

Çalışma Tasarımlarında İstatistiğin Kullanılması: Örneklem Büyüklüğü Hesaplama, Randomizasyon ve Çoklu Hipotez Düzeltmesi

Çeviri: Selçuk KORKMAZ

BÖLÜM İÇERİĞİ

1. Örneklem Büyüklüğü

- 1.1. Beta (Yanlış-Negatif) Hatası
- 1.2. Temel Örneklem Büyüklüğü Formülünün Elde Edilmesi
- 1.3. Örneklem Büyüklüğünün Hesaplanmasındaki Adımlar
- 1.4. T-Testlerini Kullanan Çalışmalar İçin Örneklem Büyüklüğü
- 1.5. Oran Farklarının Testi İçin Örneklem Büyüklüğü

2. Çalışma Katılımcılarının Randomizasyonu

- 2.1. Randomizasyonun Amaçları
- 2.2. Randomizasyon Yöntemleri
 - 2.2.a. Basit ve Blok Randomizasyon
 - 2.2.b. Sistematik Tahsis
 - 2.2.c. Tabakalı Tahsis
- 2.3. Randomizasyonla İlgili Özel Sorunlar

3. Çoklu Hipotez Testleri İçin Düzeltme

4. Özet

Değerlendirme Soruları, Cevaplar ve Açıklamalar

"Kanıtın yokluğu, yokluğun kanıtı değildir."

Carl Sagan

Doğru bir şekilde tasarlandığında, analiz edildiğinde ve yorumlandığında, randomize kontrollü çalışmalar (RKÇ), bir klinik veya halk sağlığı müdahalesinin işe yarayıp yaramadığını belirlemede en iyi yöntemdir. Randomizasyon, önemli demografik, sosyal ve biyolojik özellikleri karşılaştırma grupları arasında rastgele dağıtarak seçim yanlılığını azaltır. Teoride, randomizasyon ile gruplar birbirine yeterince benzer şekilde oluşturulmuş, bu nedenle müdahale ve kontrol grupları arasında sonuçlarda meydana gelen herhangi bir farklılık müdahalenin etkisine atfedilebilir. Çalışmaları tasarlamada istatistiğin en önemli kullanım alanlarından biri karşılaştırma grupları için örneklem büyüklüklerini belirleyerek, gruplar arasında gerçekten bir farklılık varsa çalışmanın bu farklılığı tespit etmesini sağlamaktır. RKÇ'lerde istatistiğin diğer kullanım alanları, randomizasyon yöntemleri ve çoklu hipotez testlerinin için düzeltme uygulanmasıdır.

1. ÖRNEKLEM BÜYÜKLÜĞÜ

İstatistiksel açıdan düşünüldüğünde, eğer gruplar arasında gerçekte bir farklılık varsa, istatistiksel anlamlılığın elde edilmesi için örneklem büyüklüğünün belirlenmesi oldukça önemlidir. Örneklem büyüklüğü hesaplamaları klinik araştırmaların planlanmasında

da kritik öneme sahiptir. Çünkü örneklem büyüklüğü genellikle çalışmayı gerçekleştirmek için gereken zaman ve maliyeti belirlemede en önemli faktördür.

Klinik araştırmaları değerlendirmekten ve finanse etmekten sorumlu kişiler, çalışma için ihtiyaç duyulan katılımcı sayısını tahmin etmede kullanılan varsayımları ve örneklem büyüklüğü hesaplamalarının gerçekleştirilme şeklini yakından incelemektedirler. Bu kişilerin görevlerinden biri de örneklem büyüklüğünü gözden geçirirken, önerilen araştırmanın gerçekçi olup olmadığını belirlemektir (örneğin bir RKÇ'deki müdahale ve kontrol gruplarına veya bir vaka kontrol çalışmasında vaka ve kontrol gruplarına yeterli sayıda katılımcının dahil edilip edilmediği). Literatürdeki araştırmalarda, anlamlı olması beklenen sonuçların neden istatistiksel açıdan anlamsız olduğu yetersiz örneklem büyüklüğü ile açıklanabilir.

Araştırmacılar tarafından istatistikçilere, başka nedenlerden ziyade, örneklem büyüklüğü hesaplamaları için danışılmaktadır. Örneklem büyüklüğü hesaplamaları, sıradan istatistiksel analizleri sorunsuz bir şekilde gerçekleştiren araştırmacılar için bile kafa karıştırıcı olabilir. Örneklem büyüklüğü ile ilgili öngürülerinizi sınamak için aşağıdaki üç soruyu cevaplamaya çalışın.

Örneklem büyüklüğü aşağıdaki durumlardan nasıl etkilenirdi:

1. İncelenen sonuç değişkeni yüksek miktarda değişkenlik içerseydi?

Bölüm 5 ve Şekil 5.6). Bununla birlikte, matristeki on hücre her bir değişkenin kendisi ile olan ilişkisini de içerecektir, ki bu da her zaman mükemmel bir korelasyondur. Bu durumda, elimizde 90 adet olası ilişki kalacaktır, ancak bunların yarısı "" ve diğer yarısı ise "" korelasyonudur. Aynı ikili çiftlerden elde edilen p değerleri hangi değişkenin bağımsız ve hangi değişkenin bağımlı olduğundan bağımsız olarak aynı çıkacaktır, bu durumda elimizde gerçekte sadece yarı sayıda, yani 45 adet, bağımsız ilişki bulunmaktadır. Alfa için kesim noktası olarak kullanılırsa, 100 bağımsız ilişkiden 5'inin sadece şansa bağlı olarak gerçekleşmesi beklenir.¹⁰ Örneğimizde ise ikiden biraz daha fazla ilişkinin sadece şansa bağlı olarak gerçekleşmesinin bekleneceği anlamına gelir.

Çoklu hipotezlerle ilgili problem, çoklu ilişkiler ile ilgili probleme benzerdir. Bu durumda, test edilen hipotezlerin sayısı ne kadar fazlaysa, bunlardan en az birinin sadece şansa bağlı olarak istatistiksel açıdan anlamlı bulunması o kadar olasıdır. Bu sorunu çözmenin yollarından biri, sıfır hipotezini reddetmeden önce gereken p değerini azaltmaktır (örneğin p değerini $<0,05$ yapın). Bu yaklaşım, aynı tıbbi eğitim hipotezini beş farklı hastanede test eden bir çalışmada gerçekleştirilmiştir.¹¹ Çalışmadaki alfa seviyesi 0,05 olarak belirlenseydi, her hastane %5 hata yapma olasılığı olduğu için (), beş hastaneden en az birinde sadece şansa bağlı olarak istatistiksel açıdan anlamlı bir fark bulma olasılığı neredeyse %25 olacaktı. Tüm çalışmadaki yanlış pozitif bulgu riskini 0,05'in üzerine çıkarmamak için sıfır hipotezini reddetmek üzere seçilen alfa seviyesi 5'e (hastane sayısına) bölünerek 0,01 yapılmış ve daha katı hale getirilmiştir. Bu çoklu hipotezleri düzeltme yöntemine Bonferroni düzeltmesi denir ve oldukça katı bir yöntemdir. Diğer olası düzeltme yöntemleri ise daha az katı olmakla birlikte istatistiksel olarak daha karmaşıktırlar ve farklı durumlarda kullanılmaktadırlar. Bunlar Tukey, Scheffé ve Newman-Keuls yöntemleridir.¹²

4. ÖZET

Bilimsel çalışmalar için gereken örneklem büyüklüklerinin hesaplanmasında istatistikçilere sıklıkla danışılmaktadır. Bununla birlikte, örneklem büyüklükleri ancak ve ancak araştırmacı hesaplamalarda kullanılacak bilgileri belirlediğinde tahmin edilebilir. Bu ise alfa ve beta düzeylerini, sonuç değişkenlerindeki klinik olarak önemli farkı ve beklenen varyansı içermektedir. Randomizasyon ve körleme, müdahale ve kontrol gruplarını karşılaştıran çalışmalarda yanlılığı azaltabilen yaklaşımlardır. İdeal olarak, ne katılımcılar, ne

müdahaleyi gerçekleştirenler ne de verileri toplayanlar kimin hangi tedaviyi aldığını bilmelidir. Rastgele tahsiste kullanılan temel yöntemler; basit ve blok rastgele tahsis, sistematik tahsis ve tabakalı tahsistir. Veri analizinde, araştırmacılar veri eşeleme sorununun farkında olmalı ve çoklu hipotezleri test etmenin yanlış-pozitif istatistiksel ilişki elde etme (alfa hataları) olasılığını artırdığını anlamalıdırlar.

KAYNAKLAR

1. Freiman JA, Chalmers TC, Smith Jr H, Kuebler RR: The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. N Engl J Med 1978; 299: pp. 690-695.
2. Williams HC, Seed P: Inadequate size of 'negative' clinical trials in dermatology. Br J Dermatol 1993; 128: pp. 317-326.
3. Chow SC, Wang H, Shao J: 2nd ed2007.Chapman & Hall/CRCNew York, NY
4. Kelsey JL, Whittemore AS, Evans AS, Thompson WD: 2nd ed1996.Oxford University PressNew York, NY
5. Lam JA, Hartwell S, Jekel JF: "I prayed real hard, so I know I'll get in": living with randomization. New Dir Prog Eval 1994; 63: pp. 55-66.
6. Paradise JL, Bluestone CD, Bachman RZ, et. al.: Efficacy of tonsillectomy for recurrent throat infection in severely affected children. Results of parallel randomized and nonrandomized clinical trials. N Engl J Med 1984; 310: pp. 674-683.
7. Sackett DL, Gent M: Controversy in counting and attributing events in clinical trials. N Engl J Med 1979; 301: pp. 1410-1412.
8. MacMahon B, Yen S, Trichopoulos D, Warren K, Nardi G: Coffee and cancer of the pancreas. N Engl J Med 1981; 304: pp. 630-633.
9. Feinstein AR, Horwitz RI, Spitzer WO, Battista RN: Coffee and pancreatic cancer. The problems of etiologic science and epidemiologic case-control research. JAMA 1981; 246: pp. 957-961.
10. Jekel JF: Should we stop using the P value in descriptive studies?. Pediatrics 1977; 60: pp. 124-126.
11. Jekel JF, Chauncey KJ, Moore NL, Broadus AE, Gowdy DP: The regional educational impact of a renal stone center. Yale J Biol Med 1983; 56: pp. 97-108.
12. Dawson B, Trapp RG: 4th ed2004.Lange Medical Books/McGraw-HillNew York, NY

OKUMA ÖNERİLERİ

Dawson B, Trapp RG: 4th ed2004.Lange Medical Books/McGraw-HillNew York, NY