

TEKRARLI ÖLÇÜMLER ve TASARIMLAR

Son üç hafta boyunca çeşitli ortalamalar arasındaki farklılıkları test etmek için kullanılan bir prosedürü inceledik. Şimdiye kadar farklı deneklerin farklı ortalamalara katkıda bulunduğu durumlara odaklandık; başka bir deyişle, farklı denekler farklı deneysel koşullarda yer alıyor.

Burada denekler, farklı bitkiler, farklı şirketler, farklı araziler, farklı viral türler, farklı hayvanlar olabilir. Buradaya kadar herşey normaldi. Şimdi, aynı deneklerin farklı araçlara katkıda bulunduğu durumları göz önüne alalım.

- Tekrarlı ölçüm tasarımı, deneklerin **bağımlı değişken** üzerinde **iki veya daha fazla kez ölçüldüğü** bir araştırma tasarımıdır.
- Tekrarlanan ölçüm verileri, zaman içinde tekrar tekrar gözlem yaptığınız deneylerden elde edilir. Tekrarlanan ölçümler deneyinde, deneysel birimler zaman içinde birden fazla noktada gözlemlenir. Dolayısıyla, zamanın bir noktasındaki bir gözleme bakmak yerine, zamanın birden fazla noktasındaki verilere bakacağız. Bu tür verilerle, yalnızca tek bir yanıt değişkenine bakarız, **ancak zaman içinde ölçülür**.

Bağımlı örneklem t-testi

- Bağımlı t-testi (eşleştirilmiş t-testi veya eşleştirilmiş örneklem t-testi olarak da adlandırılır), bu ortalamalar arasında istatistiksel olarak anlamlı bir fark olup olmadığını belirlemek için ilgili iki grubun ortalamalarını karşılaştırır.
- Bağımlı t-testi, "denek içi (within-subjects)" veya "tekrarlı ölçümler (repeated-measures)" istatistiksel testine bir örnektir. Bu, aynı katılımcıların birden fazla kez test edildiğini gösterir. Dolayısıyla, bağımlı t-testinde "ilgili gruplar (related groups)" her iki grupta da aynı katılımcıların bulunduğunu gösterir. Her grupta aynı katılımcıların bulunmasının mümkün olmasının nedeni, her katılımcının aynı bağımlı değişken üzerinde iki kez ölçülmüş olmasıdır.

- Örneğin, 10 katılımcının bir heceleme testindeki performansını (bağımlı değişken), hecelemeyi geliştirmek için yeni bir bilgisayarlı öğretim yöntemine tabi tutulmadan önce ve sonra ölçmüş olabilirsiniz.
- Bilgisayar eğitiminin heceleme performanslarını geliştirip geliştirmediğini bilmek istiyorsunuz. Burada bağımlı bir t-testi kullanabiliriz çünkü iki ilgili grubumuz var. İlk ilgili grup, bilgisayarlı heceleme eğitiminin başlangıcındaki (öncesindeki) katılımcılardan, ikinci ilgili grup ise aynı katılımcılardan, ancak şimdi bilgisayarlı eğitimin sonundaki katılımcılardan oluşmaktadır.

Varsayımları

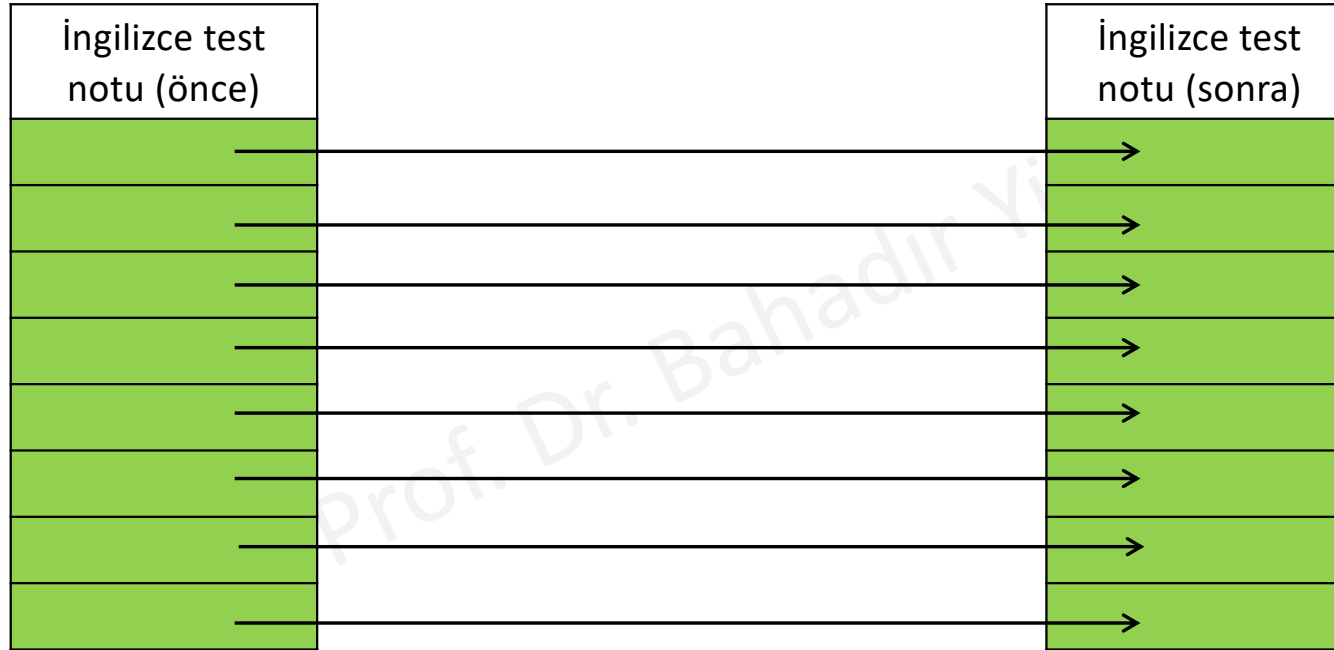
- Varsayım #1: Bağımlı değişkeniniz sürekli bir ölçekte ölçülmelidir (yani aralık veya oran düzeyinde ölçülmelidir). Bu kriteri karşılayan değişkenlere örnek olarak revizyon süresi (saat cinsinden ölçülür), zeka (IQ puanı kullanılarak ölçülür), sınav performansı (0'dan 100'e kadar ölçülür), ağırlık (kg cinsinden ölçülür) ve benzerleri verilebilir.
- Varsayım #2: Bağımsız değişkeniniz iki kategorik, "ilişkili gruplar" veya "eşleştirilmiş çiftlerden" oluşmalıdır. "ilişkili gruplar" her iki grupta da aynı deneklerin bulunduğunu gösterir.

- Varsayım #3: İlgili iki grup arasındaki farklarda önemli aykırı değerler olmamalıdır. Aykırı değerler, verileriniz içinde olağan örüntüye uymayan tekil veri noktalarıdır (örneğin, 100 öğrencinin IQ puanlarının incelendiği ve ortalama puanın 108 olduğu ve öğrenciler arasında sadece küçük bir farklılığın olduğu bir çalışmada, bir öğrencinin puanı 156'dır, bu çok sıra dışıdır ve hatta onu dünya genelindeki IQ puanlarının en üst %1'ine yerleştirebilir)
- Aykırı değerlerle ilgili sorun, bağımlı t-testi üzerinde olumsuz bir etkiye sahip olmaları ve sonuçlarınızın geçerliliğini azaltmalarındadır. Ayrıca, testin istatistiksel anlamlılığını da etkileyebilirler

- Varsayım #4: Bağımlı deęişkendeki farklılıkların ilgili iki grup arasındaki dağılımı yaklaşık olarak normal dağılım göstermelidir. Bağımlı t-testinin sadece yaklaşık olarak normal veri gerektirdiğinden bahsediyoruz çünkü normallik ihlallerine karşı oldukça "dayanıkladır", yani varsayım biraz ihlal edilebilir ve yine de geçerli sonuçlar sağlayabilir. SPSS Statistics kullanarak kolayca test edilebilen Shapiro-Wilk normallik testini kullanarak normalliği test edebilirsiniz.

- SPSS kullanarak 3. ve 4. varsayımları kontrol edebilirsiniz. Bunu yapmadan önce, verilerinizin 1. ve 2. varsayımları karşıladığından emin olmalısınız, ancak bunu yapmak için SPSS'e ihtiyacınız yoktur. Varsayım #3 ve #4'e geçerken, bunları bu sırayla test etmenizi öneririz çünkü bu, varsayıma yönelik bir ihlalin düzeltilebilir olmaması durumunda artık bağımlı t-testini kullanamayacağınız bir sırayı temsil eder (bunun yerine verileriniz üzerinde başka bir istatistiksel test uygulayabilirsiniz). Bu varsayımlar üzerinde istatistiksel testleri doğru bir şekilde çalıştırmazsanız, bağımlı t-testi çalıştırırken elde ettiğiniz sonuçların geçerli olmayabileceğini unutmayın.

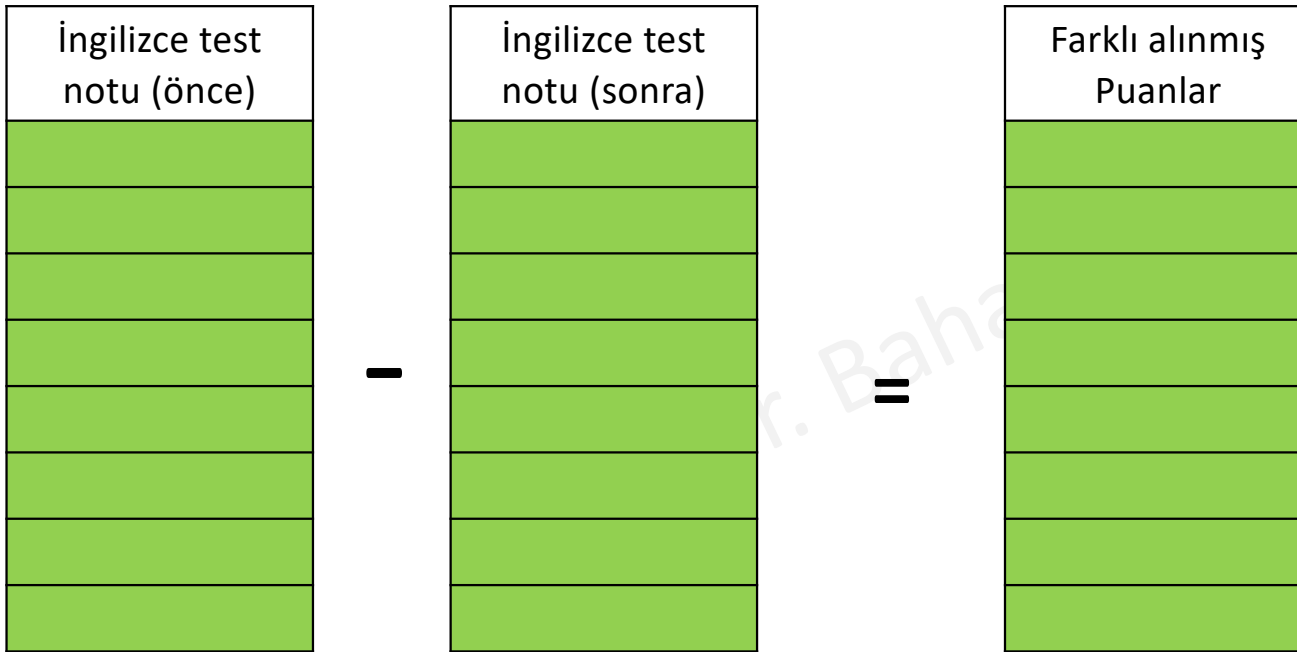
Diyelim ki bir eğitim ve araştırma kurumu yeni bir İngilizce öğrenme tekniğinin kalitesini araştırmak istiyor. Bu amaçla ilgili anakitleden bir örneklem alınıyor. Bu örnekleme eğitim öncesi teste tabi tutulup, test puanları kayıt ediliyor. Daha sonra araştırmanın amacına uygun, veya literatürden alınan bilgilere dayanılarak, belli bir süre yeni öğrenme tekniği bu örnekleme anlatılıyor. Eğitim sonunda aynı test tekrar uygulanıyor ve sonuçlar kayıt ediliyor. Acaba verilen eğitim istatistiksel olarak ne kadar etkili olmuştur.



H_0 : Verilen eğitim istatistiksel olarak faydalı olmamıştır

H_1 : Verilen eğitim istatistiksel olarak faydalı olmuştur.

Raporlama tek örneklem t testi ile benzerdir. Çünkü! İlk test ile son test puanları farkı alınarak, tek örneklem t testi yapılır. Sadece hipotez değişir.



$$H_0: \text{Farklı alınmış puanlar} = 0$$
$$H_1: \text{Farklı alınmış puanlar} \neq 0$$

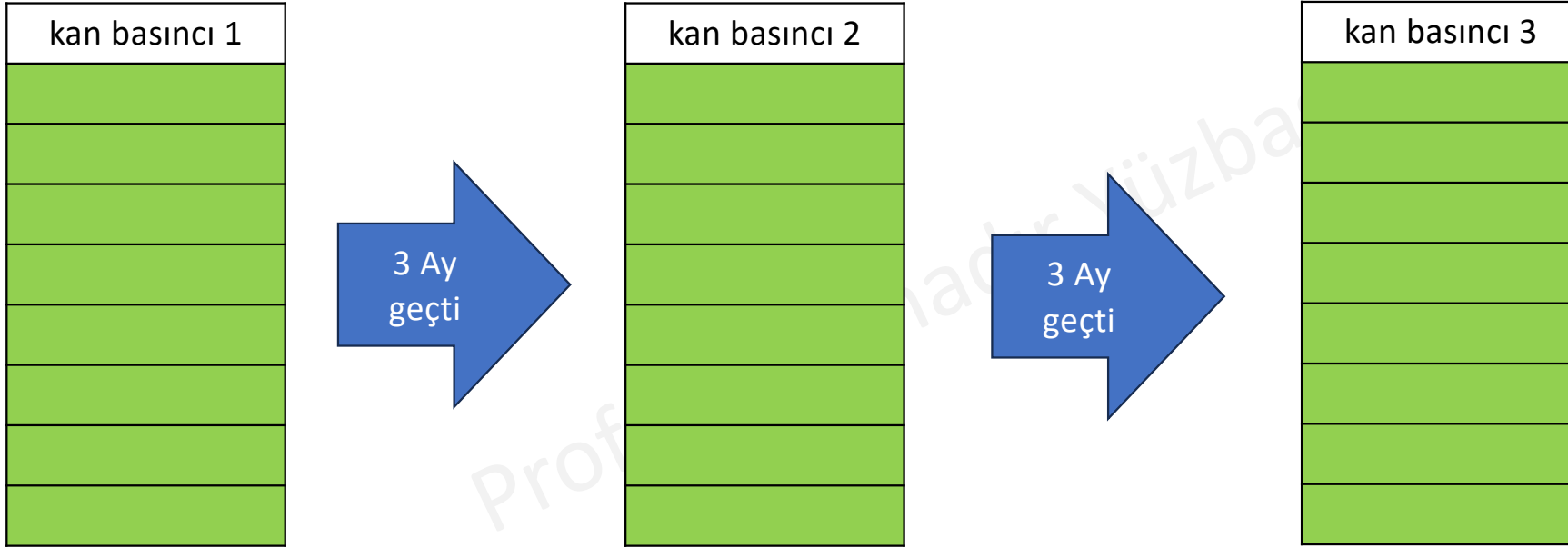
Tekrarlanan Ölçümler ANOVA

- Tekrarlanan ölçümler ANOVA, tek yönlü ANOVA'nın eşdeğeri, ancak bağımsız değil ilişkili gruplar içindir ve bağımlı t-testinin uzantısıdır. Tekrarlanan ölçümler ANOVA'sı aynı zamanda denek içi ANOVA veya ilişkili örnekler için ANOVA olarak da adlandırılır. Tüm bu isimler, tekrarlanan ölçümler ANOVA'sının doğasını, yani ilgili ortalamalar arasındaki genel farklılıkları tespit etmeye yönelik bir test olduğunu ima etmektedir.

Varsayımlar

- Normallik: Bağımsız örneklem için yapılan ANOVA testine benzer şekilde, bağımsız değişkenin her bir seviyesinin yaklaşık olarak **normal dağılım** göstermesi gerekir.
- **Küresellik (*sphericity*) kavramı**, tüm niyet ve amaçlar için, varyansların homojenliğinin tekrarlanan ölçümler eşdeğeridir. Mauchly'nin Küresellik Testi kullanılarak bu varsayım ölçülür
- Mauchly'nin Küresellik Testi, farkların varyanslarının eşit olduğu boş hipotezini test eder. Dolayısıyla, Mauchly'nin Küresellik Testi istatistiksel olarak anlamlıysa ($p < .05$), sıfır hipotezini reddedebilir ve farkların varyanslarının eşit olmadığı (yani, küresellik ihlal edilmiştir) alternatif hipotezini kabul edebiliriz..

- İki tür çalışma tasarımı için tekrarlanan ölçümler ANOVA'sı kullanarak verileri analiz edebiliriz.
- (1) üç veya daha fazla zaman noktasında ortalama skorlardaki değişiklikleri veya
- (2) üç veya daha fazla farklı koşul altında ortalama skorlardaki farklılıkları araştıran çalışmalar.
- Örneğin, (1) için, 6 aylık bir egzersiz eğitim programının kan basıncı üzerindeki etkisini araştırıyor olabilirsiniz ve kan basıncını 3 ayrı zaman noktasında (egzersiz müdahalesi öncesi, ortası ve sonrası) ölçmek isteyebilirsiniz, bu da herhangi bir egzersiz etkisi için bir zaman seyri geliştirmenize olanak tanır.
- Örneğin (2) için, farklı kişilere her bir farklı keki tattırmak yerine, aynı deneklere farklı kek türlerini (çikolata, karamel ve limon) yedirebilir ve her birini tat açısından değerlendirebilirsiniz.
- Bu iki çalışma tasarımında önemli olan nokta, aynı kişilerin aynı bağımlı değişken üzerinde birden fazla kez ölçülüyor olmasıdır (yani, buna tekrarlanan ölçümler denmesinin nedeni budur).



Örnek rapor

- Zamanın egzersiz kaynaklı zindelik üzerinde istatistiksel olarak anlamlı bir etkisi vardı, $F(df1, df2) = f\text{-değeri}, p = .000$.
- Altı aylık egzersiz-egitim programının kondisyon düzeyleri üzerinde istatistiksel olarak anlamlı bir etkisi olmamıştır, $F(df1, df2) = f\text{-değeri}, p = .123$.