

Sağlık Bilimlerinde Güç Analizi: Kavramsal ve Uygulamalı Bir Yaklaşım

Fulden Cantaş Türkış¹

Özet

Bu bölüm, sağlık bilimlerinde güç analizinin temel kavramlarını, uygulama biçimlerini ve klinik araştırmalara olan etkilerini kapsamlı şekilde ele almaktadır. Araştırma bulgularının geçerliliği ve etik uygunluğu, yalnızca analiz yöntemlerine değil, aynı zamanda uygun örneklem büyüklüğüne de bağlıdır. Güç analizi, istatistiksel anlamlılığa ulaşmak için gerekli minimum örneklem büyüklüğünü belirlemeye yönelik temel bir planlama aracıdır. Etki büyüklüğü, anlamlılık düzeyi, hedeflenen istatistiksel güç ve test türü gibi unsurlar güç analizinin temel bileşenleridir. Bölümde; a priori, posterior ve duyarlılık analizleri gibi farklı güç analizi türleri açıklanmış, bağımsız ve bağımlı grup karşılaştırmaları ile varyans analizi örnekleri üzerinden uygulamaları G*Power yazılımı ile detaylandırılmıştır. Ayrıca PASS, R ve G*Power gibi yazılımların karşılaştırması yapılmış, bu araçların avantajları ve kullanım alanları sunulmuştur. Güç analizinin klinik araştırmalardaki rolü, yalnızca istatistiksel yeterliliği değil, aynı zamanda hasta güvenliğini ve kaynak kullanımını da doğrudan etkilediği için vurgulanmıştır. Etki büyüklüğünün gelişigüzel belirlenmesi, test türüne uygun olmayan analiz yapılması gibi sık karşılaşılan hatalar da örneklerle açıklanmıştır. Son olarak, güç analizinin bilimsel, etik ve uygulamalı sorumluluklar içeren bütüncül bir planlama sürecinin parçası olduğu ve disiplinlerarası araştırmalarda ortak bir dil oluşturması gerektiği önerilmiştir.

1. Giriş

Sağlık bilimlerinde yapılan araştırmalarda ulaşılan sonuçların bilimsel geçerliliği ve güvenilirliği, yalnızca analiz aşamasında uygulanan istatistiksel yöntemlere değil, aynı zamanda bu yöntemlerin hangi büyüklükteki örneklem üzerinde uygulandığına da bağlıdır. Bir araştırmanın örneklem büyüklüğü

1 Dr. Öğr. Üyesi, Muğla Sıtkı Koçman Üniversitesi Tıp Fakültesi Biyoistatistik A.D., fuldencantas@mu.edu.tr, 0000-0002-7018-7187

yeterli değilse, klinik olarak anlamlı olabilecek farkların istatistiksel olarak gösterilememesi riski doğabilir. Bu durum, var olan etkilerin gözden kaçırılmasına ve hatalı bir şekilde “anlamsız” olarak yorumlanmasına yol açabilir. Öte yandan, gereğinden büyük örneklemle çalışılması durumunda hem zaman ve kaynak israfı yaşanmakta, hem de araştırmaya dahil edilen bireyler açısından gereksiz müdahale riski oluşmaktadır. Bu nedenle, bir çalışmanın planlama sürecinde uygun örneklem büyüklüğünün belirlenmesi, yalnızca metodolojik değil aynı zamanda etik bir gereklilik olarak da değerlendirilmektedir. (Dell vd., 2002; Biau vd., 2010)

Örneklem büyüklüğünün belirlenmesinde sıklıkla başvuru alan temel yaklaşımlardan biri güç analizidir. Güç analizi, bir araştırma sonucunun istatistiksel olarak anlamlı kabul edilebilmesi için kaç bireyin çalışmaya dahil edilmesi gerektiğini tahmin etmeye yönelik bir tekniktir. Bu analiz gerçekleştirilirken genellikle dört temel unsur göz önünde bulundurulmaktadır: (i) karşılaştırılmak istenen gruplar veya değişkenler arasındaki beklenen etki büyüklüğü, (ii) bu farkın istatistiksel olarak anlamlı kabul edilmesi için anlamlılık düzeyi, (iii) bu farkın gerçekte var olduğu durumda testin bunu hedeflenen güç ve (iv) araştırmanın istatistiksel modeli veya hipotez testinin türü. Ancak bu teknik kavramların doğrudan kullanımı çoğu zaman klinik araştırmacılar için soyut ve karmaşık olabilmektedir. Bu nedenle, güç analizinin anlaşılabilir bir şekilde sunulması ve sağlık araştırmalarında nasıl uygulanması gerektiğinin açıkça gösterilmesi önem taşımaktadır (Biau vd., 2010).

Klinik araştırmalarda karar verme süreçleri sıklıkla sınırlı kaynaklar, zaman baskısı ve hasta güvenliği gibi faktörler ile şekillendirilmektedir. Bu çerçevede, çalışmalardan elde edilen verilerin güvenilirliğini artırmak ve klinik uygulamalara yansıtılabilir bulgular elde etmek için sağlam temellere dayanan bir örneklem planlamasının yapılması gerekmektedir. Güç analizi, bu planlamanın merkezinde yer almakta ve araştırmanın en başında yapılması gereken temel adımlardan biri olarak kabul edilmektedir. Özellikle yeni bir tedavi yönteminin, tanı aracının veya klinik uygulamanın etkisinin değerlendirilmek istendiği çalışmalarda, yeterli sayıda bireyin araştırmaya dahil edilip edilmediği, araştırmanın sonucunu doğrudan etkileyen belirleyici bir faktör haline gelmektedir.

Sağlık bilimlerinde yürütülen araştırmaların çok disiplinli bir yapı kazanmasıyla birlikte, istatistiksel planlamaya yönelik farkındalık ve yetkinlik ihtiyacı da artmaktadır. Klinik uzmanlar, hemşireler, halk sağlığı araştırmacıları ve biyoistatistikçiler tarafından ortaklaşa yürütülen araştırmalarda, güç analizinin doğru biçimde yapılması sayesinde hem etik standartlara uygunluk

sağlanmakta hem de elde edilen sonuçların bilime ve uygulamaya katkısı artırılmaktadır. Bu bağlamda, güç analizinin yalnızca teknik bir hesaplama aracı olarak değil, aynı zamanda araştırma kalitesini artırmaya yönelik stratejik bir yaklaşım olarak değerlendirilmesi gerekmektedir.

Bu bölümde, güç analizinin temel kavramları ile güç analizi türleri açıklanacak, farklı araştırma türlerine göre güç analizi ve temel istatistiksel testler için nasıl uygulandığı adım adım ele alınacaktır. Ayrıca, güç analizinin hangi yazılımlar aracılığıyla yapılabildiği tanıtılacak ve sağlık bilimlerinde çalışan araştırmacıların, çalışmalarının bilimsel niteliğini artırmaya yönelik pratik bilgiler edinmesi hedeflenecektir.

2. Güç Analizinin Temel Kavramları

Bir araştırmanın planlama aşamasında, elde edilmek istenen sonuçların güvenilirliğini ve yorumlanabilirliğini sağlamak amacıyla belirli istatistiksel kavramlara dayalı olarak örneklem büyüklüğü hesaplanmaktadır. Bu hesaplamaların temelinde ise güç analizi yer almaktadır. Güç analizinin doğru bir biçimde uygulanabilmesi için bazı temel kavramların anlaşılması gerekmektedir. Bu kavramlar, araştırma sürecinin istatistiksel temellerini oluşturmakta ve örneklem büyüklüğü ile araştırma sonuçları arasındaki ilişkiyi açıklamaktadır.

2.1. Etki Büyüklüğü

Etki büyüklüğü (Effect size), araştırma kapsamında incelenen iki grup ya da değişken arasındaki farkın veya ilişkinin büyüklüğünü nicel olarak ifade eden bir ölçüdür. Bu büyüklük, istatistiksel anlamlılığın ötesinde, bulgunun klinik ya da pratik önemini de yansıtmaktadır. Örneğin bir tedavi uygulamasının, mevcut yöntemle karşılaştırıldığında ne kadar farklı bir sonuç ürettiği, etki büyüklüğü ile ifade edilir. Güç analizlerinde, etki büyüklüğü genellikle önceden belirlenmekte veya daha önce yapılmış benzer çalışmalardan elde edilen değerler temel alınarak tahmin edilmektedir. Etki büyüklüğünün büyümesi durumunda, istatistiksel olarak anlamlı bir sonuca ulaşmak için daha küçük bir örneklem yeterli olabilmektedir (Cohen, 1988; Bulut, 2023).

2.2. Anlamlılık Düzeyi

Anlamlılık düzeyi (α), araştırma sonucunda “fark vardır” şeklinde bir sonuca ulaşılrken, gerçekte böyle bir fark bulunmama olasılığını ifade etmektedir. Genellikle %5 olarak kabul edilen bu değer, araştırmacı tarafından önceden belirlenmektedir. Bu düzey, elde edilen sonucun tesadüfen ortaya

çıkma olasılığını kontrol altında tutmak amacıyla kullanılmaktadır. Anlamlılık düzeyinin düşürülmesi, daha kesin sonuçlar elde edilmesini sağlamakla birlikte, aynı zamanda daha büyük bir örnekleme ihtiyaç duyulmasına da neden olabilmektedir (Fisher, 1970; Kul, 2011).

2.3. Güç

Güç ($1-\beta$), araştırmada var olan gerçek bir farkın ya da ilişkinin istatistiksel olarak saptanabilme olasılığıdır. Yüksek güç değerleri, araştırmanın duyarlılığını artırmakta ve yanlış şekilde “fark yok” sonucuna ulaşma riskini azaltmaktadır. Literatürde genellikle en az %80 düzeyinde güç hedeflenmektedir. Bu, var olan bir etkinin %80 olasılıkla tespit edilebileceği anlamına gelir. Gücün yetersiz olması durumunda, araştırmanın sonuçları geçersiz ya da yanıltıcı olabilir. Gücü artırmak için ya örneklem büyüklüğü artırılmalı ya da etki büyüklüğüne ilişkin daha belirgin bir fark hedeflenmelidir (Cohen, 1992).

2.4. Gruplara Dağıtım Oranı

Araştırmalarda örneklem büyüklüğünün belirlenmesinde kritik bir faktör olan gruplara dağıtım oranı (Allocation Ratio, N_2/N_1), katılımcıların deney ve kontrol gruplarına hangi oranda atanacağını ifade eder. Bu oran genellikle deney grubu için bir, kontrol grubu için ise iki veya üç katılımcı şeklinde ayarlanır (Örneğin: 1:2 veya 1:3). Optimal güç elde etmek için 1:1'den farklı oranlar kullanılabilir; özellikle deney grubu büyüklüğü, kontrol grubundan daha fazla sayıda katılımcı gerektiren durumlarda bu önemlidir (Kang vd., 2008).

2.5. Örneklem Büyüklüğü

Güç analizinin çıktısı olarak elde edilen örneklem büyüklüğü (sample size), araştırmaya dahil edilmesi gereken birey sayısını göstermektedir. Bu sayı, yukarıda açıklanan üç temel kavrama (etki büyüklüğü, anlamlılık düzeyi ve güç) bağlı olarak değişmektedir. Örneklem büyüklüğünün eksik ya da fazla olması, araştırma bulgularının bilimsel geçerliliğini ve klinik uygulamalara katkısını olumsuz yönde etkileyebilmektedir. Bu nedenle, örneklem büyüklüğünün doğru tahmin edilmesi, yalnızca istatistiksel bir gereklilik değil, aynı zamanda etik bir zorunluluk olarak kabul edilmektedir (Lenth, 2001).

2.6. Testin Türü ve Araştırma Deseni

Güç analizinde dikkate alınması gereken bir diğer unsur ise kullanılacak olan istatistiksel testin türüdür. Örneğin iki grubun ortalamalarının

karşılaştırılacağı bir çalışmada kullanılacak yöntem ile bir gruba uygulanan tedavi etkisinin araştırıldığı eş ölçümlere dayalı bir çalışmada gerekecek örneklem büyüklüğü farklılık gösterecektir. Ayrıca, çalışmanın kesitsel mi, uzunlamasına mı, deneysel mi yoksa gözlemsel mi olduğu gibi metodolojik özellikler de güç analizini etkileyen faktörler arasında yer almaktadır (Kul, 2011).

3. Güç Analizi Türleri

Güç analizi, araştırmanın planlandığı evreye, veri toplama sürecine ve elde edilen bulgulara göre farklı biçimlerde uygulanabilmektedir. Her bir güç analizi türü, farklı amaçlara hizmet etmekte olup, araştırma sürecinin farklı aşamalarında kullanılmaktadır. Bu bağlamda güç analizleri; önceden yapılan (a priori), veri toplandıktan sonra yapılan (posterior, post hoc) ve duyarlılık (sensitivity) analizleri olmak üzere üç temel grupta incelenmektedir.

3.1. A Priori Güç Analizi

A priori güç analizi, araştırmanın yürütülmesinden önce gerçekleştirilen ve çalışmada yer alacak örneklem büyüklüğünün belirlenmesini amaçlayan analiz türüdür (Cohen, 1988). Bu yaklaşımda araştırmacı, klinik açıdan anlamlı kabul edilen bir fark veya ilişki düzeyini (etki büyüklüğü), kabul edilebilir anlamlılık düzeyini ve hedeflenen istatistiksel gücü önceden belirlemelidir. Belirtilen bu değerler doğrultusunda, analizde ihtiyaç duyulan minimum örneklem büyüklüğü hesaplanmaktadır (Özdamar, 2017).

Klinik araştırmalarda etik standartların sağlanabilmesi, gereğinden fazla sayıda bireyin gereksiz yere araştırmaya dahil edilmesinin önlenmesi ve araştırmanın maliyet etkin biçimde yürütülebilmesi açısından a priori güç analizine başvurulması önerilmektedir. Özellikle deneysel çalışmalar ve müdahale araştırmaları gibi bireysel düzeyde etkilerin değerlendirildiği tasarımlarda, bu analiz yöntemi yaygın olarak kullanılmaktadır.

3.2. Posterior Güç Analizi

Posterior güç analizi, veriler toplandıktan ve analizler tamamlandıktan sonra yapılan ve elde edilen sonuca ilişkin istatistiksel gücün geriye dönük olarak hesaplandığı analiz türüdür. Bu yaklaşım, çalışmada anlamlı bir fark ya da ilişki saptanamamışsa, bunun gerçekten bir etkisizlik durumunu mu yoksa yetersiz örneklem büyüklüğü nedeniyle mi oluştuğunu değerlendirmek amacıyla kullanılmaktadır (Görgülü ve Ergin, 2023; Bulut, 2023).

Ancak bu analiz türünün yorumlanmasında dikkatli olunmalıdır. Çünkü posterior güç, elde edilen p değerine doğrudan bağlıdır ve anlamlı

olmayan sonuçlarda genellikle düşük çıkmaktadır. Bu nedenle, posterior güç analizinin yalnızca tamamlayıcı nitelikte kullanılması, çalışmanın metodolojik eksikliklerini örtbas etmek amacıyla kullanılmaması önerilmektedir.

3.3. Duyarlılık ve Güç Eğrileri

Bazı durumlarda, araştırmacılar önceden belirlenmiş bir örneklem büyüklüğü ile çalışmak zorunda kalabilirler. Bu gibi durumlarda, mevcut örneklem büyüklüğüyle hangi büyüklükteki farkların tespit edilebileceği veya hangi etki düzeylerinin istatistiksel olarak yakalanabilir olduğu duyarlılık analizi ile değerlendirilebilmektedir. Duyarlılık (sensitivity) analizi, sabit tutulan örneklem büyüklüğü ve anlamlılık düzeyine karşılık olarak, belirli büyüklükteki etkilerin tespit edilme olasılığını hesaplamaktadır.

Bu analiz türü aynı zamanda araştırma planlamasında farklı senaryoların değerlendirilmesine olanak tanımakta ve karar vericilere çeşitli örneklem-etki büyüklüğü-güç kombinasyonlarını görselleştirme fırsatı sunmaktadır. Bu amaçla kullanılan güç eğrileri, araştırmacının hedeflediği güç düzeyine ulaşabilmesi için hangi örneklem büyüklüğünün gerekli olduğunu grafiksel olarak göstermektedir (Walters, 2004). Böylece, sınırlı kaynaklar altında en uygun örneklem büyüklüğünün belirlenmesi kolaylaşmaktadır.

4. Farklı Araştırma Desenlerinde Güç Analizi

Güç analizinin uygulanması, araştırmanın istatistiksel tasarımı kadar, karşılaştırılmak istenen değişkenlerin türüne ve analizde kullanılacak yöntemle göre de farklılık göstermektedir (Kul, 2011; Özdamar, 2017; Bulut, 2023; Görgülü ve Ergin, 2023). Her istatistiksel testin kendine özgü varsayımları ve yapısal özellikleri bulunduğu için, uygun örneklem büyüklüğünün belirlenmesi sürecinde kullanılacak testin dikkate alınması gerekmektedir. Bu bölümde, sağlık bilimlerinde yaygın olarak kullanılan araştırma desenleri açıklanacaktır.

4.1. Bağımsız İki Grubun Karşılaştırılması

Bağımsız iki grubun ortalamaları arasındaki farkın karşılaştırılması temeline dayanan hipotez testi bağımsız örneklem t-testidir. Bu hipotez testinde normal dağılan iki toplumdan rastgele çekilen iki farklı örneklemin ortalamaları karşılaştırılır. Bağımsız grupların örneklem büyüklüklerinin eşit olması şart değildir (Özdamar, 2015).

4.2. Bağımlı İki Ölçümün Karşılaştırılması

Aynı hastalardan farklı zamanlarda alınan ölçümlerin (eş ölçümler, matched pairs) ortalamaları arasındaki farkın karşılaştırılması temeline dayanan hipotez testi eşleştirilmiş örneklem t-testidir (Alpar, 2016). Örneğin; bir tedavi yönteminin hastalık üzerindeki etkisi araştırılırken aynı hastalardan tedavi öncesi ve tedavi sonrasında alınan ölçümler eş ölçümlerdir. Bu testte aynı hastalardan farklı zamanlarda elde edilen bu ölçümler arasında korelasyon (ilişki) olduğundan güç analizinde bu korelasyon düzeyi de dikkate alınmalıdır (Bulut, 2023).

4.3. Bağımsız k ($k > 2$) Grubun Karşılaştırılması

Üç ya da daha fazla grubun ortalamalarının karşılaştırılmak istendiği çalışmalarda tek yönlü varyans analizi (ANOVA) kullanılmaktadır. Bu tür çalışmalarda, gruplar arası farkın hangi büyüklükte olacağının yanı sıra gruplar içi varyasyonun da örneklem büyüklüğü üzerinde etkisi bulunmaktadır (Sümbüloğlu ve Sümbüloğlu, 2012). Güç analizinde, gruplar arası ortalama farkların etkisi genellikle eta-kare (η^2) gibi bir etki büyüklüğü ölçütü ile ifade edilmektedir. Bu tür analizlerde, grupların sayısı arttıkça, aynı güç düzeyini sağlayabilmek için daha büyük bir örneklem ihtiyacı duyulmaktadır (Görgülü ve Ergin, 2023).

5. Güç Analizinde Kullanılan Yazılımlar ve Uygulamalar

Güç analizinin etkin bir şekilde gerçekleştirilebilmesi için çeşitli yazılım araçlarından yararlanılmaktadır. Bu yazılımlar sayesinde araştırma türüne uygun örneklem büyüklüğü hesaplamaları hızlı ve sistematik bir biçimde yapılabilmekte, farklı senaryolar test edilebilmektedir. Kullanılan yazılımlar arasında hem kullanıcı dostu arayüz sunan görsel platformlar hem de daha esnek hesaplamalar yapılabilen programlama tabanlı araçlar yer almaktadır. Bu yazılımlardan en sık kullanılanları PASS (Power Analysis and Sample Size), R programlama dili ve G*Power'dır.

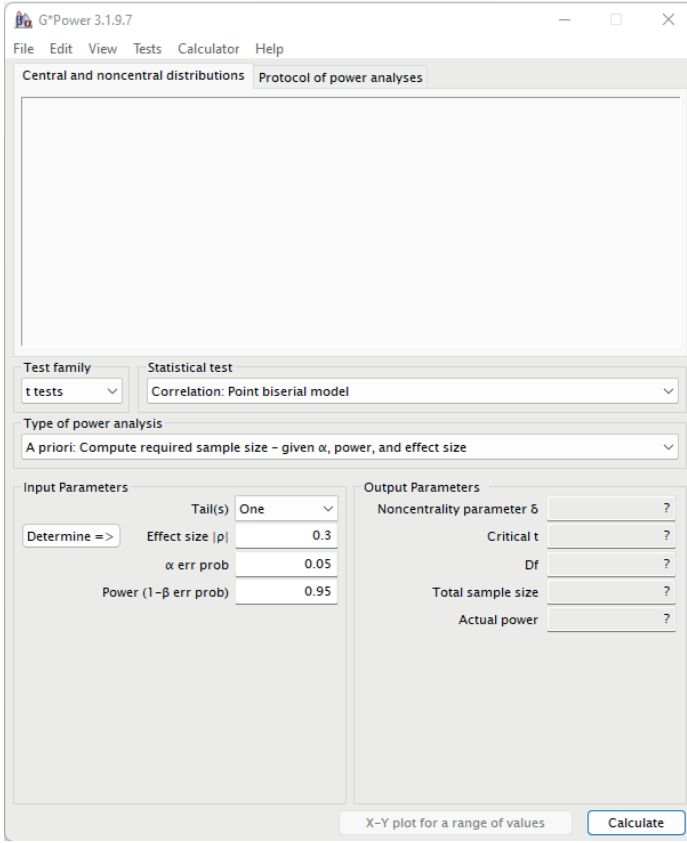
PASS, daha ileri düzey istatistiksel analizler ve özel tasarımlar için geniş seçenekler sunan bir ticari yazılımdır. Özellikle lojistik regresyon, Cox regresyon, tekrarlı ölçümler, non-inferiority/superiority tasarımları ve çoklu hipotez düzeltmeleri içeren çalışmalar için gelişmiş örneklem hesaplama modüllerine sahiptir. Ayrıca duyarlılık analizleri ve güç eğrileri gibi ileri senaryoların değerlendirilmesine de olanak tanımaktadır. Lisans gerektirmesi nedeniyle erişimi sınırlı olmakla birlikte, çok merkezli ya da faz çalışmaları gibi yüksek maliyetli araştırmalarda örneklem planlaması için sıklıkla kullanılmaktadır (Hintze, 2011).

Açık kaynak kodlu bir istatistiksel programlama dili olan R, güç analizi uygulamaları için çeşitli paketler sunmaktadır. Bunlar arasında en çok kullanılanları *pwr*, *WebPower* ve *simr* paketleridir. Bu paketler, standart senaryoların ötesinde özelleştirilmiş analizlerin yapılabilmesine olanak sunduğu için araştırmalarda esneklik sağlamaktadır. Programlama bilgisi gerektirmesi nedeniyle başlangıç düzeyindeki kullanıcılar için öğrenme süreci daha uzun olabilmektedir (Champely vd., 2018).

G*Power, Düsseldorf Üniversitesi'nden Erdfelder, Faul ve Buchner tarafından geliştirilmiş, kullanımı kolay ve ücretsiz bir güç analizi programıdır (Faul vd, 2007). Hem a priori hem de posterior güç analizleri yapma imkanı sunarak, araştırmanın planlama aşamasından sonuçların yorumlanmasına kadar geniş bir yelpazede destek sağlar. G*Power programını indirmek için en güvenilir ve güncel kaynak, genellikle programın geliştiricisi olan Düsseldorf Üniversitesi'ne ait resmi web sitesidir (www.psychologie.hhu.de). Program hem Windows hem de Mac OS X işletim sistemleri için farklı sürümler sunmaktadır. Windows kullanıcıları genellikle Windows XP, Vista, 7, 8, 10 ve hatta 11 gibi daha yeni sürümlerle uyumlu olan 3.1.x serisi sürümleri indirebilirken, Mac kullanıcıları da OS X 10.7'den başlayarak daha güncel sürümleri destekleyen versiyonlara ulaşabilirler. İndirme işlemi genellikle ilgili işletim sistemine uygun bağlantıya tıklanarak gerçekleştirilir ve indirilen dosyanın bir sıkıştırılmış klasör (.zip) olması durumunda, kurulumdan önce klasör içeriğinin açılması gerekebilir (Görgülü ve Ergin, 2023).

6. G*Power Programı ile Güç Analizi

Bu bölümde G*Power programı (Şekil 1) ile bağımsız iki grup, bağımlı iki ölçüm ve bağımsız k grup ($k > 2$) karşılaştırmasına ilişkin analizlere dayalı a priori güç analizlerinin nasıl yapılacağı en sade şekli ile anlatılacaktır.



Şekil 1. G*Power 3.1.9.7 programı açılış penceresi

G*Power ile güç analizi yapmak genellikle aşağıdaki adımları içerir:

Spesifik İstatistiksel Testin Seçimi: “Tests” menüsü içinden ya da test ailesi (test family) seçildikten sonra istatistiksel testler arasında araştırmanın amacına uygun spesifik test belirlenir (örneğin, “Means: Difference between two independent means (two groups)” veya “ANOVA: Fixed effects, omnibus, one-way”).

Güç Analizi Türünün Belirlenmesi: Genellikle araştırmanın planlama aşamasında a priori güç analizi tercih edilir. Bu analiz türünde beklenen etki büyüklüğü, α hatası ve çalışmada hedeflenen istatistiksel güç ($1 - \beta$) değerleri girilerek gerekli örneklem büyüklüğü hesaplanır. Posterior güç analizinde ise örneklem büyüklüğü, α hatası ve elde edilen etki büyüklüğü girilerek çalışma sonunda ulaşılan güç değeri hesaplanır.

Input (Girdi) Parametrelerinin Belirlenmesi: Bu adımda, seçilen güç analizi türüne göre gerekli parametreler programa girilir:

Etki Büyüklüğü (Effect Size): Etki büyüklüğü, literatürdeki referans araştırmalarda yer alan olan ya da pilot çalışmadan elde edilen ortalama, standart sapma, karşılaştırılan grupların birim sayısı, değişkenler arasındaki korelasyon düzeyi gibi değerler kullanılarak hesaplanır.

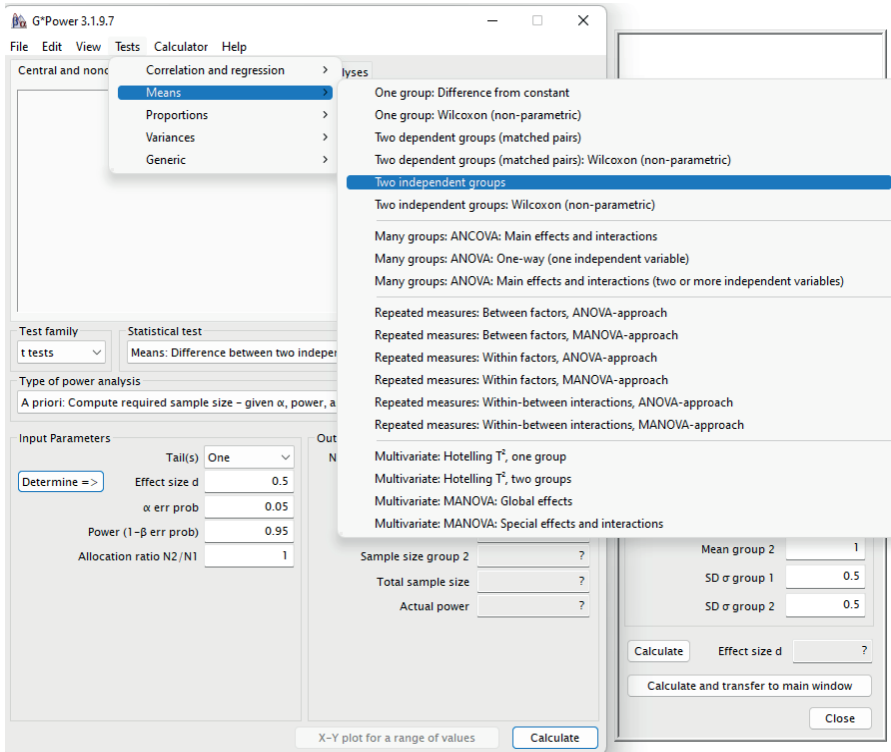
α Hatası: Yanlış pozitif sonuç elde etme olasılığıdır (genellikle 0,05 olarak ayarlanır).

Hedeflenen Güç (Power): Genellikle en az 0,80 olarak önerilir (Özdamar, 2017).

Hesaplama ve Yorumlama: Tüm parametreler girildikten sonra, G*Power programı gerekli hesaplamaları yapar ve istenen çıktıları (Total sample size) sunar. Bu çıktılar, araştırmanın tasarımını optimize etmek için kullanılır.

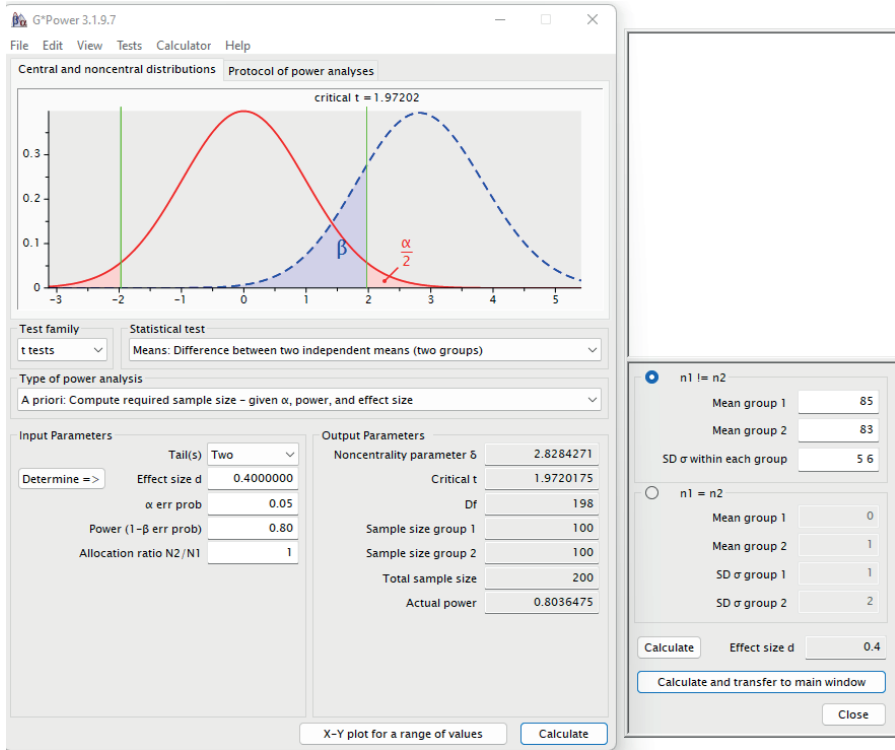
6.1. Bağımsız İki Grup Karşılaştırmasına Dayalı Örneklem Büyüklüğü Hesaplaması

Bağımsız iki grup karşılaştırmasına dayalı örneklem büyüklüğü hesaplanırken öncelikle “bağımsız örneklem t testi” anlamına gelen “two independent groups” hipotez testi “means” sekmesi içinden seçilir (Şekil 2).



Şekil 2. Bağımsız iki grup karşılaştırma testine ilişkin ekran görüntüsü

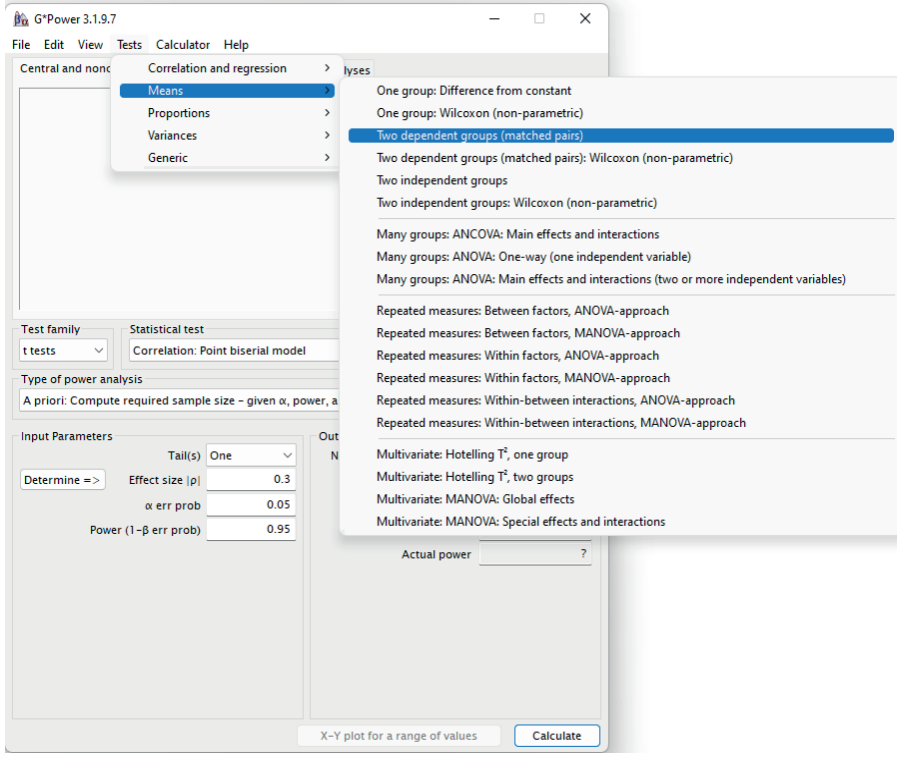
Açılan ekranda input parametresi olarak α hata düzeyi 0,05; hedeflenen güç 0,80 ve analiz sonucunda karşılaştırılacak grupların büyüklüğünün birbirine eşit olması isteniyorsa grupların büyüklüğünün birbirine oranı ($N2/N1$) değeri 1 olarak girilir. Burada etki büyüklüğü değeri hesaplanırken referans yayındaki grupların büyüklüklerinin birbirine eşit olup olmadığına dikkat edilmelidir. Eğer referans yayındaki grupların büyüklüğü birbirine eşit ise etki büyüklüğü değeri hesaplanırken ($n1=n2$) yaklaşımı ile grupların ortalama ve standart sapma değerleri girilerek “calculate and transfer to main window” yardımı ile etki büyüklüğü hesaplanarak input bölümüne taşınır. Referans yayının grupları farklı büyüklükte ise etki büyüklüğünün hesaplanması için ($n1 \neq n2$) yaklaşımı ile grupların ortalama değerleri ile standart sapma değerleri “SD σ within each group” kutusuna aralarında boşluk bırakarak yazılır. Şekil 3’teki örnek değerlere göre ortalamaları sırası ile 85, 83 ve standart sapma değerleri sırası ile 5 ve 6 olan iki grubun etki büyüklüğü grupların büyüklüklerinin farklı olduğu varsayımına göre 0,4 olarak hesaplanmıştır. Buna göre çalışmaya her iki gruptan en az $n=100$ ’er katılımcı olmak üzere toplam en az $n=200$ katılımcı alınması gerektiği hesaplanmıştır.



Şekil 3. Bağımsız iki grup karşılaştırma testine ilişkin örneklem büyüklüğü hesaplama ekranı

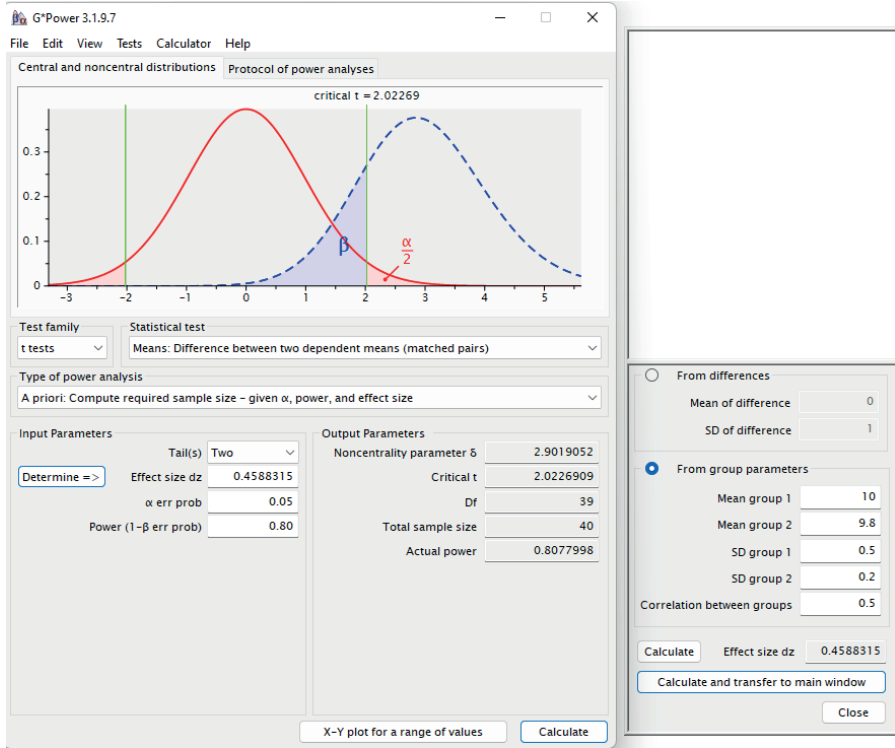
6.2. Bağımlı İki Ölçüm Karşılaştırmasına Dayalı Örneklem Büyüklüğü Hesaplaması

Bağımlı iki ölçüm karşılaştırmasına dayalı örneklem büyüklüğü hesaplanırken öncelikle “eşleştirilmiş örneklem t testi” anlamına gelen “two dependent groups (matched pairs)” hipotez testi “means” sekmesi içinden seçilir (Şekil 4).



Şekil 4. Bağımlı iki ölçüm karşılaştırma testine ilişkin ekran görüntüsü

Açılan ekranda input parametresi olarak α hata düzeyi 0,05; hedeflenen güç 0,80 olarak girilir. Burada etki büyüklüğü değeri hesaplanırken referans yayındaki bağımlı ölçümlerin ortalama ve standart sapma değerleri ile bu ölçümler arasındaki korelasyon katsayısının değeri “correlation between groups” kutusuna girilir. Referans çalışmada korelasyon katsayısı belirtilmemiş ise pilot bir çalışma yapılarak korelasyon düzeyi hesaplanmalıdır. Pilot çalışma yapmak maliyetli ya da mümkün değil ise korelasyon katsayısı orta düzeyde 0,5 seçilerek etki büyüklüğü hesaplanabilir (Görgülü ve Ergin, 2023). Şekil 5’teki örnek değerlere göre ilk ve son ölçümün ortalama değerleri sırası ile 10 ve 9,8; standart sapma değerleri ise 0,5 ve 0,2’dir. İki bağımlı ölçüm arasında orta düzeyde korelasyon olduğu varsayımına göre etki büyüklüğü 0,458 olarak elde edilmiştir. Yapılan analiz sonucunda çalışmaya dahil edilmesi gereken örneklem büyüklüğü en az $n=40$ olarak hesaplanmıştır.



Şekil 5. Bağımlı iki ölçüm karşılaştırma testine ilişkin örneklem büyüklüğü hesaplama ekranı

6.3. Bağımsız k ($k > 2$) Grup Karşılaştırmasına Dayalı Örneklem Büyüklüğü Hesaplaması

Bağımsız k ($k > 2$) grup karşılaştırmasına dayalı örneklem büyüklüğü hesaplanırken öncelikle “tek yönlü varyans analizi (ANOVA)” anlamına gelen “Many groups: ANOVA: one-way (one independent variable)” hipotez testi “means” sekmesi içinden seçilir (Şekil 6).

G*Power 3.1.9.7

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

Test family: F tests

Statistical test: ANOVA: Fixed effects, omnibus, one-way

Type of power analysis: A priori: Compute required sample size - given α , power, and effect size

Input Parameters

Determine =>

Effect size f: 0.25

α err prob: 0.05

Power (1- β err prob): 0.95

Number of groups: 5

Output Parameters

Noncentrality parameter λ : ?

Critical F: ?

Numerator df: ?

Denominator df: ?

Total sample size: ?

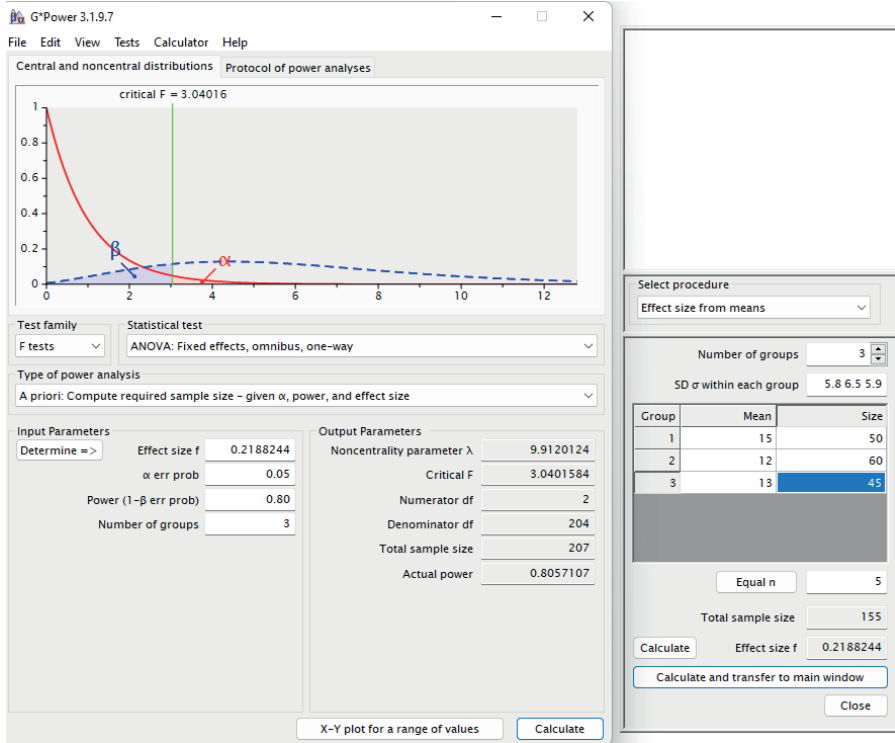
Actual power: ?

X-Y plot for a range of values

Calculate

Şekil 6. Bağımsız k ($k > 2$) grup karşılaştırma testine ilişkin örneklem büyüklüğü hesaplama ekranı

Açılan ekranda input parametresi olarak karşılaştırılacak grup sayısı değeri (number of groups) ile α hata düzeyi 0,05; hedeflenen güç 0,80 olarak girilir. Burada etki büyüklüğü değeri çoğunlukla referans yayındaki karşılaştırma gruplarının büyüklüğü, ortalama ve standart sapma değerleri yaklaşımı ile hesaplanır. Grupların ortalama ve büyüklükleri ilgili kutudaki boşluklara, standart sapma değerleri ise aralarında boşluk bırakarak yazılır. Şekil 7'deki örnek değerlere göre karşılaştırılan üç gruptaki katılımcı sayısı sırası ile $n=50$, $n=60$, $n=45$ olmak üzere; grupların ortalama değerleri sırası ile 15, 12, 13 ve standart sapma değerleri ise sırasıyla 5,8; 6,5 ve 5,9 olarak girilmiş ve etki büyüklüğü değeri 0,219 olarak elde edilmiştir. Yapılan analiz sonucu çalışmaya toplam en az $n=207$ katılımcının alınması gerektiği hesaplanmıştır.



Şekil 7. Bağımsız k ($k > 2$) grup karşılaştırma testine ilişkin örneklem büyüklüğü hesaplama ekranı

7. Güç Analizinin Klinik Araştırmalara Yansımaları

Sağlık bilimlerinde yürütülen araştırmaların yalnızca bilimsel bilgi üretmekle kalmayıp, aynı zamanda klinik uygulamalara yön verme amacı da taşıdığı bilinmektedir. Bu nedenle, araştırma sonuçlarının geçerliliği, güvenilirliği ve genellenebilirliği, sağlık hizmetlerinin kalitesi üzerinde doğrudan belirleyici bir etkiye sahiptir. Güç analizi, bu bağlamda yalnızca istatistiksel planlamanın bir unsuru olarak değil, aynı zamanda klinik kararların bilimsel temellere dayandırılması açısından da stratejik bir araç olarak değerlendirilmelidir.

Uygun biçimde yapılmış bir güç analizi sayesinde, bir çalışmanın belirli bir klinik etkisi olup olmadığını ortaya koyabilecek yeterlilikte olup olmadığı önceden belirlenebilmektedir. Bu durum, araştırmadan elde edilecek bulguların yanlış bir şekilde “etki yok” şeklinde yorumlanmasının önüne geçilmesini sağlamaktadır. Çünkü kimi zaman klinik olarak anlamlı bir etkinin var olmasına rağmen, yetersiz örneklem büyüklüğü nedeniyle

istatistiksel olarak anlamlı sonuçlara ulaşamamaktadır. Böyle bir durumda, potansiyel olarak faydalı bir tedavi yöntemi ya da müdahale yanlışlıkla etkisiz kabul edilebilmekte, bu da sağlık hizmetlerinin niteliğini ve hasta yararını olumsuz etkileyebilmektedir.

Öte yandan, olması gerekenden fazla örnekleme yürütülen çalışmalar da hem etik hem de pratik açılardan sakıncalar doğurabilmektedir. Gereksiz yere daha fazla hastanın çalışmaya dahil edilmesi, hasta güvenliği açısından ek riskler oluşturabileceği gibi, kaynakların israfına da yol açabilmektedir. Ayrıca, istatistiksel olarak anlamlı bulunan sonuçların gerçekte klinik açıdan önemsiz düzeyde etkiler taşıması da mümkündür. Bu gibi durumların önlenmesi için araştırmanın başında etkili bir güç analizinin gerçekleştirilmesi gerekmektedir.

Klinik araştırmaların değerlendirilmesinde yalnızca p-değerine odaklanılması, sonuçların yorumlanmasını dar bir çerçeveye indirgemekte; buna karşın güç analizi ile birlikte etki büyüklüğü ve güç değeri gibi ek ölçütlerin raporlanması, bulguların daha kapsamlı ve doğru şekilde değerlendirilmesini sağlamaktadır. Nitekim saygın tıbbi dergiler ve CONSORT, STROBE gibi araştırma raporlama yönergeleri tarafından da, güç analizinin raporlanması önerilmekte, bu analiz aracılığıyla çalışmanın istatistiksel yeterliliğinin belgelenmesi beklenmektedir. (Schulz vd., 2010; von Elm vd., 2007)

Bu çerçevede değerlendirildiğinde, güç analizinin yalnızca teknik bir hesaplama olmaktan öte, klinik araştırmalarda metodolojik sağlamlık ve etik uygunluk açısından temel bir bileşen olduğu görülmektedir. Güç analizine dayalı planlamaların yapılması, hem araştırmacıların daha doğru bilimsel sonuçlara ulaşmalarını sağlamak hem de bu sonuçların hasta bakımına entegrasyonunu kolaylaştırmaktadır.

8. Güç Analizinde Sık Yapılan Hatalar

Güç analizi, araştırma sürecinin erken dönemlerinde karar vericilere önemli avantajlar sağlayan bir araç olmakla birlikte, yanlış veya eksik uygulandığında hem metodolojik sorunlara hem de etik tartışmalara yol açabilmektedir. Bu nedenle, güç analizinin uygulanmasında sık karşılaşılan hataların bilinmesi ve bu hatalardan kaçınılması, araştırma kalitesinin artırılması açısından büyük önem taşımaktadır.

8.1. Etki Büyüklüğünün Gelişigüzel Belirlenmesi

En sık karşılaşılan hatalardan biri, güç analizinde kullanılacak etki büyüklüğünün herhangi bir bilimsel temele dayandırılmadan, örneğin

“orta düzey etki” varsayımıyla belirlenmesidir. Bu yaklaşım, örneklem büyüklüğünün gerçekte gerekenden çok az ya da çok fazla olmasına yol açabilmektedir. Etki büyüklüğü belirlenirken, benzer konuda daha önce yayımlanmış çalışmalardan yararlanılması, pilot veri kullanılması ya da klinik uzman görüşünün alınması önerilmektedir.

8.2. Posterior Güç Analizine Aşırı Güven Duyulması

Veri toplama süreci tamamlandıktan sonra yapılan posterior güç analizleri, sıklıkla çalışmanın yeterli olup olmadığını “kanıtlama” amacıyla kullanılmaktadır. Ancak bu tür analizler, elde edilen sonuçlara doğrudan bağımlı olduğu için yanıltıcı olabilir. Özellikle anlamlı olmayan sonuçlar elde edildiğinde yapılan posterior analizlerde düşük güç değeri çıkması beklenen bir durumdur ve bu, çalışmanın kalitesi hakkında doğrudan bilgi vermez. Bu nedenle posterior güç analizlerinin sınırlılıklarının farkında olunmalı ve bu analizler, yalnızca tamamlayıcı nitelikte kullanılmalıdır.

8.3. Anlamlılık Düzeyi ve Güç Değerinin Sabit Kabul Edilmesi

Güç analizlerinde anlamlılık düzeyi (genellikle %5) ve güç düzeyi (genellikle %80 veya %90) sıklıkla sabit varsayımlarla ele alınmakta, farklı senaryolar yeterince test edilmemektedir. Oysa bu değerlerin farklı kombinasyonları altında güç analizinin tekrarlanması, araştırma planlaması açısından daha bilinçli ve esnek bir yaklaşım sunar. Özellikle sınırlı kaynaklara sahip çalışmalar için bu tür senaryo testlerinin yapılması önem arz etmektedir.

8.4. Örneklem Büyüklüğü Hesaplamasının Veri Toplama Sonrası Yapılması

Bazı araştırmalarda, örneklem büyüklüğü hesaplamasının veriler toplandıktan sonra yapıldığı görülmektedir. Bu durum, güç analizinin amacına aykırıdır ve bilimsel açıdan geçerliliğini yitirmesine neden olur. Güç analizi, araştırma tasarımının bir parçası olarak veri toplanmadan önce yapılmalıdır. Aksi takdirde, hesaplamaların yalnızca göstermelik olduğu yönünde etik sorunlar gündeme gelebilir.

8.5. Uygun Olmayan Test Türüyle Güç Hesaplaması Yapılması

Güç analizinin, araştırmada gerçekten kullanılacak istatistiksel test türüne uygun olarak yapılması gerekir. Örneğin regresyon planlanan bir çalışmada sadece iki grup karşılaştırmasına göre güç analizi yapılması, örneklem büyüklüğünü olduğundan az gösterebilir. Benzer şekilde, parametrik varsayımların karşılanmadığı durumlarda kullanılacak test türünün hesaplamada dikkate alınmaması da yaygın bir hatadır.

9. Sonuç ve Öneriler

Sağlık bilimlerinde yürütülen araştırmaların metodolojik sağlamlığının artırılması ve elde edilen bilimsel bulguların güvenilirliğinin sağlanması, büyük ölçüde araştırma sürecinin erken aşamalarında yapılan istatistiksel planlamaların niteliğine bağlıdır. Bu bağlamda, güç analizi yalnızca örneklem büyüklüğünün tahmini için kullanılan teknik bir araç olarak değil; aynı zamanda etik, bilimsel ve uygulamalı sorumlulukları da içeren bütüncül bir değerlendirme aracı olarak ele alınmalıdır.

Bu bölümde, güç analizinin temel kavramları açıklanmış, farklı araştırma desenlerinde nasıl uygulanması gerektiği örneklerle gösterilmiş ve klinik araştırmalara olan etkisi detaylı biçimde ele alınmıştır. Ayrıca, yaygın yazılımların kullanımına ilişkin pratik bilgiler sunulmuş, analiz sürecinde karşılaşılabilecek başlıca hatalara dikkat çekilmiştir. Elde edilen tüm bu bilgiler ışığında, aşağıda bazı temel öneriler sıralanmıştır:

Araştırma planlaması yapılırken güç analizine mutlaka yer verilmelidir. Özellikle yeni bir tedavi ya da girişimin etkisinin değerlendirilmesinin amaçlandığı çalışmalarda, uygun örneklem büyüklüğünün belirlenmesi bilimsel ve etik açıdan zorunlu kabul edilmelidir.

Etki büyüklüğü, literatüre dayalı biçimde ya da pilot veriler aracılığıyla rasyonel şekilde belirlenmelidir. Gelişigüzel tahmin edilen etki büyüklükleri, örneklem yetersizliğine veya kaynak israfına neden olabilir.

Yalnızca a priori değil, gerektiğinde duyarlılık ve senaryo analizleri de planlamaya dahil edilmelidir. Bu sayede, sınırlı kaynak koşullarında optimal çözüm yolları oluşturulabilir.

Güç analizi için kullanılan yazılımlar dikkatli seçilmeli, çalışmanın istatistiksel yapısıyla uyumlu araçlar tercih edilmelidir. Özellikle çok değişkenli modellerde ya da parametrik olmayan testlerin kullanıldığı durumlarda gelişmiş yazılımlara ihtiyaç duyulabilir.

Güç analizi süreci şeffaf biçimde raporlanmalı ve kullanılan varsayımlar açıkça belirtilmelidir. Böylece araştırmanın tekrarlanabilirliği artar ve bilimsel güvenilirlik sağlanmış olur.

Sonuç olarak, güç analizine ilişkin farkındalığın artırılması ve disiplinlerarası araştırma ekiplerinde bu kavramın ortak bir dil olarak benimsenmesi, sağlık bilimlerinde yapılan çalışmaların kalitesine doğrudan katkı sağlayacaktır. Araştırmacıların yalnızca analiz aşamasında değil, araştırmanın her safhasında istatistiksel planlamaya gerekli önemi göstermeleri; elde edilen bilimsel bilginin klinik uygulamalara güvenle aktarılabilmesi açısından

büyük önem taşımaktadır. Bu bağlamda, güç analizinin yalnızca istatistiksel değil; aynı zamanda bilimsel, etik ve klinik bir sorumluluk alanı olduğu unutulmamalıdır.

Referanslar

- Dell, R. B., Holleran, S., & Ramakrishnan, R. (2002). Sample size determination. *ILAR Journal*, 43(4), 207–213.
- Biau, D. J., Jolles, B. M., & Porcher, R. (2010). P value and the theory of hypothesis testing: an explanation for new researchers. *Clinical Orthopaedics and Related Research*, 468(3), 885–892.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Bulut, Y. (2023). İstatistiksel güç ve örneklem büyüklüğü: GPower uygulamalı.* Akademisyen Kitabevi.
- Fisher, R. A. (1970). Statistical methods for research workers. In *Breakthroughs in statistics: Methodology and distribution* (pp. 66–70). New York, NY: Springer New York.
- Kul S. (2011) Klinik Araştırmalarda Örnek Genişliği Belirleme. *Plevra Bülteni* 2011;2:129-32
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Kang, M., Ragan, B. G., & Park, J. H. (2008). Issues in outcomes research: an overview of randomization techniques for clinical trials. *Journal of athletic training*, 43(2), 215–221.
- Lenth, R. V. (2001). Some practical guidelines for effective sample size determination. *The American Statistician*, 55(3), 187–193.
- Özdamar, K. (2017). *Modern bilimsel araştırma yöntemleri* (3. baskı). Nisan Kitabevi.
- Görgülü, Ö., & Ergin, R. (2023). G*Power programıyla örnek büyüklüğü hesaplama ve güç analizi (Power analysis).* Nobel Akademik Yayıncılık.
- Walters, S. J. (2004). Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the SF-36. *Health and quality of life outcomes*, 2, 1–17.
- Özdamar, K. (2015). *Paket programlar ile istatistiksel veri analizi – 1* (10. baskı). Nisan Kitabevi.
- Alpar, R. (2016). *Spor, sağlık ve eğitim bilimlerinden örneklerle uygulamalı istatistik ve geçerlik-güvenirlik* (4. baskı). Detay Yayıncılık.
- Sümbüloğlu, K., & Sümbüloğlu, V. (2012). *Biyoistatistik*. Hatiboğlu Yayınları.
- Hintze, J. (2011). *PASS (power analysis and sample size system) quick start manual*.
- Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., ... & De Rosario, M. H. (2018). Package ‘pwr’. R package version, 1(2).

- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2), 175-191.
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, c332.
- Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Götzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *The lancet*, 370(9596), 1453-1457.