

Bağımlı Değişken ile Ortak Değişken Arasındaki Korelasyona Göre Kovaryans Analizinin Gücü

The Power of Covariance Analysis According to the Correlation Between Dependent Variable and Covariate

- Emre DEMİR^a,
- Yasemin YAVUZ^b,
- Serdal Kenan KÖSE^b

^aBiyostatistik AD,
Hitit Üniversitesi

Tıp Fakültesi,
Çorum, TÜRKİYE

^bBiyostatistik AD,
Ankara Üniversitesi
Tıp Fakültesi,
Ankara, TÜRKİYE

Received: 26.03.2019

Received in revised form: 27.07.2019

Accepted: 21.08.2019

Available Online: 20.12.2019

Correspondence:

Emre DEMİR

Hitit Üniversitesi

Tıp Fakültesi,

Biyostatistik AD, Çorum,

TÜRKİYE/TURKEY

emredemir82@gmail.com

ÖZET Amaç: Kovaryans analizi (ANCOVA), bir çalışmada etkisi incelenen faktör dışında bağımlı değişken ile ilişkisi bulunan sürekli bir veya birden fazla değişkenin etkisinin istatistiksel olarak kontrol edilmesini sağlayan bir yöntemdir. Bu çalışmada randomize olan araştırmalarda ortak değişken ile sonuç değişkeni arasında korelasyon bulunan modellerde ortak değişkenin etkisinin göz ardı edildiği varyans analizi (ANOVA) ile ortak değişken etkisinin artıldığı ANCOVA'nın korelasyon düzeyine bağlı olarak gücünün değişimi karşılaştırılmıştır. **Gereç ve Yöntemler:** Araştırmamızda Cohen'in küçük, orta ve büyük f etki büyüklüklerine bağlı 3 farklı senaryo kullanılarak n=30, 50 ve 100 örneklem büyüklükleri için benzetim çalışması yapılmıştır. Her bir benzetimde bağımlı değişken ile ortak değişken arasında 0'dan 1'e kadar 0,1 birimlik artış ile 11 farklı korelasyon bulunacak şekilde veri seti üretilerek her senaryoda 5000 benzetim yapılmıştır. Benzetim uygulamaları için R paket programı kullanılmıştır. R paketinde veri üretmek için "MBESS" ve güç hesaplamaları için "pwr" açık kaynak kodlu kütüphaneler kullanılmıştır. **Bulgular:** Küçük ve orta etki büyüklüklerine sahip benzetim sonuçlarında sonuç değişkeni ile ortak değişken arasında 0,20 korelasyon bulunduğu ANOVA ile ANCOVA arasındaki güç değerinin ayrılmaya başladığı ve korelasyon değeri 0,50 ve sonrasında ANCOVA'nın ANOVA'ya göre çok daha fazla güce ulaştığı görülmüştür. **Sonuç:** Ortak değişken ile sonuç değişkeni arasında 0,2 ve üzeri korelasyon bulunduğu ANOVA yerine ANCOVA kullanılarak daha az sayıda örneklem büyüklüğü ile araştırmada yeterli güce ulaşılabilecektir.

Anahtar Kelimeler: Kovaryans analizi; ANCOVA; ortak değişken; korelasyon; güç analizi; benzetim

ABSTRACT Objective: Covariance analysis (ANCOVA) is a statistical procedure that enables one to compare groups on some quantitative dependent variable while simultaneously controlling for the effects of selected other continuous variables, which co-vary with the dependent. The control variables are called the "covariates". In this study, we aimed to compare the statistical power of variance analysis (ANOVA), in which covariate was omitted, and ANCOVA for different values of the correlation r between the covariate and the dependent variable in randomized studies. **Material and Methods:** In our study, a simulation was conducted for n = 30, 50 and 100 sample sizes by using 3 different scenarios of Cohen's small, medium and large f effect sizes. In each simulation, a data set was generated with 11 different correlations from 0 to 1 with an increase of 0.1 units between the dependent variable and the covariate, and 5000 simulations were made in each scenario. R package was used for simulation applications. The "MBESS" open source library was used to generate data and the "pwr" library was used for power analysis in the R package. **Results:** The results of simulation for small and medium effect sizes, showed that statistical power for ANOVA and ANCOVA started to decompose when the covariate-dependent variable correlation was higher than 0.2. ANCOVA reached much more statistical power than ANOVA when covariate-dependent variable correlation was higher than 0.5. **Conclusion:** The use of ANCOVA instead of ANOVA may considerably reduce the number of patients required when the covariate-dependent variable correlation was higher than 0.2.

Keywords: Covariance analysis; ANCOVA; covariate; correlation; power analysis; simulation

Nedensellik bağıntısının araştırıldığı çalışmalarda çalışma sonunda ulaşılan bulguların araştırılan etken dışında başka faktör/lere bağlı olup olmadığının belirlenmesi oldukça önemlidir. Çünkü araştırılan etkenin sonuç değişkeni üzerindeki etkisi ancak sonuç değişkeni ile ilişkili diğer değişken/lerin etkilerinin arıtılması sonucunda yansız olarak elde edilebilir.¹ Özellikle belirli bir müdahalenin ya da etkenin etkisinin araştırıldığı çalışmalarda maruziyetin yarattığı etkiyi vaka ve kontrol gruplarında en az denek sayısı ile karşılaştırma yaparak yansız olarak belirlemek temel amaçtır. Maruziyetin etkisinin araştırıldığı klinik denemelerde araştırma öncesi “Randomizasyon” yapılarak deney ve kontrol gruplarında hasta (yaş, cinsiyet v.b.) ve hastalık (evre, aile öyküsü v.b.) ile ilgili özelliklerin dengeli dağılması yoluyla yanlılığın önlenmesi varsayılır.²⁻⁴ Fakat pratikte özellikle denek sayısının küçük olduğu çalışmalarda bu varsayımın gerçekleşemediği durumlar ile karşılaşmaktadır. Belirli bir etkenin etkisinin araştırıldığı randomize olmayan gözlemsel çalışmalarda ise vaka-kontrol grupları birçok ortak değişken/faktör açısından benzer özelliklere sahip olamamaktadır.⁵ Dengeli dağılmayan ortak değişken (karıştırıcı değişken) ile karşılaştırılmak istenen bağımlı değişkenin (sonuç değişkeni) ilişkili olduğu durumlarda tedavi etkisini gösteren istatistiksel test sonuçları yanlı olarak elde edilebilir.⁶⁻⁸

Bu durumda “Müdahale Öncesi Düzeltme (pre adjustment)” ve “Müdahale Sonrası Düzeltme (post adjustment)” yöntemleri olmak üzere iki temel yaklaşımdan söz edilebilir. Müdahale öncesi düzeltme yöntemleri genellikle çalışma tasarımı sırasında randomizasyon, tabakalı randomizasyon, propensity skora dayalı eşleştirme, kısıtlama yoluyla ortak değişkenin gruplarda dengeli dağıtılması yoluyla gerçekleştirilirken, müdahale sonrası düzeltme, analiz evresinde; tabakalandırma, ANCOVA gibi çeşitli istatistiksel yöntemlerin kullanılması yoluyla yapılır. Böylece ortak değişkenin etkisinin arıtılmasını sağlayarak yansız sonuçlar elde edilmesine olanak sağlanır. Örneğin tedavilerin hastaların kolesterolü üzerine etkilerinin incelendiği bir çalışmada yaşın kolesterolü etkilediği bilimsel olarak kanıtlanmış ise yaş çalışmaya ortak değişken olarak alınmalı ve etkisi arındırılmalıdır.

Ön-test ölçümlerinin gruplarda benzer olmadığı Önce-Sonra (pre-post) karşılaştırmalarında ise ön-test ölçümleri ortak değişken olarak alınarak analizlerin yapılması ile başlangıç koşulları eşitlenmektedir.⁹ Örneğin yüzücülerde üç farklı yüzme stiline (kurbağa, kelebek ve serbest stil) biceps çevre uzunluğu artışına etkisinin araştırıldığı bir çalışmada, biceps çevre uzunlukları sadece yüzme stiline değil deneyin başındaki değerlerden de etkilenecektir. Dolayısıyla modele deneyin başındaki biceps çevre uzunlukları ortak değişken olarak alınmalı ve başlangıç değerlerinin etkisi arındırılarak deneme etkileri test edilmelidir.

Müdahale etkisinin sürekli sonuçlu bir değişken ile incelendiği çalışmalarda analiz evresinde en sık kullanılan yaklaşım grup ortalamalarını karşılaştıran istatistiksel testleri kullanmaktır. ANCOVA, bir çalışmada etkisi incelenen faktör dışında bağımlı değişken ile ilişkisi bulunan sürekli bir veya birden fazla değişkenin (covariate-ortak değişken) etkisinin arıtılarak istatistiksel olarak kontrol edilmesini sağlayan bir yöntemdir.^{10,11} ANCOVA yardımıyla sonuç değişkeni ile ilişkili diğer değişkenlerin analize alınarak etkilerinin arındırılması ve böylece daha az denek ile istenen güçte sonuç elde edilmesi mümkündür.¹²⁻¹⁵ ANCOVA grup içi hata varyansını küçülterek açıklanamayan hata varyansının bir kısmını deneye dâhil edilen diğer ortak değişkenler yoluyla açıklayarak hata varyansını azaltır ve F değerini büyütür. Özet bir ifadeyle ortak değişkenden kaynaklanan değişim ölçülebilir ve hata varyansından ayırt edilebilir dolayısıyla hata varyansının azaltılması ile testin gücü artar.¹⁶

ANCOVA tasarımında ortak değişken aralıklı veya oransal veriler içermeli ve hatasız bir şekilde ölçülmüş olmalıdır. ANOVA'nın önemli varsayımlarından grupların bağımsızlığı, normallik dağılımı ve varyansların homojenliği varsayımları ANCOVA'nda temel varsayımlarındandır.¹⁰⁻¹⁴ Bu varsayımlara ek olarak ANCOVA'da ortak değişken ile etkisi test edilen faktörün (bağımsız değişkenin) bağımsızlığı (ortak değişken araştırma gruplarından bağımsızdır) ve regresyon doğrularının eğimlerinin homojenliği varsayımları

rının sağlanması gerekmektedir. Bu varsayımda faktör-ortak değişken etkileşim etkisi olmamalı, grup içi regresyon katsayıları homojen olmalıdır. Ortak değişken gruplarda bağımlı değişken üzerinde aynı etkiye sahip olmalıdır. Özet bir ifadeyle ortak değişken ve bağımlı değişken arasındaki doğrusal ilişkinin gücü ve yönü bağımsız değişkenin düzeylerinde benzer olmalıdır. ANCOVA modelinde birden fazla ortak değişken kullanılacaksa da bunlar arasında da güçlü bir korelasyon bulunmamalıdır.^{10,17,18}

ANCOVA'nın önemli varsayımlarından biride ortak değişken ile bağımlı değişken arasında doğrusal bir korelasyon olmasıdır.¹⁰⁻¹⁴ Ancak bu korelasyonun düzeyi üzerinde literatürde yeterince çalışma bulunmamaktadır. Çalışmamızda nedensellik bağlantısı araştırılan Randomize araştırmalarda etkisi incelenen değişken dışında bağımlı değişken ile ilişkili ortak değişkenin etkisinin artırıldığı (ANCOVA ile) durumun bu ortak değişkenin göz ardı edildiği duruma göre (ANOVA ile) testin güç (power)'ünde ne kadarlık bir artış yarattığını göstermek amaçlanmıştır. Bu amaçla çalışmamızda korelasyonun düzeyine göre kovaryans analizinin gücünün nasıl değiştiği bir benzetim çalışması ile gösterilmiştir.

GEREÇ VE YÖNTEMLER

Uygulamada veri üretmek için R paket programında açık kaynak kodlu "MBESS" (Version 4.4.3; Kelley, 2018) ve güç hesaplamaları için "pwr" (Version 1.2-1; Champely, 2017) paketi kullanılmıştır.^{19,20} "pwr" paketi Cohen'in (1988) etki büyüklükleri ve notasyonlarını kullanarak temel güç hesaplamaları için bazı fonksiyonlara sahip bir pakettir.²¹

"MBESS" kütüphanesinde "*ancova.random.data*" fonksiyonu, ANCOVA varsayımlarının sağlandığı rastgele verileri üretmek için çok değişkenli normal dağılım kullanmakta ve ortak değişken modelde rastgele bir değişken olarak kabul edilmektedir. Bu işlev MASS paketindeki *mvnorm* parametresini kullandığı için MASS paketinin kurulmasını gerektirmektedir.²² Bu fonksiyon, hem randomize hem de randomize olmayan tasarımlar için gruplar arasında homojen kovaryans matrisini varsaymaktadır.

Tek yönlü ve tek ortak değişken bulunan ANCOVA modeli için gruplar arasında ortak değişken ortalamasının benzer olduğu randomize tasarım için *ancova.random.data* (*ancova.random.data(mu.y, mu.x, sigma.y, sigma.x, rho, J, n, randomized = TRUE)*) fonksiyonu kullanılarak çok değişkenli normal dağılıma uyan farklı etki büyüklüklerine dayalı senaryolarda veriler üretilmiştir.

Araştırmamızda Cohen'in küçük, orta ve büyük f etki büyüklüklerine bağlı 3 farklı senaryo kullanılarak n=30, 50 ve 100 örneklem büyüklükleri için benzetim çalışması yapılmıştır. Her bir benzetimde bağımlı değişken ile ortak değişken arasında 0'dan 1'e kadar 0,1 birimlik artış ile 11 farklı korelasyon bulunacak şekilde veri seti üretilerek her senaryoda 5000 benzetim yapılmıştır. Birinci benzetim çalışmasında Cohen'in önerdiği f=0,1 küçük etki büyüklüğünü elde edebilmek için bağımlı değişkenin gruplarda standart sapması 1 ve yaklaşık olarak ortalamaları sırasıyla 3; 3,1 ve 3,25, ortak değişkenin standart sapması 1 ve ortalaması 3 olacak şekilde çok değişkenli normal dağılımdan türetilmiştir (Senaryo 1). İkinci benzetim çalışmasında f=0,25 orta etki büyüklüğünü elde edebilmek için bağımlı değişkenin gruplarda standart sapması 1 ve yaklaşık olarak ortalamaları sırasıyla 3; 3,3 ve 3,6, ortak değişkenin standart sapması 1 ve ortalaması 3 olacak şekilde çok değişkenli normal dağılımdan türetilmiştir (Senaryo 2). Üçüncü benzetim çalışmasında ise f=0,4 büyük etki büyüklüğünü elde edebilmek için bağımlı değişkenin gruplarda standart sapması 1 ve yaklaşık olarak ortalamaları sırasıyla 3; 3,5 ve 4, ortak değişkenin standart sapması 1 ve ortalaması 3 olacak şekilde çok değişkenli normal dağılımdan türetilmiştir (Senaryo 3).

ANOVA ve ANCOVA benzetim verilerine bağlı güç değerlerini elde etmek için öncelikle Eşitlik 1'de verilen eta kare (η^2) hesaplanmıştır. Daha sonra eşitlik 2'de verilen Cohen'in f değeri hesaplanmış ve "pwr" paketinde Cohen'in f değerine bağlı güç değerleri elde edilmiştir.

$$\eta^2 = \frac{\text{kareler toplamı (gruplar arası)}}{\text{kareler toplamı (toplam)}} = \frac{\sum_{j=1}^p n_j (\bar{x}_j - \bar{x})^2}{\sum_{j=1}^p \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2} \quad (1)$$

$$\text{Cohen's } f = \sqrt{\frac{\eta^2}{1-\eta^2}} \quad (2)$$

Bu çalışma Helsinki Deklarasyonu 2013 prensiplerine uygun olarak yürütülmüştür.

BULGULAR

Benzetim çalışması öncesinde “*ancova.random.data*” fonksiyonu ile üretilen veri setleri için ANCOVA varsayımlarının sağlanıp sağlanmadığı test edilmiştir. Normal dağılım varsayımı Shapiro-Wilk testi, varyansların homojenliği varsayımı Levene testi, faktör ile ortak değişken bağımsızlığı ve regresyon eğrilerinin homojenliği varsayımları ANOVA testi ile değerlendirilmiş ve 3 farklı senaryo için üretilen tüm örneklem büyüklüklerinde varsayımların sağlandığı görülmüştür.

1. senaryoda n=30 için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla p=0,622, p=0,971, p=0,966, Levene test sonuçları F(2, 87) = 1,013: p = 0,367, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları F(2, 87) = 0,665: p = 0,517 ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları F(2, 84) = 0,757: p = 0,472 olarak bulunmuştur. n=50 için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla p=0,876, p=0,783, p=0,119, Levene test sonuçları F(2, 147) = 0,086: p = 0,918, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları F(2, 147) = 1,245: p = 0,291 ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları F(2, 144) = 0,814: p=0,445 olarak bulunmuştur. n=100 için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla p=0,941, p=0,996, p=0,645, Levene test sonuçları F(2, 297) = 1,06: p = 0,347, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları F(2, 297) = 0,944: p = 0,390 ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları F(2, 294) = 0,100: p=0,905 olarak bulunmuştur.

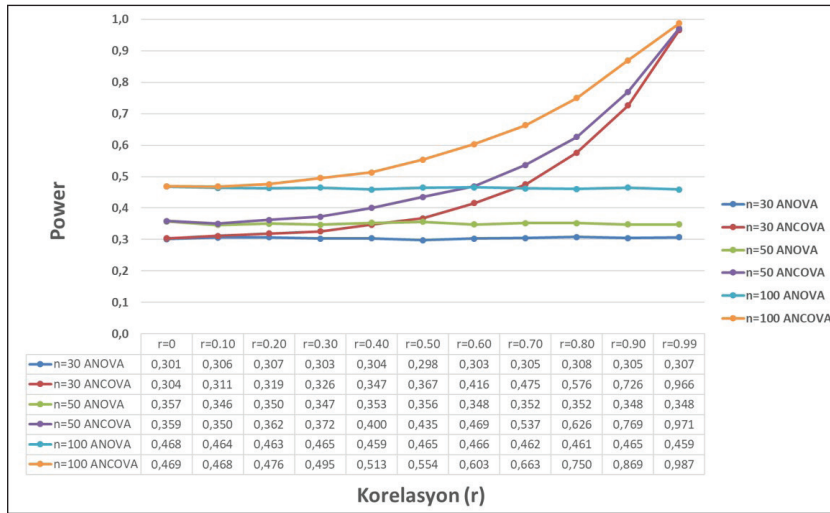
2. senaryoda n=30 için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla p=0,346, p=0,223, p=0,658, Levene test sonuçları F(2, 87) = 1,06: p = 0,350, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları F(2, 87) = 0,546: p = 0,581 ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları F(2, 84) = 0,037: p = 0,964 olarak bulunmuştur. n=50 için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla p=0,816, p=0,397, p=0,721, Levene test sonuçları F(2, 147) = 0,251: p = 0,778, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları F(2, 147) = 0,961: p = 0,385 ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları F(2, 144) = 0,937: p=0,394 olarak bulunmuştur. n=100 için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla p=0,767, p=0,705, p=0,162, Levene test sonuçları F(2, 297) = 1,534: p = 0,217, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları F(2, 297) = 0,054: p = 0,947 ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları F(2, 294) = 1,058: p=0,348 olarak bulunmuştur.

3. senaryoda n=30 için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla p=0,756, p=0,785, p=0,241, Levene test sonuçları F(2, 87) = 0,216: p = 0,806, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları F(2, 87) = 0,803: p = 0,451 ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları F(2, 84) = 0,517: p = 0,598 olarak bulunmuştur. n=50 için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla p=0,634, p=0,776, p=0,508, Levene test sonuçları F(2, 147) = 0,370: p = 0,691, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları F(2, 147) = 1,263: p = 0,286 ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları F(2, 144) = 0,949: p=0,389

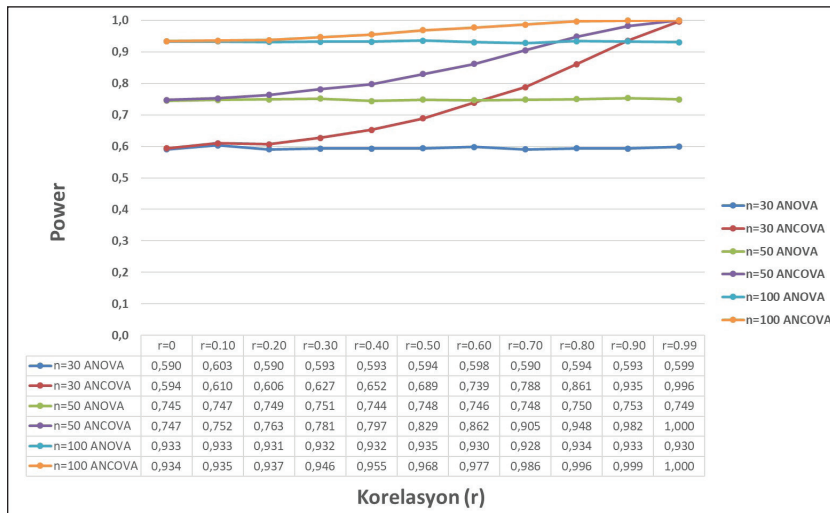
olarak bulunmuştur. $n=100$ için Shapiro-Wilk test sonuçları 3 farklı grup için sırasıyla $p=0,495$, $p=0,949$, $p=0,722$, Levene test sonuçları $F(2, 297) = 0,239$; $p = 0,787$, faktör ile ortak değişken bağımsızlığı için ANOVA test sonuçları $F(2, 297) = 0,026$; $p = 0,974$ ve regresyon eğrilerinin homojenliği varsayımı için etkileşime ilişkin ANOVA test sonuçları $F(2, 294) = 0,731$; $p=0,482$ olarak bulunmuştur.

Sonuç olarak $n=30, 50$ ve 100 örneklem büyüklüğünde üretilen benzetim verilerinde ANCOVA varsayımlarının sağlandığı görülmüştür.

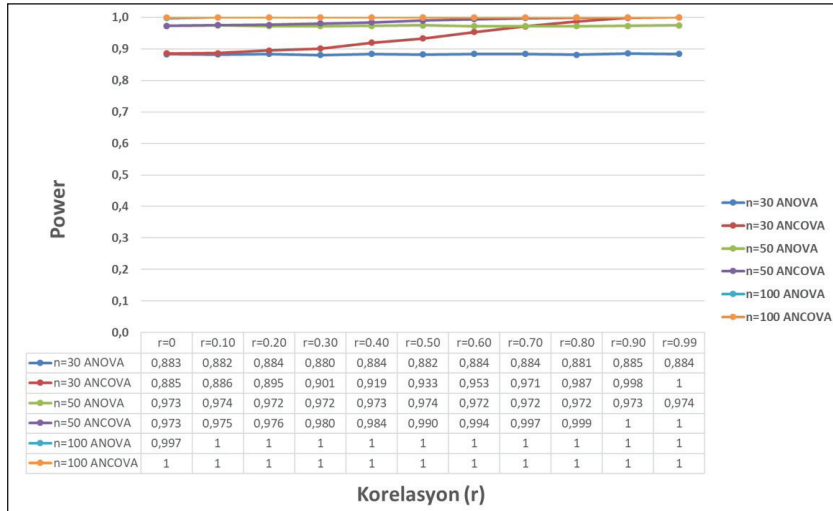
5000 benzetim sonucunda ANOVA ve ANCOVA güç analizi sonuçları senaryo 1 için Şekil 1’de, senaryo 2 için Şekil 2’de ve senaryo 3 için Şekil 3’te verilmiştir.



ŞEKİL 1: Cohen $f=0,10$ küçük etki büyüklüğü için $n=30, 50$ ve 100 örneklem büyüklüğüne göre gücün değişimi ($\mu_1 \approx c(3; 3, 1; 3, 25)$, $\sigma_1 = 1$, $\mu_2 = 3$, $\sigma_2 = 1$)



ŞEKİL 2: Cohen $f=0,25$ orta etki büyüklüğü için $n=30, 50$ ve 100 örneklem büyüklüğüne göre gücün değişimi ($\mu_1 \approx c(3; 3, 3; 3, 6)$, $\sigma_1 = 1$, $\mu_2 = 3$, $\sigma_2 = 1$)



ŞEKİL 3: Cohen f=0,40 büyük etki büyüklüğü için n=30, 50 ve 100 örneklem büyüklüğüne göre güç değişimi (mu.y=c(3; 3,5; 4), sigma.y=1, mu.x=3, sigma.x=1)

Şekil 1'e göre ANOVA kullanıldığında n=30 için güç değerlerinin %30 civarında, n=50 için güç değerlerinin %35 civarında ve n=100 için güç değerlerinin %46 civarında olduğu görülecektir. ANOVA yerine ANCOVA kullanıldığında ise n=30 için güç değerlerinin %30 civarında başlayıp korelasyon değeri 1'e yaklaştıkça gücün %97'ye yaklaştığı, n=50 için güç değerlerinin %35 civarında başlayıp korelasyon değeri 1'e yaklaştıkça gücün %97'ye yaklaştığı, n=100 için güç değerlerinin %47 civarında başlayıp korelasyon değeri 1'e yaklaştıkça gücün %99'a yaklaştığı görülecektir.

Şekil 2'ye göre ANOVA kullanıldığında n=30 için güç değerlerinin %60 civarında, n=50 için güç değerlerinin %75 civarında ve n=100 için güç değerlerinin %93 civarında olduğu görülecektir. ANOVA yerine ANCOVA kullanıldığında ise n=30 için güç değerlerinin %60 civarında başlayıp korelasyon değeri 1'e yaklaştıkça gücün %100'e yaklaştığı, n=50 için güç değerlerinin %75 civarında başlayıp korelasyon değeri 1'e yaklaştıkça gücün %100 olduğu, n=100 için güç değerlerinin %93 civarında başlayıp korelasyon değeri 1'e yaklaştıkça gücün %100 olduğu görülecektir.

Şekil 3'e göre ANOVA kullanıldığında n=30 için güç değerlerinin %88 civarında, n=50 için güç değerlerinin %97 civarında ve n=100 için güç değerlerinin %100 civarında olduğu görülecektir. ANOVA yerine ANCOVA kullanıldığında ise n=30 için güç değerlerinin %88 civarında başlayıp korelasyon değeri 1'e yaklaştıkça gücün %100'e yaklaştığı, n=50 için güç değerlerinin %97 civarında başlayıp korelasyon değeri 1'e yaklaştıkça gücün %100 olduğu, n=100 için güç değerlerinin tüm korelasyonlarda %100 olduğu görülecektir.

TARTIŞMA

ANCOVA ve ANOVA yöntemleri gruplar arasında sürekli bir ölçüm açısından istatistiksel olarak anlamlı farklılığın araştırılması amacıyla kullanılan testlerdir. Literatürde ANCOVA genel olarak iki amaçla kullanılmaktadır. Bunlardan birincisi bireylerin deney gruplarına rastgele atandığı klinik denemelerde, müdahale öncesi değerlerin ortak değişken olarak alınarak hata varyansının azaltılması yoluyla istatistiksel gücün artırılmasının amaçlandığı çalışmalardır. Böylece araştırılan etkinin tahmininde daha fazla hassasiyet elde edilmiş olur. İkincisi ise, gruplara rastgele atanmanın yapılmadığı gözlemsel çalışmalarda, ortak değişkenin grup değişkeniyle ilişkili olduğu durumlarda ANCOVA ile ortak değişkenin karıştırıcı (confounder) etkisinin arındırılmasının amaçlandığı çalışmalardır.^{14,23} Ancak bazı araştırmacılar ikinci

amaç için ANCOVA kullanımının uygun bir yöntem olmadığını iddia etmektedirler.²⁴ Literatürde bu amaç için Propensity skora dayalı eşleştirme analizinin kullanılması önerilmiştir. Çalışmamızda randomize çalışmalarda ortak değişken ile sonuç değişkeni arasındaki (örneğin ön-test ölçümü ile etkisi araştırılan son-test ölçümü arasındaki korelasyon) korelasyon düzeyine bağlı olarak gücün değişimi incelenmiştir.

Van Breukelen (2006) hastaların tedavi ve kontrol gruplarına randomize atıldığı bir çalışma tasarlamış ve tedavinin etkinliğini tedavi öncesi değerleri ortak değişken olarak alıp ANCOVA ve tedavi öncesi ve sonrası fark değerlerini hesaplayıp t-testi ile karşılaştırmıştır. İki testin gücü karşılaştırıldığında ANCOVA'nın t-testine göre daha fazla güce sahip olduğu sonucuna varılmıştır.²⁵

Borm ve ark. (2007), müdahale etkinliğinin incelendiği çalışmalarda, gruplarda müdahale öncesi ve sonrası ölçüm farklarının karşılaştırılması yerine başlangıç ölçümlerinin ortak değişken olarak analize ANCOVA ile dâhil edilmesinin gereken hasta sayısını önemli ölçüde azaltabileceğini ifade etmişler ve çalışma öncesi örneklem büyüklüğü hesaplamalarında bunun dikkate alınmasını önermişlerdir.¹³

Egbewale ve ark. (2014), başlangıç ölçümlerinin dengesiz olduğu randomize denemelerde Cohen'in küçük, orta ve büyük etki büyüklüklerine dayalı olarak belirledikleri tedavi etkisi ile farklı senaryolar oluşturarak gerçekleştirdikleri benzetim çalışması sonucunda ANOVA ve ANCOVA'nın gücünü karşılaştırmışlar ve özellikle ön-test son-test ölçümleri arasında 0,3'ten daha büyük korelasyon bulunduğu yanlılığı azaltmak için ANOVA yerine ANCOVA'nın kullanılmasını önermişlerdir. Orta ve yüksek düzeyde korelasyon bulunduğu ise ANOVA kullanımına karşı dikkat edilmesi gerektiğini ifade etmişlerdir.²⁴

Wu ve Lai (2015), korelasyon olmadığı durumu, zayıf, orta ve yüksek korelasyonu temsil edecek şekilde ön-test ve son-test skorları arasında sırasıyla 0; 0,3; 0,6 ve 0,8 korelasyon bulunacak şekilde ürettikleri veri ile yapmış oldukları benzetim çalışması sonucunda ön-test ve son-test ölçümleri arasında korelasyon bulunmadığında ANCOVA ve ANOVA'nın benzer güç ürettiği sonucuna ulaşmışlardır. Ancak korelasyon bulunduğu durumlarda ANCOVA'nın ANOVA'dan daha güçlü olduğu, ön-test ve son-test ölçümleri arasındaki korelasyon arttıkça ANCOVA ile güç kazanımının arttığı benzetim çalışması ile gösterilmiştir. Bu çalışmada aykırı veri bulunmayan normal dağılım verileri için korelasyon bulunmadığında ($r=0$) hem ANOVA hemde ANCOVA'nın %43 ile %46 civarında güce ulaştığı, korelasyon düzeyinin 0,8 olduğunda ise ANOVA %43 ile %46 güce ulaşırken ANCOVA'nın %63 ile %67 güce ulaştığı raporlanmıştır. Sonuç olarak korelasyonun varlığındaki istatistiksel güç kazancı ve tip I hatanın daha iyi korunması nedeniyle ön-test son-test tasarımlarında ANCOVA'nın kullanılması önerilmiştir. Araştırma sonuçları çalışmamızın sonuçları ile benzerdir. Ayrıca bizim çalışmamızda, bu çalışmada olduğu gibi benzetimler Cohen'in etki büyüklüklerine dayalı olarak gerçekleştirilmiştir.²⁶

Vickers ve Altman (2001) ve Venter ve ark. (2002)'da ANCOVA'nın t testine göre daha güçlü olduğunu ve araştırmalarda daha az örneklem büyüklüğü ile çalışılabilmesine olanak sağlayacağını belirtmişlerdir.^{27,28} Literatürde genellikle korelasyonun gücünden örneklem büyüklüğünün nasıl etkilendiği gösterilmeye çalışılmış ve bazı çalışmalarda korelasyon değeri arttıkça ANCOVA kullanılması ile daha az örneklem büyüklüğü ile araştırmanın yapılacağı sonucuna ulaşılmıştır.^{15,29,30} Çalışmamızda ise farklı etki ve örneklem büyüklüklerinde artan korelasyon değerlerinin çalışmanın gücünü nasıl değiştirdiği incelenmiş ve hangi değerden sonra farklılık gözlemlendiği belirlenmeye çalışılmıştır.

SONUÇ

Çalışmamızın sonuçları genel olarak literatürdeki çalışmalarla paralellik göstermekte olup özellikle küçük ve orta etki büyüklüğüne sahip bir araştırmada sonuç değişkeni ile ortak değişken arasında 0,20 korelasyon bulunduğu ANOVA ile ANCOVA arasındaki güç değerinin ANCOVA lehine ayrılmaya

başladığı ve korelasyon değeri 0,50 ve üzerine çıktığında ANCOVA ile daha güçlü sonuçlar elde edildiği görülmüştür. Daha genel bir ifadeyle araştırmanın amacına uygun istatistiksel yöntemlerin seçimi araştırmacıların daha az sayıda örneklem büyüklüğü ile yeterli güce ulaşmalarını sağlayacak bu da daha az maliyet ve daha kısa sürede araştırmaların sonlandırılmasına fırsat verecektir.

Finansal Kaynak

Bu çalışma sırasında, yapılan araştırma konusu ile ilgili doğrudan bağlantısı bulunan herhangi bir ilaç firmasından, tıbbi alet, gereç ve malzeme sağlayan ve/veya üreten bir firma veya herhangi bir ticari firmadan, çalışmanın değerlendirme sürecinde, çalışma ile ilgili verilecek kararı olumsuz etkileyebilecek maddi ve/veya manevi herhangi bir destek alınmamıştır.

Çıkar Çatışması

Yazarlar herhangi bir çıkar çatışması veya finansal destek bildirmemiştir.

Yazar Katkıları

Fikir/Kavram: Emre Demir; **Tasarım:** Emre Demir, Yasemin Yavuz, S. Kenan Köse; **Denetleme/Danışmanlık:** Emre Demir, Yasemin Yavuz, S. Kenan Köse; **Veri Toplama Ve/Veya İşleme:** Emre Demir; **Analiz Ve/Veya Yorum:** Emre Demir, Yasemin Yavuz, S. Kenan Köse; **Kaynak Taraması:** Emre Demir; **Makalenin Yazımı:** Emre Demir, Yasemin Yavuz, S. Kenan Köse; **Eleştirel İnceleme:** Emre Demir, Yasemin Yavuz, S. Kenan Köse.

KAYNAKLAR

1. Kaspar EÇ, Bekiroğlu N, Genceli M. [Propensity score in observational studies and an application in medical sciences]. Türkiye Klinikleri J Biostat. 2010;2(1):1-10.
2. Rosenberger WF, Lachin JM. Randomization in Clinical Trials: Theory and Practice. 1st ed. New York: John Wiley; 2002. p.288. [Crossref]
3. Kanik A, Taşdelen B, Erdoğan S. [Randomization in clinical trials]. Marmara Medical Journal. 2011;24(3):149-55. [Crossref]
4. Gökmen D, Alkan A, Bakırarar B, Siddikoğlu D, Özgür EG, Özalp Ateş FS, et al. Bilimsel Araştırma Yöntemleri. 1. Baskı. Ankara: Ankara Üniversitesi Basımevi; 2018. p. 130.
5. D'agostino RB. Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statist Med. 1988;17:2265-81. [Crossref]
6. Hill HA, Kleinbaum DG. Bias in observational studies. In: Armitage P, ed. Encyclopedia of Biostatistics. 2nd ed. Chichester: John Wiley and Sons; 2005. 323-330. [Crossref]
7. Hoffmeister H. Bias in observational studies. In: Hoffmeister H, Szklo M, Thamm M. (eds) Epidemiological Practices in Research on Small Effects. Springer, Berlin, Heidelberg; 1998. p. 59-69. [Crossref]
8. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc. 1984;79(387):516-24. [Crossref]
9. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology. 1974;66(5):688-701. [Crossref]
10. Leppink, J. Analysis of covariance (ANCOVA) vs. moderated regression (MODREG): why the interaction matters. Health Professions Education. 2018;4(3):225-32. [Crossref]
11. Kim HY. Statistical notes for clinical researchers: analysis of covariance (ANCOVA). Restor Dent Endod. 2018;43(4):e43. PMID: 30483467 [Crossref] [PubMed] [PMC]
12. Tang Y. Exact and approximate power and sample size calculations for analysis of covariance in randomized clinical trials with or without stratification. Statistics in Biopharmaceutical Research. 2018;10(4):274-86. [Crossref]
13. Borm GF, Fransen J, Lemmens WA. A simple sample size formula for analysis of covariance in randomized clinical trials. J Clin Epidemiol. 2007;60(12):1234-8. PMID: 17998077 [Crossref] [PubMed]
14. Egbewale BE, Lewis M, Sim J. Bias, precision and statistical power of analysis of covariance in the analysis of randomized trials with baseline imbalance: a simulation study. BMC Med Res Methodol. 2014;14:49. PMID: 24712304 [Crossref] [PubMed] [PMC]
15. Shieh G. On tests of treatment-covariate interactions: an illustration of appropriate power and sample size calculations. PLoS One. 2017;12(5):e0177682. PMID: 28545117 [Crossref] [PubMed] [PMC]
16. Algina J, Olejnik S. Conducting power analyses for ANOVA and ANCOVA in between-subjects designs. Eval Health Prof. 2003;26(3):288-314. PMID: 12971201 [Crossref] [PubMed]
17. Kraemer HC. A source of false findings in published research studies: adjusting for covariates. JAMA Psychiatry. 2015;72(10):961-2. PMID: 26244634 [Crossref] [PubMed]

18. Levy KJ. A Monte Carlo study of analysis of covariance under violations of the assumptions of normality and equal regression slopes. *Educational and Psychological Measurement (EPM)*. 1980;40(4):835-40. [[Crossref](#)]
19. Kelley K. MBESS (Version 4.0.0 and higher) [computer software and manual]. 2017. <http://cran.r-project.org>.
20. Champely S, Ekstrom C, Dalgaard P, Gill J, Weibelzahl S, Anandkumar A, et al. pwr: Basic Functions for Power Analysis 1.2-1. 2017. p.2-22.
21. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988. p.567.
22. Venables WN, Ripley BD. *Modern Applied Statistics with S*. 4th ed. New York, NY: Springer; 2002. p.495. [[Crossref](#)]
23. Keselman HJ, Huberty CJ, Lix LM, Olejnik S, Cribbie RA, Donahue B, et al. Statistical practices of educational researchers: an analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research (RER)*. 1998;68(3):350-86. [[Crossref](#)]
24. Miller GA, Chapman JP. Misunderstanding analysis of covariance. *J Abnorm Psychol*. 2001;110(1):40-8. PMID: 11261398 [[Crossref](#)] [[PubMed](#)]
25. Van Breukelen GJ. ANCOVA versus change from baseline: more power in randomized studies and more bias in nonrandomized studies. *J Clin Epidemiol*. 2006;59(9):920-5. PMID: 16895814 [[Crossref](#)] [[PubMed](#)]
26. Wu XW, Lai D. Comparison of statistical methods for pretest-posttest designs in terms of type I error probability and statistical power. *Communications in Statistics-Simulation and Computation*. 2015;44(2):284-94. [[Crossref](#)]
27. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. *BMJ*. 2001;323(7321):1123-4. PMID: 11701584 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
28. Venter A, Maxwell SE, Bolig E. Power in randomized group comparisons: the value of adding a single intermediate time point to a traditional pretest-posttest design. *Psychol Methods*. 2002;7(2):194-209. PMID: 12090410 [[Crossref](#)] [[PubMed](#)]
29. Clifton L, Birks J, Clifton DA. Comparing different ways of calculating sample size for two independent means: a worked example. *Contemporary Clinical Trials Communications*. 2019;13:100309. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
30. Wan F. Analyzing pre-post randomized studies with one post-randomization score using repeated measures and ANCOVA models. *Stat Methods Med Res*. 2019;28(10-11):2952-74. PMID: 30084297 [[Crossref](#)] [[PubMed](#)]