişkenin (örneğin ağırlık) ortalamalarını karşılaştırmasını sağlar. Eşleştirilmiş t-testi, araştırmacının bir müdahaleden önce ve sonra bir grup çalışma katılımcısında sürekli bir değişkendeki ortalamayı değerlendirmesini sağlar. Ortalamalar arasındaki farkı karşılaştıran t-testlerinin aksine, z-testleri oranlar arasındaki farkı karşılaştırır.

KAYNAKLAR

1. Moore DS, McCabe GP. Introduction to the Practice of Statistics. 7th ed. New York, NY: Freeman; 2010.

OKUMA ÖNERİLERİ

- Dawson B, Trapp RG. Basic & Clinical Biostatistics. 4th ed. New York, NY: Lange Medical Books/McGraw-Hill; 2004.
- Silver N. The Signal and the Noise: Why So Many Predictions Fail but Some Don't. New York, NY: Penguin Press; 2012.

DEĞERLENDİRME SORULARI

1. Yerel bir hastanede çalışırken tanıştığınız ilk üç tıp stajyeri, yaygın olarak reçete edilen bir anti depresan almaya başladıklarından beri kendilerini çok daha iyi hissettiklerini bildiriyor. İstemeyerek stajın depresyonla ilişkili olduğu sonucuna ulaşıyor-sunuz. Bu sonuca ulaşmak aşağıdakilerden hangisi için bir örnektir?
 - A. Hipotez testi
 - B. Tümevarım
 - C. Tümdengelim
 - D. İnterpolasyon
 - E. Ekstrapolasyon
2. 100 öğrenciden bu bölümü okumaları istendiğinde, her öğrencinin metni belli belirsiz anlaması için geçen ortalama süre 3.2 saat, %95 güven aralığı 1.4-5.0 saattir. Bu durumda %95 güven aralığı aşağıdakilerden hangisini temsil eder?
 - A. Örneklem ortalaması, 3.2 1 standart sapmadır.
 - B. Popülasyon ortalaması, 3.2 1 standart hatadır.
 - C. Gerçek popülasyon ortalaması için mümkün olan maksimum ve minimum değerlerdir.