

EĞİTİMDE PROGRAM DEĞERLENDİRME VE İSTATİSTİKSEL YÖNTEMLER

Yrd. Doç. Dr. Mustafa Akdağ •

Özet

Bu çalışmanın amacı, eğitimde program geliştirme çalışmalarında önemli ve sürekli bir yer tutan program değerlendirme işinin gerekliliği ve bu konudaki uygulanmakta olan değerlendirme yaklaşımları ve araştırma yöntemleri üzerinde durulmuştur. Program değerlendirme çalışmalarında yaygın bir şekilde kullanılan ki-kare, t testi, varyans analizi tekniklerinin hangi durumlarda nasıl kullanılması gerektiği; SPSS paket programı dikkate alınarak incelenmiştir.

Anahtar sözcükler: Program değerlendirme, araştırma yöntemi, istatistiksel analiz

Günümüz eğitim kurumlarının en temel işlevlerinden birisi şüphesiz Milli eğitim teori ve politikaları doğrultusunda bireyler yetiştirmektir. Her ülkede olduğu gibi bizde de yetiştireceğimiz insan tipinin temel karakteristikleri, eğitim ile ilgili olarak çıkarılan yasa ve yönetmeliklerde belirtilmiştir.

Eğitime yapılan yatırımların ülke kalkınmasına dönüşmesi için devlet; öğretmen yetiştirme, okul binalarının yapımı, araç-gereç temini vs. gibi çeşitli önlemler almaktadır. Bunların başında da eğitim programları gelmektedir. Programlar, eğitilenlerin davranış standartlarından, öğrenme ve öğretim etkinliklerine kadar tüm çalışmalara kılavuzluk etmektedir. Böylece eğitim programları, milli eğitim teori ve politikasının uygulamaya dönüştürülmesinde çok önemli bir köprü görevi görmektedir (Hakan, 1991: 3).

Gerek örgün, gerekse yaygın eğitim kurumlarındaki eğitimin niteliğinin artırılması, büyük ölçüde uygulanan programlara bağlıdır. Birey, toplum ve bilim alanındaki gereksinim ve değişimler dikkate alınarak, programlarda meydana gelen aksaklık ve eksikliklerinin giderilmesi; kısaca, sürekli yapılan program geliştirme çalışmaları, eğitimin niteliğinin olumlu yönde gelişmesine hizmet etmektedir (Erden, 1998: 3).

Program geliştirme çalışmaları masa başında gelişmiş güzel yapılan çalışmalar değil, eğitim süreci ile ilgili bütün koşulların, bireylerin, araç ve gereçlerin geliştirilmesine yönelik uygulamalı bir süreçtir. Ayrıca, bu süreçte Merkez (Milli Eğitim Bakanlığı) ile okullar arasında çift yönlü bir iletişim ve işbirliğinin sağlanması gerekir (Varış, 1996: 31). Bu işbirliğinin sağlanması amacıyla il ve ilçe Milli Eğitim Müdürlüklerinin bünyesinde “Program Hazırlama ve Geliştirme Komisyonları” kurulmuştur. Bu komisyonlar;

Yürürlükte bulunan programların alandaki uygulamalarını izlemek; yönetici, öğretmen, öğrenci velileri ve öğrencilerin görüşlerini almak; toplanan bilgileri ve dökümanları birleştirerek gerekli değerlendirmeleri yapıp, karşılaşılan problemler için düşünülen çözüm tekliflerini de içeren raporları hazırlamak ve ilgililere sunmak,... (MEB. Tebliğler Dergisi, 1995).

• İnönü Üniversitesi Eğitim Fakültesi. Bilgisayar ve Öğretim Teknolojileri Öğretmenliği Bölümü
makdag@inonu.edu.tr

görevlerini yerine getirirler. Söz konusu komisyonların görevleri ve çalışma esasları incelendiğinde program geliştirme çalışmalarında, değerlendirmenin önemli ve sürekli yer tuttuğu; her ikisinin de birbirini tamamlayan süreçler olduğu anlaşılmaktadır.

Program Değerlendirmenin Gerekliliği

Eğitim programının başarılı olabilmesi için bütün öğrencilerin programda öngörülen hedeflere ulaşmış olması gerekir. Ancak, bu durum her zaman gerçekleşmeyebilir. Bu nedenle, program uygulandıktan sonra, aksayan öğelerin olup olmadığı, varsa bunların nelerden kaynaklandığını belirlemek ve gerekli düzeltmeleri yapmak amacıyla programın değerlendirilmesine gereksinim duyulur. Program değerlendirme süreci: Veri toplama, verileri ölçütlerle karşılaştırıp yorumlama ve etkililik hakkında karar verme basamaklarından oluşmaktadır (Demirel, 1998: 159-160).

Program geliştiriciler, program hakkında değer yargılarına sahip olmak ve bu yargılar doğrultusunda programı kabul etmek, red etmek veya düzeltmeye tabi tutmak durumundadır. Program hakkında bu şekilde verilecek olan kararların isabetli olabilmesi için yapılacak olan ölçme ve değerlendirme işlemlerinin geçerli ve güvenilir olması gerekir. Aksi halde, değer yargıları da yanlış olabilir (Turgut, 1983 a: 216).

Hazırlanmış olan her program ne kadar eksiksiz ve kusursuz hazırlanmış olursa olsun henüz uygulamaya konulmadığı için denencel niteliktedir. Program hakkında bu aşamada verilecek olan yargılar bizi yanıltabilir. Kesin yargıya, program gerçek ortamlarda denenip uygulandıktan ve sonuçları görüldükten sonra varılabilir.

Program Değerlendirme Yaklaşımları

Eğitim programlarının değerlendirilmesi konusunda, program geliştirme ve değerlendirme uzmanları tarafından çeşitli yaklaşım ve modeller ortaya konulmakla birlikte, bunlar incelendiğinde ortak yönlerin çok olduğu görülür.

Turgut (1985: 106) öğretimin değerlendirilmesinde; öğretim programına, öğretim stratejisine ve öğretimi tamamlamış öğrencilerin öğrenme derecelerine bakılarak programın etkililiği hakkında karar verilebileceğini belirtmektedir. Bu üç tür değerlendirme yaklaşımından her biri program geliştirme çalışmalarına katkı sağlar ve birbirini tamamlar. Bunlardan sadece birini ele alıp programın etkililiği hakkında değer yargısına varmak, yanıltıcı olabilir. Çünkü, bu yaklaşımlar aynı bilgileri vermedikleri gibi eşdeğer kullanılışığa da sahip değildir.

Ertürk (1986: 114-115) ise değerlendirme yaklaşımlarını altı grupta toplamıştır: Yetişek tasarısına bakarak, ortama bakarak, başarıya bakarak, erişişe bakarak, öğrenmeye bakarak, ürüne bakarak yapılan değerlendirme.

Program değerlendirmedeki bu farklı yaklaşımların yanında, son zamanlarda ortaya çıkan farklı program değerlendirme modelleri de görülmektedir. Bu modellerden bazıları şunlardır: Hedefe Dayalı Değerlendirme Modeli, Metfessel-Michael Değerlendirme Modeli, Provus'un Farklar Yaklaşımı ile Değerlendirme Modeli, Stake'in Uygunluk Olasılık Modeli, Stufflebeam'in Bağlam, Girdi, Süreç ve Ürün Modeli (Demirel, 1998: 162-167).

Program Değerlendirmede Kullanılan Araştırma Yöntemleri

Araştırma yöntemleri, yapılan çalışmanın amacına, kullanılan veri toplama ve çözümleme tekniklerine, değişkenlerin kontrol edilebilme derecesine, verilerin kaynaklarına göre gruplandırılabilir. İlgili literatür incelendiğinde çok çeşitli gruplandırmaların yapıldığı görülür. Ancak, bunların benzerlikleri, farklılıklarından daha çoktur.

Program değerlendirme çalışmalarında kullanılan araştırma yöntemleri dört grupta incelenebilir. Bunlar: Deneysel yöntem, survey (betimsel) yöntemi, gözlem yöntemi ve tarihi yöntemdir (Erden, 1998: 42-56; Kaptan, 1991: 45-88; Karasar, 1999: 76-104).

Deneysel Yöntem

Bu yöntemin de kendi içinde çeşitleri bulunmaktadır:

1. Denk kontrol gruplu öntest-sontest deseni.- Bu desende deney ve kontrol grubundaki bireyler şans yöntemiyle belirlenir. Araştırmanın başında her iki gruba öntest verilir. Daha sonra deney grubuna değerlendirilecek program uygulanırken, kontrol grubuna ise herhangi bir program uygulanır.

Bu uygulamada öntest sonuçları, grupların ön bilgiler açısından denkliliğinin sağlanmasında kullanılır. Grupların gerek öntest, gerekse sontest puan ortalamaları arasındaki farkın anlamlı olup olmadığı, bağımsız gruplarda "t" testi ile test edilir. Grupların öntest puanı ortalamaları arasındaki fark anlamlı değilse; grupların sontest puanı ortalamaları arasındaki farkın anlamlı olup olmadığı test edilir. Fark anlamlı olup, deney grubunun lehine ise, deney grubuna uygulanan programla ilgili denenen işlemin, kontrol grubunda yapılan işlemde daha etkili olduğu şeklinde değerlendirme yapılabilir.

Bu desende her iki grubun kendi içlerinde öntest-sontest puanları arasındaki farkın anlamlılığına da bakılarak grupların erişimleri ortaya konulabilir. Bu durumda, bağımlı gruplarda "t" testi kullanılır. Böylece, grupların erişimleri dikkate alınmak suretiyle, elde edilen sonuçlar daha anlamlı bir şekilde yorumlanabilir. Bu desenin en önemli sınırlılığı, grupların şans yöntemiyle okul ortamında her zaman oluşturulamamasıdır.

2. Denk olmayan kontrol gruplu öntest-sontest deseni.- Bu desen, bir önceki desene benzer. Gruplar şans yöntemiyle seçilmez. Ancak, bunun yerine mümkün olduğunca deney grubuna benzeyen kontrol grubu kullanılır. Burada okulun düzeni bozulmadığı için, I. desene nazaran uygulanması daha kolaydır. Desenin uygulanması, istatistiksel analizler, bunların yorumlanması I. desendeki gibidir.

3. Denk kontrol gruplu sontest deseni.- Bu desen de I. desene benzemektedir. Tek farkı öntestin uygulanmamasıdır. Gruplara yalnızca sontest uygulanır. Bu desen, denenen işlem ve öntest ile sontest arasında süre kısa ise; öntestin uygulanması için yeterli zaman olmadığında tercih edilebilir. Bu durumda grupların şans yöntemiyle oluşturulması, deney öncesi benzerliği sağlamak için yeterli sayılabilir. Bu nedenle önteste gerek duyulmaz. Sontest puan ortalamaları arasındaki fark, bağımsız gruplarda “t” testi ile kontrol edilir.

4. Tek gruplu öntest-sontest deseni.- Değerlendirilecek olan program, başka herhangi bir program ile karşılaştırılmıyacaksa bu desen kullanılır. Burada tek bir grup alınır, bu gruba program uygulanmadan önce öntest ve program uygulandıktan sonra da sontest uygulanır. Öntest ve sontest puan ortalamaları arasındaki farkın anlamlı olup olmadığına bağımlı gruplarda “t” testi ile bakılır. Bu fark, istatistiksel olarak anlamlı bulunup, sontestin lehine ise programın başarılı olduğu yorumu yapılabilir. Ancak, bu desenin en önemli sınırlılığı, programın neye göre iyi, neye göre kötü olduğunu söylemenin mümkün olmamasıdır.

5. Tek grup zaman serili desen.- Programın uygulandığı grup aynı zamanda kontrol grubu olarak kabul edilir. Aynı ölçme aracı, uygulamadan önce ve sonra birkaç kez verilir. Bu desende uygulamadan önceki ve sonraki gelişmeler de gözönüne alınabilir. Bir gruba öntest ve sontest uygulanıp aradaki farkın değerlendirilmesi yapılırken, bu ölçmelerden önceki ve sonraki ölçümlerin sonucuna da bakmak gerekir. Bu desende elde edilen verilerin analizi ve yorumları güçtür. Test ortalamaları arasındaki farkın anlamlılığı bağımlı gruplarda “t” testi ile kontrol edilebilir.

6. Denk olmayan kontrol gruplu zaman serili desen.- Bu desen, bir önceki desenin kontrol gruplusudur. Kontrol grubu, değerlendirmeyi daha da güçlendirir.

Survey (Betimsel) Yöntemi

Programların aksaklık ve eksikliklerinin belirlenmesinde programla ilgili olan öğretmenlerin, öğrencilerin, yöneticilerin, velilerin, konu alanı ve program geliştirme uzmanlarının görüşlerinin alınmasına ihtiyaç duyulur. Böylece mevcut durum yansıtılmaya çalışılır. Survey araştırmalarında veri toplamak için genelde anket ve görüşme teknikleri kullanılır.

Gözlem Yöntemi

Bu yöntem, psiko-motor davranışların ölçümünde ve sınıf içi faaliyetler hakkında bilgi toplamada kullanılır. Gözlem yoluyla hem niceliksel, hem de nitelikse veriler toplanabilir. Yapılan gözlemin objektif olabilmesi için çeşitli önlemlerin alınması gerekir.

Tarihi (Tarama) Yöntem

Program değerlendirme çalışmalarında değerlendirilmesi yapılacak olan program ögesine ilişkin bazı bilgilerin, yazılı dökümanlardan elde edilmesi yoluna gidilir. Bu gibi durumlarda tarihi yöntem kullanılır. Bu yöntemin kapsamına giren okul arşivindeki öğrenci, öğretmen ve diğer personel hakkındaki bilgiler, öğretmen planları, ders notları, not defterleri, örgüt içindeki yazışmalar, okulun araç-gereç durumu kayıtları vs. programlar hakkında önemli bilgiler verir.

İstatistiksel Yöntemler

Bu kısımda program değerlendirme çalışmalarında gerek deneysel, gerekse survey tipi araştırmalarda yoğun bir şekilde kullanılan ki-kare, “t” testi, varyans analizi istatistiksel analiz yöntemlerinin hangi durumlarda kullanılması gerektiği ve bilgisayarda kullanılan SPSS (10.0 versiyonu) paket programında bu yöntemlerin nasıl yapıldığı konusu üzerinde durulacaktır.

Araştırmalarda kullanılacak istatistiksel yöntemler; araştırmanın problem, alt problem ve denencelerine göre değişiklik gösterebileceği gibi; veri toplamada kullanılacak ölçek türleri de yöntem seçimini etkilemektedir (Karasar, 1999: 146; Siegel, 1977: 21). Ölçek türleri aynı zamanda veri türlerini ifade eder. Araştırmalarda genel olarak dört ölçek türü kullanılmaktadır (Turgut, 1983: 16-19; Arıcı, 1990: 13-18; Karasar, 1999: 143-144; Siegel, 1977: 25-34).

1. Sınıflama (Nominal) ölçekleri.- Bireylerin tembel-çalışkan, zengin-fakir olarak sınıflandırılması, ayrıca bir grubun öğrenim durumu, yaş grupları, öğrencilerin okudukları bölümler, cinsiyet durumları vs. bu ölçek türündendir. Burada her grup bir sayı, bir harfle adlandırılır. Örneğin, kızların 1, erkeklerin 2 olarak kodlanması gibi. Bu ölçekle elde edilmiş veriler üzerinde frekans, mod ve ki-kare hesaplamaları yapılabilir.

2. Sıralama (Ordinal) ölçekleri.- Sınıflandırma kategorileri kendi aralarında sıralanabiliyorsa bu ölçek kullanılır. Örneğin, herhangi bir derste öğrencilerin başarılarına göre “pekiyi-iyi-orta-zayıf” gibi gruba ayrılması, ya da boyların “küçük, orta, büyük” şeklinde sıralanıp, bunların dizideki yerini belirten numaraların (1,2,3,...gibi) verilmesi bu ölçek türüne girer. Burada temel ölçüt “birinden büyük ya da küçük olmak”tır. “Ne kadar büyük olduğu” önemli değildir. Yani gruplar arasında farkların ne kadar olduğu matematiksel olarak bilinmez. Bu tür verilerde frekans, ranj, mod, yüzdelik hesaplamaları yapılabilir.

3. Eşit aralıklı (Interval) ölçekler.- Ölçüm birimleri arasında matematiksel anlamın çıktığı türdür. Bu tür ölçeklerde bağıl bir başlangıç noktası vardır ve ölçek eşit birim ve aralıklara bölünmüştür. Eğitimde kullanılan bazı standart testlerin, bu tür ölçek niteliğinde puanlar verecek şartlarda ölçeklenmesine çalışılır. Genellikle, standartlaştırılmış puanların eşit aralıklı ölçek niteliğinde

olduğu kabul edilir. Termometre ve takvim eşit aralıklı ölçeklere örnek olarak verilebilir. Bu ölçeklerden elde edilen verilere ortalama, standart sapma, korelasyon vb. hesaplamalar uygulanabilir.

4. Eşit oranlı (Ratio) ölçekler.- Bir değişkenin iki değerinin birbirine oranının, başka iki değerinin birbirine olan oranına eşitse eşit oranlı ölçekler söz konusudur. Başlangıcında mutlak bir sıfır noktası vardır. Aralıklar arasında oranlamalar yapılabilir. Uzunluk, ağırlık vs. ölçümleri bu gruba girer. Bu tür ölçekler, önceki ölçeklerin en gelişmişidir. Oranlı ölçeklerin verilerinden toplama, çıkarma, ortalama vs. alınabilir ve verilere her türlü istatistiksel yöntemler uygulanabilir.

Bilimsel araştırmalarda elde edilen verilere uygulanacak olan istatistiksel teknikler, yukarıda açıklanan ölçek türlerine göre değişiklik gösterir. Bu nedenle, araştırmaların başlangıcında, hangi tür ölçek kullanılacağı, buna uygun istatistiksel tekniklerin neler olabileceğinin belirlenmesinde yarar vardır. Örneğin: Sınıflama ve sıralama ölçeklerle elde edilmiş verilere parametrik olmayan istatistiksel teknikler uygulanabilirken; eşit aralıklı ve oranlı ölçeklerden elde edilen verilere ise parametrik testler uygulanabilir. Parametrik testler, örneklemden alınan puanları toplar, böler, çarpar. Dolayısıyla bu testler, sadece sayısal verilerde kullanılır. Sayısal olmayan değerlendirmelere bu testler uygulanamaz. T testi, F testi, Pearson korelasyonları, aritmetik ve geometrik ortalama, standart sapma gibi değerlendirmeler parametrik türdür. Ölçümler, sınıflama ve sıralama ise, örneklem normal olmayan dağılımlardan alınmışsa parametrik olmayan testler kullanılmalıdır. Parametrik olmayan teknikler verilerin sayısal olmasına bakmayıp, onların sırasına (mertebe) bakar, hatta sıralamanın bile mümkün olmadığı sınıflandırmaya dayalı verilere de uygulanabilir. Ki-kare, frekans, yüzdelik, mod, medyan parametrik olmayan testlere örnek olarak gösterilebilir (Siegel, 1977: 32; Ergün, 1995: 186).

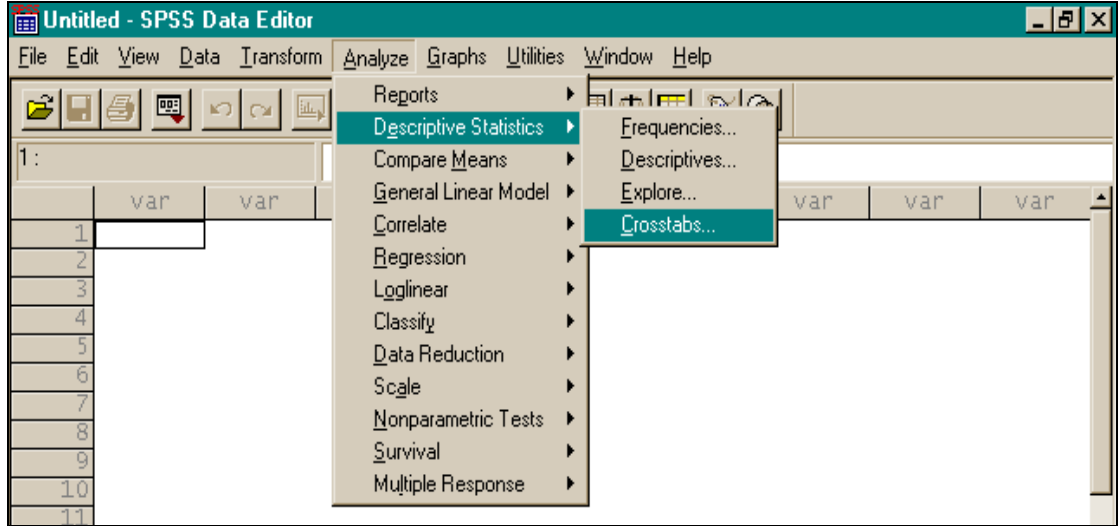
Parametrik testlerin uygulanabilmesi için örneklemin normal bir dağılım göstermesi ve varyansların eşitliği varsayımının sağlanması gerekir (Siegel, 1977: 22). Eğer bu varsayımlar karşılanmıyorsa; parametrik olmayan testler kullanılır. Örneğin: Bağımsız t testi yerine Mann-Whitney U Testi; bağımlı t testi yerine Wilcoxon-İşaretli-Sıralama Testi; F testi yerine Kruskal-Wallis testi uygulanır (Balcı, 1995: 269).

Ki-Kare Testi

Sınıflama (nominal) ölçekle elde edilen verilere uygundur. Elde edilen verilerin frekansa dayalı olması gerekir. Bu frekanslar iki ya da daha çok kategoriye yerleştirilir. Erkek-kız; geçti-kaldı, kısa-orta-uzun gibi veriler bu gruba girer. Örneğin: Uygulanan iki öğretim metodunun başarı durumuna etkisi; kişilerin eğitim düzeyleri ile belli konulardaki görüşleri (tutumları) arasındaki ilişkiler vs. ki-kare ile test edilir (Ergün, 1995: 85) .

SPSS programında ki-kare testi için; *Analyze/Descriptive Statistics/Crosstabs...* yolundan “Crosstabs” penceresine geçilir. Burada çapraz tablonun satır (row) ve sütun (Column) değişkenleri ilgili kutulara aktarılır. “Statistics...” düğmesiyle “Crosstabs: Statistics” penceresine geçilir ve burada

“Chi-square” kutusu işaretlenir. Elde edilecek çapraz tabloda, genelde satırlar (Row) bağımsız değişkeni (independent variable), sütunlar da (Column) bağımlı değişkeni (dependent variable) gösterir.



Şekil 1: Ki-kare Testi Uygulaması

Bu test; “iki veya daha fazla grubun belli kategorilere dağılmış olan frekans veya görüşleri arasında fark (veya ilişki) var mıdır?”; H_0 , “gruplar arasında bir fark yoktur” prensibine dayanır. İki değişken arasındaki ilişki (relationship) bulunmaması (H_0) hipotezinin test edilmesi gerekir. Test sonucunda p değeri, başta öngörülen anlamlılık düzeyinden (yaygın olarak kullanılan anlamlılık düzeyleri 0.05 veya 0.01’dir) küçük eşit ya da çıkarsa ($p \leq 0.05$), H_0 reddedilir ve gruplar arasında fark (ilişki) olduğu kabul edilir (Ergün, 1995: 86). Ayrıca, test sonucunda bulunan ki-kare değeri; ki-kare tablo değerinden eşitse/büyükse yine “farklılık vardır” hipotezi kabul edilir. Örneğin; belli bir konuda görüşleri alınan kişilerin, görüşlerinin cinsiyete göre farklılaşıp farklılaşmadığı, ki-kare ile test edildiğinde bulunan ki-kare değeri= 17.99, serbestlik derecesi= 5, p değeri= 0.002 bulunmuş olsun. Anlamlılık düzeyi de ister 0.05, isterse 0.01 kabul edilsin, örneğimizde bulunmuş olan p değeri (0.002), bunlardan küçük olduğu ($p < 0.01$) için H_0 “yokluk” hipotezi red edilecek, dolayısıyla, kişilerin görüşlerinin cinsiyet faktörüne göre farklılaştığı sonucuna varılacaktır. Ayrıca, örneğimizde bulunan ki-kare değeri (17.99), tablo değerinden (0.01 anlamlılık düzeyi için: 15.09) büyük olduğu için yine aynı sonuca varılacaktır.

Ki-kare testinin uygulanabilmesi için çapraz tablodaki hücrelerde beklenen değerin sıfır olmaması ve beklenen değeri 5’in altında olan hücre sayısının, toplam hücre sayısının % 20’sini geçmemesi gerekir (Karasar, 1999: 243; Siegel, 1977: 196). Bu sağlanmıyorsa çözüm olarak birbirine yakın kategorilerin anlamı bozmayacak şekilde birleştirilmesi yoluna gidilir. Ayrıca, 2*2 tablolarda N sayısının durumuna bağlı olarak Fisher testi ve Yates düzeltmesi uygulanır (Siegel, 121-122).

Ki-kare değerinin kendisi iki değişken arasındaki ilişkinin derecesini göstermede yetersizdir. Bunun için 2*2 tablolarda Phi katsayısı, dörtten çok hücreli tablolarda ise kontincensi katsayısı

hesaplanır. Bu korelasyon katsayıları, ki-kare tablosundaki değişkenler arası ilişkinin büyüklüğünü tahmin imkânı sağlar (Balcı, 1995: 268-269; Ergün, 1995: 94).

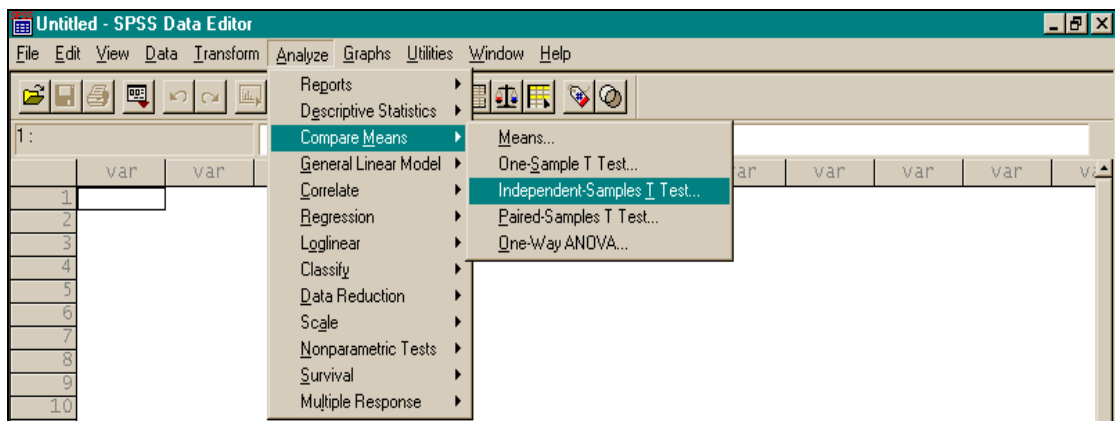
T testi

Gerek t testi, gerekse F testleri eşit aralıklı ya da oranlı ölçeklerle elde edilmiş verilere uygulanır. Kısaca, verilerin sayısal türden olması gerekir. Örneğin: Cinsiyet ve başarı arasındaki ilişki belirlenirken, kızlar ve erkeklerin iki grup halinde düşünülmesi ve başarı puanlarının karşılaştırılması t testi çözümlemesini gerektirir (Karasar, 1999: 2458).

Bu test iki grubun ortalamaları arasındaki farkın test edilmesinde kullanılır. t testinin, bağımsız t testi ya da bağımlı t testi olarak uygulanması, ortalamaların ilişkili ya da ilişkisiz olmasına bağlıdır. Ortalamalar; aynı gruptan alınmışsa bağımlı; farklı gruplardan alınmışsa bağımsız durum söz konusudur (Kaptan, 1991: 251).

Test analizi sonucunda bulunan anlamlılık düzeyi, genellikle kabul edilen 0.05 veya 0.01 seviyelerinden küçükse/eşitse ortalamalar arasında anlamlı bir fark olduğu kabul edilir (Ergün, 1995: 168).

SPSS programında bağımsız gruplardaki t testi için; **Analyze/Compare Means/Independent-Samples T Test...** yolundan “Independent-Samples T Test” penceresine geçilir. Burada “Grouping Variable” kutusuna bağımsız değişken aktarılır. “Define Groups...” kısmında ise bağımsız değişkenin sınıkları, Group 1 ve Group 2 kutularına yazılır. Gruplar bu şekilde belirlendikten sonra da bağımlı değişken (ler) “Test Variable(s)” kutusuna aktarılarak işlem gerçekleştirilir. Örneğin: Bir öğretim yönteminin etkili olup olmadığı araştırılıyor olsun. Öğretimin sonunda deney ve kontrol grubundaki öğrencilere uygulanan sontest sonuçları arasında anlamlı bir farklılık olup olmadığını inceleyelim. Bu durumda uygulanan bağımsız gruplardaki t testi sonuçları şu şekilde bulunmuş olsun: t değeri= 3.26, serbestlik derecesi= 160, p değeri= 0.001. Örneğimizde bulunan p değeri (0.001), kabul edilen 0.05 ya da 0.01 anlamlılık düzeyinden küçük olduğu ($p < 0.05$) için grupların sontest sonuçları arasında anlamlı farklılık olduğu söylenebilir. Bulunan fark, deney grubunun lehine ise, uygulanan öğretim yönteminin etkili olduğu sonucuna varılabilir.



Şekil 2: Bağımsız Gruplarda “t” testi Uygulaması

SPSS programında bağımlı gruplardaki t testi için ise; **Analyze/Compare Means/Paired-Samples T Test...** yolundan “Paired-Samples T Test” penceresine geçilerek yapılır.

Varyans Analizi

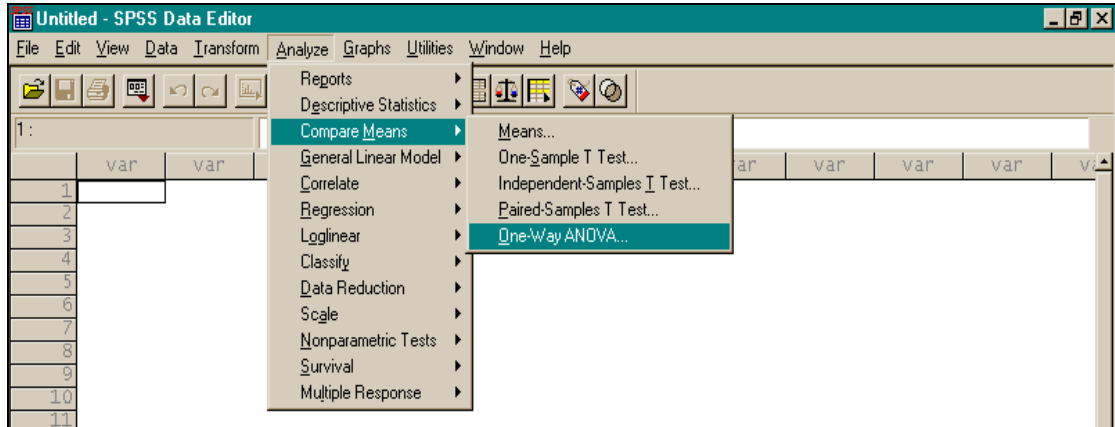
Bir veri kümesini betimleyen iki özellikten biri aritmetik ortalama, diğeri de varyanstır. İki ortalama arasındaki farkın test edilmesinde t testi kullanılır. Karşılaştırılacak olan ortalamalar ikiden fazla olduğunda önce varyans analizi yapılır. varyansın karşılaştırılmasında F testi kullanılır. Varyans analizinin birçok çeşitleri vardır (Baykul, 1999: 394). Burada en yalın olan tek yönlü varyans analizi incelenmiştir.

Tek yönlü varyans analizi (Anova).- Örneğin: Dört ayrı öğretim yönteminin her birinin örnekleme yoluyla seçilmiş dört ayrı gruba uygulanması sonucunda; bu yöntemlerden hangisinin daha etkili olduğunu anlamak başka bir değişle yöntemler arasında etkililik yönünden bir fark olup olmadığını anlamak için tek yönlü varyans analizi kullanılabilir. Burada, dört ayrı grubun ortalamalarının karşılaştırılması söz konusudur. (Kaptan, 1991: 259).

Ortalamalarının karşılaştırılması içinde; eğer karşılaştırılacak gruplar ikiden fazla ise, t testi yapılmadan önce varyans analizi yapılır. Varyans analizi sonucunda, F değerinin anlamlı bulunması, gruplardan en az iki tanesinin farklı olduğunu göstermektedir. Bu durumda hangi grupların (ortalamaların) farklı olduğunu bulmak için spesifik bir tür t testi kullanılır. Bu amaçla, çoklu karşılaştırmalar t testlerinden biri (Scheffe testi, Tukey testi, Duncan çoklu genişlik testi, Newman-Keuls testi) kullanılabilir (Kaptan, 1991: 259; Balcı, 1995: 266). Örneğin; bir F testi sonucunda gruplar arası serbestlik derecesi= 4, gruplar içi serbestlik derecesi= 357, F değeri= 2.817, p değeri= 0.025 bulunmuş olsun. Örneğimizdeki p değeri (0.025), kabul edilen 0.01 anlamlılık düzeyinden büyük olduğu ($p > 0.01$) için gruplar arasında anlamlı bir farklılık yoktur. Dolayısıyla grupları ikili karşılaştırmaya gerek bulunmamaktadır. Eğer anlamlılık düzeyi 0.05 olarak alındığı takdirde, örneğimizdeki p değeri, bundan küçük olacağı ($p < 0.05$) için gruplar arasında anlamlı bir farklılık bulunduğu sonucuna varılabilir. Bunun devamında ise, grupların ikiyeşerli karşılaştırmalarının yapılması gerekir. Ayrıca, bu örneğimizde, bulunan F değeri (2.817), F değerler tablosundan, -örneğin; 0.01 düzeyinde, pay için= 4, payda için= 357 serbestlik derecesine göre- okunan değerden (3.32) küçük olduğu için, bulunan F değerinin anlamlı olmadığı, dolayısıyla gruplar arasında anlamlı bir farklılığın bulunmadığı da söylenebilir.

SPSS programında tek yönlü varyans analizi yapmak için; **Analyze/Compare Means/One-Way ANOVA...** yolundan “One-Way ANOVA” penceresine geçilir. Burada “Dependent List” kutusuna bağımlı değişkenler; “Factor” kutusuna ise bağımsız değişkenler aktarılır. “Post Hoc...” düğmesiyle “One-Way ANOVA: Post Hoc Multiple Comparisons” penceresine geçilir ve burada grupların ikiyeşerli karşılaştırılmasında kullanılacak olan test (Scheffe, Tukey, Duncan...) işaretlenir.

“Options...” düğmesiyle de “One-Way ANOVA: Options” penceresine geçilir ve buranın “Statistics” kısmında “Descriptive” seçeneği işaretlenir.



Şekil 3: Tek Yönlü Varyans Analizi Uygulaması

T testi ve F testi; fark-ilişki vardır, yoktur şeklinde bir sonuç verir. Ancak, ilişki varsa bunun miktarını vermez. İlişkinin miktarı bulunmak istenirse, t testini izleyerek ondan sonra Pearson r'si bulunabilir. F testinden sonra ise eta-korelasyon oranı uygulanabilir (Balcı, 1995: 270).

Kaynakça

- Arıcı, Hüsnü. **İstatistik Yöntem ve Uygulamalar**. Ankara: Meteksan Matbaacılık, 1990.
- Balcı, Ali. **Sosyal Bilimlerde Araştırma Yöntem, Teknik ve İlkeler**. Ankara: 72 TDFO Bilgisayar-Yayınçılık, 1995.
- Baykul, Yaşar. **İstatistik Metodlar ve Uygulamalar**. Anı Yayıncılık, 1999.
- Demirel, Özcan. **Kuramdan Uygulamaya Eğitimde Program Geliştirme**. Ankara: Kardeş Kitap ve Yayınevi, 1998.
- Erden, Münire. **Eğitimde Program Değerlendirme**. Ankara: Anı Yayıncılık, 1998.
- Ergün, Mustafa. **Bilimsel Araştırmalarda Bilgisayarla İstatistik Uygulamaları SPSS For Windows**. Ankara: Ocak Yayınları, 1995.
- Ertürk, Selahattin. **Eğitimde Program Geliştirme**. Ankara: Yelken Tepe Yayınları, 1986.
- Hakan, Ayhan (Ed.) **Eğitim Bilimlerinde Çağdaş Gelişmeler**. Eskişehir: Anadolu Üniversitesi Açıköğretim Fakültesi Yayınları, 1991.
- Kaptan, Saim. **Bilimsel Araştırma ve İstatistik Teknikleri**. Ankara: Tekışık Web Ofset Tesisleri, 1991.
- Karasar, Niyazi. **Bilimsel Araştırma Yöntemi**. Ankara: Nobel Yayın Dağıtım, 1999.
- MEB. **Tebliğler Dergisi**. Sayı: 2428, 1995.
- Siegel, Sidney. **Davranış Bilimleri İçin Parametrik-Olmayan İstatistikler**. (Çev.:Yurdal Topsever), Ankara: Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Yayınları, 1977.
- Turgut, M. Fuat. **Eğitimde Ölçme ve Değerlendirme Metotları**. Ankara: Saydam Matbaacılık, 1983.
- Turgut, M. Fuat. “Öğretimin Değerlendirilmesi” **Eğitim Yönetiminde Denetleme ve Değerlendirme Sempozyumu**. Ankara: Ankara Üniversitesi Eğitim Bilimleri Fakültesi, 25-26 Nisan 1985.
- Turgut, M. Fuat. “Program Değerlendirme” **Cumhuriyet Döneminde Eğitim**. İstanbul: Milli Eğitim Bakanlığı Yayınları, 1983 a, ss. 215-234.
- Varış, Fatma. **Eğitimde Program Geliştirme Teori ve Teknikler**. Ankara: Alkım Kitapçılık Yayıncılık, 1996.