

Fen Bilimleri, Bilgisayar ve Matematik Eğitimi Alanlarındaki Tez Çalışmalarının İstatistiksel Açından İncelenmesi

Tolga Kabaca¹,

Yavuz Erdoğan²

Özet

Bu araştırmanın amacı, fen bilimleri ve matematik eğitimi alanlarında yapılan yüksek lisans ve doktora tezi çalışmalarındaki istatistiksel hataları incelemektir. Bu amacı gerçekleştirmek üzere, İstanbul, Ankara ve İzmir'deki üç devlet üniversitesinin kütüphanelerinde bulunan fen bilimleri ve matematik eğitimi alanlarıyla ilgili tezlerin listesi çıkarılmıştır. Bu listeden; Bilgisayar, Biyoloji, Fen Bilgisi, Fizik, Kimya ve Matematik eğitimi alanları ile ilgili 129 adet tez çalışması random olarak seçilmiştir. İncelenen tez çalışmalarındaki hatalar yedi farklı alt boyutta değerlendirilmiştir. Bu boyutlar; veri toplama araçlarının geçerlilik ve güvenilirlikleri ile ilgili hatalar, örneklem seçimi ile ilgili hatalar, betimsel istatistiklerin kullanımına yönelik hatalar, normal dağılım tespitine yönelik hatalar, parametrik ve parametrik olmayan testlerin kullanımı ile ilgili hatalar, ifade hataları ve biçimsel hatalardır. Araştırma sonucunda, birçok tez çalışmasında istatistiksel açıdan hatalı kullanımların olduğu gözlenmiştir. Bu hatalar, bazı alt boyutlarda yıllara bağlı olarak değişim göstermektedir. Araştırma sonucunda elde edilen bulgular tartışılarak, istatistik tekniklerinde yaşanan gelişmelere göre çözüm önerileri sunulmuştur.

Anahtar Kelimeler: Tez araştırmaları, İstatistik, İstatistiksel hatalar

Examining Of Statistical Properties of the Thesis Studies in the Fields of Science, Computer and Mathematics Education

Abstract

The purpose of this study is to examine the statistical mistakes of master and doctoral theses in the field of science and mathematical education. With this aim in mind, from the libraries of three state universities which are in İstanbul, Ankara and İzmir, the list of master and doctoral theses about science and mathematical education was determined. From this list, 129 different theses were chosen randomly in the field of biology, chemistry, computer, mathematic, physics and science education. These theses were examined in seven different factors. These factors can be stated as reliability and validity mistakes, sample choosing mistakes, descriptive statistics mistakes, determining normality distribution mistakes, using parametric and non-parametric statistical techniques, expression and formal mistakes. At the end of the study, it was found that science and mathematical education's theses have many statistical mistakes. In addition, these mistakes are changed in respect of years.

Key Words: Thesis researches, Statistic, Statistical mistakes.

¹ Öğr.Gör.Dr. Uşak Üniversitesi Fen Edebiyat Fakültesi, Matematik Bölümü, UŞAK. e-posta: tolgakabaca@yahoo.com

² Öğr.Gör.Dr. Marmara Üniversitesi, Atatürk Eğitim Fakültesi, BÖTE Bölümü, İstanbul. e-posta: yavuzerdogan@gmail.com

Giriş

Bilimsel yöntem, olgusal nitelikli problem çözmenin, bilim üretmenin bilinen ve belli süreçleri olan, en güvenilir yolu olarak kabul edilir. Bu yöntem; açık seçiktir, denetlenebilir, eleştirici ve düzelticidir ve tümüyle yansızlık temeline oturtulmuştur. Bilimsel yöntem, belli kestirimlerin doğru sayılıp sayılmayacağına karar verebilmek için yürütülen sınamaları içerir ve problemlere doyurucu çözümler buluncaya kadar bu sınamalar sürdürülür. Bu amaçla, duyarlılığı yüksek ölçme değerlendirme süreçlerini içerir ve istatistik biliminden yararlanır. İstatistik bir tür bilim grameri olup, matematiğin bilimsel araştırmalara uygulanışından doğmuştur (Karasar, 2003, s.207). Bir araştırmacı, çalışmasının amaçlarını yerine getirebilmek için istatistik tekniklerden yararlanmak durumundadır. Hatta bir bilimsel araştırmanın tasarlanma aşamasında bile istatistik bilgisi kaçınılmaz olabilmektedir. Örneğin, örneklemelerden alınan istatistikler yardımıyla geleceğe yönelik yordamalar yapılabilir. Eğer regresyon analizi bilinmezse, geleceğe yönelik yordamsal çözümlemeler yapmak oldukça zordur. Aynı şekilde, parametrik olmayan teknikler hakkında bilgi sahibi olmayan bir araştırmacı, sınıflama ölçeği ile elde edilen kategorik verilerle veya sıralama ölçekli verilerle parametrik teknikler kullanma hatasına düşebilir. Bunun gibi birçok örnek sıralanabilir. Bu hatalar, bulguların yüzeysel olarak değerlendirilmesine, hatta yanlış yorumlanmasına neden olabilir.

Bilimsel makaleler, tamamlanmış bir bilimsel çalışmanın patenti niteliğindedir ve bir ülkenin bilimsel ilerleyişi bu çalışmalara göre değerlendirilir. Ülkemizde yayımlanan binlerce makaleye karşın atıf sayılarının yetersiz olması, üretilen makalelerin bilimsel ölçütler açısından tekrar ele alınması gerekliliğini ortaya koymaktadır. Bir çalışmanın bilimselliği de kullandığı araştırma ve istatistiksel yöntemlerin sağlamlığıyla ölçülür (Erkuş, 2004, s.176).

Bu noktalardan hareketle, mevcut araştırmada; fen bilimleri ve matematik eğitimi alanlarında yapılan yüksek lisans ve doktora tezi çalışmalarındaki istatistiksel hatalar incelenmiş ve çözüm önerileri sunulmuştur. Bu amacı gerçekleştirmek üzere, tez çalışmaları aşağıda sunulan alt başlıklar doğrultusunda değerlendirilmiştir:

- Tezlerde kullanılan istatistiksel tekniklerin seçimindeki hatalar
- İstatistiksel analizler ile araştırma amaçlarının uyumuna yönelik hatalar,
- Örneklem seçimi ile ilgili hatalar,
- Betimsel istatistikler ile ilgili hatalar,
- Veri toplama araçlarının geçerlik ve güvenilirlik çalışmaları ile ilgili hatalar,
- Normal dağılım testleri ile ilgili hatalar,
- Parametrik testlerin seçimi ile ilgili hatalar,
- Parametrik ve parametrik olmayan testlerin kullanımı ile ilgili hatalar,
- İfade hataları ve biçimsel hatalar.

Yöntem

Bu çalışmada, fen bilimleri ve matematik eğitimi alanlarında yapılan yüksek lisans ve doktora tezi çalışmalarındaki istatistiksel hataları incelemek amacıyla nitel araştırma yönteminden yararlanılmıştır. Nitel araştırmalar, toplanan verilerin analizinde tümevarıma dayalı bir yaklaşım izler. Araştırmacı, topladığı verilerden yola çıkarak, incelediği probleme ilişkin ana temaları ortaya çıkararak, anlamlı bir yapıya kavuşturur (Glaser ve Straasfasf, 1967; akt.; Şimşek ve Yıldırım, 2000. s.19). Bu çalışmada ise 129 adet tez çalışması incelenerek, tezler hakkında genel bir kanaat oluşturulması hedeflenmiştir.

Araştırma Grubu

Doküman incelemelerinde, evren-örneklem çerçevesi belirlenirken, sosyal bilimlerde sıklıkla kullanılan klasik yaklaşımlardan yararlanılabilir (Şimşek ve Yıldırım, 2000. s.151). Bu çalışmada evren, İstanbul, Ankara ve İzmir’de bulunan üç devlet üniversitesinin kütüphanelerinde yer alan fen bilimleri ve matematik eğitimi alanlarında tamamlanmış tezlerdir. Örneklem ise bu tezlerden rasgele seçilen 129 adet tez çalışmasıdır. Biyoloji eğitimi ile ilgili 14, Fen Bilgisi eğitimi ile ilgili 26, Fizik eğitimi ile ilgili 21, Kimya eğitimi ile ilgili 23, Matematik eğitimi ile ilgili 29 ve Bilgisayar eğitimi ile ilgili 16 çalışma örnekleme dâhil edilmiştir.

Veri Toplama Araçları

Verilerin toplanmasında, nitel araştırmalarda veri toplama yöntemlerinden biri olan doküman incelemesi yaklaşımı kullanılmıştır. Doküman incelemesi, araştırılması hedeflenen olgu ve olgular hakkında bilgi içeren yazılı materyallerin analizini kapsar (Şimşek ve Yıldırım, 2000, s.140). Bu araştırmada, tez çalışmalarındaki istatistiksel hataların incelenmesi amacıyla, 11 maddelik tez analiz formu geliştirilmiştir.

İşlem

İstanbul, Ankara ve İzmir’deki üç üniversitenin kütüphanelerindeki fen bilgisi ve matematik eğitimi alanında tamamlanmış tezler geliştirilen tez analiz formu yardımı ile tek tek incelenmiştir. Tez analiz formu yardımı elde edilen bilgiler daha sonra derlenerek analizlerin yapılabilmesi için uygun hale getirilmiştir.

Verilerin Çözümlemesi

Tezlerde karşılaşılan istatistik hatalarının değerlendirilmesi amacı ile yüzde istatistiksel teknikten yararlanılmış, ayrıca bazı istatistik hatalarının yıllara bağlı değişiklik gösterip göstermediği Kay-kare bağımlılık testi kullanılarak analiz edilmiştir. Araştırmacının istatistiksel işlemleri SPSS 13.0 paket programıyla gerçekleştirilmiştir. Araştırma kapsamında kullanılan tüm istatistiksel işlemlerde anlamlılık düzeyi 0.05 olarak kabul edilmiştir.

Bulgular

Bu bölümde, tez analiz formunda yer alan her bir madde tek tek incelenerek, yorumlanmıştır.

Tez Çalışmalarında Kullanılan İstatistiksel Teknikler ve İstatistiksel Tekniklerin Seçimindeki Hatalar

Tez çalışmalarının büyük çoğunluğunda bağımlı ve bağımsız grup t-testleri kullanılmıştır (N=95, %73.6). Daha sonra, Tek Yönlü Varyans Analizi (ANOVA) tercih edilmiştir (N=59, %45.7). Bu bulgulardan da anlaşılabacağı üzere, Fen Bilimleri ve Matematik Eğitimi alanındaki tezlerde genel olarak fark analizleri tercih edilmiştir. Korelasyon analizleri sadece 11 (%8.5) tezde kullanılmıştır (tezlerde kullanılan korelasyon analizlerinin tamamı Pearson Momentler Korelasyon katsayısı ile gerçekleştirilmiştir). Bu arada 10 (%7.8) çalışmada Kay-kare testi kullanılmış, fakat bu analizlerin tamamı amacından farklı olarak hatalı şekilde kullanılmıştır. Tezlerden 9 (%6.9) tanesinde non-parametrik testler kullanılmış, 19 tezde ise regresyon analizi kullanılmıştır.

İstatistik tekniklerle ilgili dikkati çeken başka bir unsur; hiçbir tez çalışmasında MANOVA veya MANCOVA tekniklerinden yararlanılmamıştır. Sadece birkaç tez çalışmasında ANCOVA analizi kullanılmıştır. Araştırmalarda birden fazla bağımlı değişken varken, tek bağımlı değişken varmış gibi istatistiksel analizler yapılmıştır. Örneğin, hem başarı hem de kaygının birlikte araştırıldığı bir çalışmada MANOVA gibi çok değişkenli istatistikler kullanmak yerine, her bir bağımlı değişken tek tek ele alınmış ve ANOVA gibi tek yönlü varyans analizi ile istatistikler gerçekleştirilmiştir.

Dikkati çeken başka bir bulgu; deneysel araştırmalarda grupların eşitliği, sadece ilgili bağımlı değişkene göre gerçekleştirilmiş ve kontrol değişkenleri dikkate alınmamıştır. Örneğin, kontrol ve deney gruplarının $\bar{X}$ dersi başarıları, sadece $\bar{X}$ başarı testine göre karşılaştırılmış; fark görülmediği durumlarda uygulamaya geçilmiştir. Aslında, $\bar{X}$ başarısına etki ettiği düşünülen $\bar{Y}$ tutumları veya Q kaygısı, başarı kadar uygulama sonucunu etkileme potansiyeline sahiptir. Dolayısı ile bu tür değişkenlerin Kovaryans Analizleri veya Kısmi Korelasyon tekniği kullanılarak kontrol altında tutulması, daha sağlıklı sonuçlar elde edilmesini sağlayabilir.

İstatistiksel Analizler ile Araştırma Amaçlarının Uyumuna Yönelik Hatalar

Tez çalışmalarındaki istatistiksel analizlerin tezin amaçlarına uygunluğu incelendiğinde, 94 tezin (%72.8) amaçlarıyla uyumlu olduğu; diğerlerinin ise (%27.2) uyumsuz olduğu belirlenmiştir. Bazı tezlerde dikkate değer hatalar gözlenmiştir. Bu hatalardan bazı örnekler aşağıda sunulmuştur¹.

- a. Araştırmanın bağımlı ve bağımsız değişkenleri tam olarak belirlenmemiş ve bu duruma bağlı olarak yanlış anlaşılacak alt problemler kurgulanmıştır. Bağımlı değişken, araştırmacının açıklanmasını istediği durumdur; bağımsız değişken ise bağımlı değişkene etki ettiği düşünülen değişkendir (Levin ve Rubin, 1990. s.521). Bu hatayla ilgili bazı örnekler aşağıda sunulmuştur;

¹ Çalışmada, örneklendirmeler yapılırken, etik olmayacağı düşüncesiyle bazı gizlemeler yapılmıştır. Gizlemeler; X, Y, Q gibi harflerle ifade edilmiştir.

- “X kavramı testi, cinsiyet üzerinde farklılığa neden olmakta mıdır?”. Bu cümleye göre, bağımlı değişken cinsiyettir. Bu problemin eğitim sahasının dışında olduğu değerlendirilmektedir.
- “Ortaöğretimde İngilizce Y eğitiminin başarıya etkisi var mıdır?”. Bu çalışmada neyin araştırıldığı net olarak ifade edilmemiştir. Tez incelendiğinde, Q eğitimi İngilizce olarak verildiğinde, Q başarısının bundan etkilenip etkilenmediğinin araştırıldığı gözlenmiştir.
- b. Amaçlar bölümünde hipotezler, H0 (Yokluk Hipotezi) şeklinde kurgulanmış. Hâlbuki amaçlar bölümünde yer alan hipotezler, H1 (araştırma hipotezi) şeklinde sunulmalıdır. Çünkü araştırmacı yargıç rolündedir (Karasar, 2003, s.70) ve bir ilişkiyi/farkı sınar. H0 değişkenler arasındaki ilişkisizliği savunan bir ifadedir. Araştırmacı, değişkenler arasında ilişki olmadığına inanıyorsa, o zaman bu ilişkiyi test etmesi doğru olmayabilir. Bazı örnekler;
- “X öğretim metoduyla ders anlatılan deney grubunun sontest başarısı ile kontrol grubunun sontest başarısı arasında anlamlı farklılık yoktur”.
- “Öğrencilerin Y dersi tutumları ile Y dersi başarıları arasında anlamlı ilişki yoktur”.
- c. Hipotezler yanlış şekilde ifade edilmiştir. ;
- “Kontrol grubunun ön testleri, son testlerinden az oranda anlamlı yüksektir”. Sosyal bilimlerde, iki ölçüm sonucu arasında tespit edilen .05 veya .01 düzeyindeki farklılıklar yeter şarttır. Bu sınırların “az oranda” veya “çok oranda” gibi ifadelerle güçlendirilmesi akademik literatüre uygun değildir.
- “X öğretim metodu, ülkemiz eğitim sisteminde alternatif bir ölçme değerlendirme aracı olarak kullanılacaktır”. Bu ifade, bu araştırma sonucunda test edilemeyecek bir hipotezdur.

Örneklem Seçimi ile İlgili Hatalar

Evren, araştırma sonuçlarının genellenmek istendiği elemanlar bütünü; örneklem ise evreni temsil yeterliliği olduğu kabul edilen küçük kümedir (Karasar, 2003, s.111). Bu araştırma kapsamında incelenen tezlerin büyük çoğunluğunda (N=90, %69.7) örneklem seçimi ile ilgili hatalar yapıldığı belirlenmiştir. Tezlerin çoğunluğunda yapılan örneklem hataları şu şekilde özetlenebilir;

- d. Deneyisel çalışmaların tamamına yakınında evren-örneklem ikilisi kullanılmış ve yanlış genellemelere gidilmiştir. Deneyisel araştırmaların asıl amacı, ulaşılan sonuçları bir evrene genellemek değildir (bu yorum; deneyisel çalışmalarda genelleme yapılamaz şeklinde anlaşılmamalıdır). Bu tür araştırmalarda sonuçlar, uygulama yapılan grup için geçerlidir. Uluslararası literatürde yer alan deneyisel araştırmalarda, evren örneklem yerine, katılımcılar (participants) veya çalışma grubu (study sample) gibi ifadeler tercih edilmektedir. Bu hatayla ilgili bazı örnekler aşağıda sunulmuştur;
- Bir ilköğretim okulunun 7. sınıflarında deneyisel bir araştırma yapılmış ve sonuçlar okulun tamamına genellenmiştir. İstatistiksel bakımdan, 7.sınıftaki bir gruba ait bulguların 1. sınıf öğrencilere genellenmesi gerçekçi olmayabilir.
- Kadıköy İlçesine ait 2 anaokulunda deneyisel bir araştırma yapılmış ve sonuçlar Kadıköy’deki tüm anaokullarına genellenmiştir. Bu yaklaşım doğru olmayabilir. Çünkü Kadıköy ilçesi gerek sosyo-ekonomik, gerekse sosyo-kültürel değişkenler

bakımında oldukça kozmopolit bir ilçedir. Birkaç anaokulunda yapılan çalışmanın Kadıköy gibi bir ilçeye genellenmesi araştırmacıları yanıltabilir.

Betimsel İstatistikler ile İlgili Hatalar

Betimsel istatistikler, gözlemlerimiz sonucunda elde ettiğimiz bilgileri yorumlamak için kullandığımız temel istatistiklerdir (ortalama, medyan, standart sapma vb) ve çoğu araştırmalar bu tür çözümlemeleri gerektirir niteliktedir (Karasar, 2003, s.209). İncelenen tezlerin 77'sinde (%59.7) betimsel istatistikler verilirken, 52'sinde (%40.3) verilmemiştir.

Veri Toplama Araçlarının Geçerlik ve Güvenirlik Çalışmaları ile İlgili Hatalar

Araştırmalarda kullanılan ölçme araçlarının geçerliği, ölçmek istediği özelliği başka şeylerle karıştırmadan doğru olarak ölçmesi; güvenilirliği ise ölçüm sonuçları arasındaki tutarlılıktır (Gronlund ve Linn 1990, s.48). Dolayısı ile geçerlik-güvenirlik çalışması tam olarak yapılmamış bir ölçeğin ölçüm sonuçlarına da güvenilemez. Çünkü bu ölçek, başka bir amacı test ediyor olabilir. Mevcut araştırmadaki tezlerden sadece 30'unda (%23.2) geçerlik-güvenirlik çalışması yapılırken, 40 tanesinde (% 31) hiçbir çalışma yapılmamıştır. Geri kalan 59 (%45.7) tez çalışmasında ise sadece Cronbach Alfa güvenirlik katsayısı incelenmiştir. Geçerlik-güvenirlik çalışmalarının yıllara bağlı olarak değişip değişmediğini belirlemek amacıyla uygulanan Kay-kare analizi sonucunda, anlamlı bağımlılık saptanmıştır ($\chi^2=4.78$; $p<0.05$). Bu bulguya göre 2000 yılından sonra gerçekleştirilen tezlerde geçerlik-güvenirlik çalışmalarına daha fazla yer verildiği belirlenmiştir.

Bununla birlikte, incelenen 129 tezden hiçbirisinde faktör analizi gibi yapı geçerliğini test eden geçerlik analizlerine ve kriter geçerliğine rastlanmamıştır. Geçerlik çalışmaları genel olarak uzman görüşleri ile güvenirlik çalışmaları ise içtutarlılık katsayısıyla sınırlandırılmıştır. Bazı tezlerde çok yüksek Cronbach Alfa güvenirlik katsayıları belirlenmiştir ($\alpha=.98$ gibi). Aslında, geçerlik çalışması yapılmayan bir ölçeğin, çok yüksek içtutarlılık göstermesi, o ölçeğin çok iyi bir ölçüm aracı olduğunu da işaret etmez. Çünkü ölçüm sonuçları arasındaki tutarlılık yanlış bir amaç üzerinde de sağlanabilir. Dikkat çekici bir başka bulgu ise bazı araştırmalarda kullanılan kişisel bilgi anketlerinin içtutarlılık katsayıları belirlenmiştir. Bu tarz anketlerde, bağımlı değişkene etki ettiği düşünülen bağımsız değişkenler verilir ve anket maddeleri genel olarak kategorik yapıdadır. Kategorik veriler eşit aralıklı ölçeklerden elde edilmediğinden dolayı geçerlik-güvenirlik analizleri yapmak sağlıklı olmayabilir.

Normal Dağılım Testleri ile İlgili Hatalar

İstatistiksel sonuç çıkarmada elimizdeki verilerin normal dağılım göstermesi çok önemlidir ve verilerin normal dağılımdan gelip gelmediğinin mutlaka test edilmesi gerekmektedir (Ergün, 1995, s.63; Akgül, 2005, s.108). Normal dağılım gösteren verilerin ortalama, ortanca, tepe değerleri birbirine eşittir ve dağılım süreklidir (Bluman, 1995, s.228). Araştırma kapsamında incelenen tezlerin 36'sında (%27.9) normal dağılım testleri uygulanırken, geri kalan 93 (%68.7) tezde ise normal dağılım testlerinin uygulanmadığı gözlenmiştir. Normallik varsayımları,

incelenen tezlerin tamamında Kolmogorov-Smirnov Uyum İyiliği Testi ile denirken; Shapiro Wilk gibi alternatif testlerden hiç faydalanılmamıştır. Normal dağılım testlerinin yıllara göre kullanım oranları incelenmiş ve .05 düzeyinde anlamlı ilişkiler tespit edilmiştir ($\chi^2=7.634$; $p<0.05$). Bu bulguya göre, 2000 yılı öncesinde 5 tezde normallik varsayımı kullanılırken, 2000 yılından sonra ise 31 tezde kullanılmıştır.

Parametrik Testlerin Seçimi ile İlgili Hatalar

Araştırmalarda parametrik testlerin kullanılabilmesi bazı şartlara bağlıdır. Bu şartlar şu şekilde özetlenebilir (Bluman, 1995, s.492; Ergün, 1995, s.63; Orhunbilge, 1997, s.247; Levin ve Rubin, 1998, s.596; Akgül, 2005, s.164;).

- Araştırmada kullanılan veriler eşit aralıklı ölçeklerden elde edilmelidir.
- Bağımlı değişkene ilişkin ölçüm sonuçları normal dağılım sergilemelidirler.
- Ölçüm araçları geçerli ve güvenilir olmalıdır.

İncelenen tezlerden bahsi geçen şartlara haiz sadece 17 (%13.2) çalışma belirlenmiştir. Geri kalan 103 (%79.8) tezde ise parametrik testlerin seçimi usulüne göre yapılmamıştır. Bazı tezler; parametrik istatistikler kullanılmadığından değerlendirilmeye tabi tutulmamışlardır. Parametrik testlerin kullanımı ile ilgili hatalar incelendiğinde 99 tezde geçerlik-güvenirlik çalışmasının yeterli olmadığı, 93 tezde ise verilerin normallik varsayımının test edilmediği belirlenmiştir. Diğer yandan, birçok tezde kategori başına düşen denek sayısı son derece düşük tutulmuştur. Her bir grupta 30 denek sayısına ulaşılması matematiksel bir sınır olmamakla birlikte, gözlem sonuçlarına dayalı önemli bir varsayımdır. Bununla birlikte, bazı tez çalışmalarında, kategori başına 2-3 kişiyle hatta 0 kişiyle (!) bile karşılaştırmalar yapılmıştır. Bu durum istatistiksel açıdan kabul edilemez. Parametrik testlerin seçimi ile ilgili hatalar yıllara bağlı olarak incelenmiş ve bağımlılık olmadığı belirlenmiştir ($\chi^2=0.546$; $p>0.05$).

Parametrik ve Parametrik Olmayan Testlerin Kullanımı ile İlgili Hatalar

İncelenen tezlerden 84 tanesinde (%65.1) istatistiksel hatalar saptanmıştır. İstatistiksel hataların yıllara göre değişimi Kay-kare testi ile incelenmiş, fakat anlamlı bağımlılık tespit edilememiştir ($\chi^2=0.620$; $p>0.05$). Bununla birlikte tezlerde dikkate değer hatalar gözlenmiştir. Bu hatalardan bazı örnekler aşağıda sunulmuştur:

- a. Bir gruba ait 2 farklı veri seti, Bağımsız Grup t-Testi ile karşılaştırılmıştır. Örneğin deney grubunun X başarıları ile X tutumları Bağımsız Grup t-Testi ile karşılaştırılmıştır. Hâlbuki Bağımsız Grup t-Testleri, kendi içinde iki kategoriye ayrılmış bir süreksiz değişken cinsinden bir sürekli değişkenin karşılaştırılması amacıyla kullanılır. Yani, iki grubun bir değişken cinsinden mukayesesinde kullanılır (Büyüköztürk, 2002, s.39, Akgül, 2005, s.204). Tek gruba ait farklı veriler arasında fark analizi değil, korelasyon analizleri yapılabilir.

Başka bir tezde ise bir öğrenci grubunun 5 farklı testten aldıkları notlar, Tek Yönlü Varyans Analizi (ANOVA) ile karşılaştırılmıştır. Bu durumda ANOVA kullanamayız. Çünkü ANOVA, bir gruba ait 5 farklı sürekli değişkeni değil, kendi içinde 5 farklı kategoriye ayrılmış bir süreksiz değişkene göre bir sürekli değişkenin karşılaştırılması amacıyla kullanılır.

- b. Bazı tezlerde, bir ölçekte yer alan tüm sorular Bağımsız Grup t-Testiyle ayrı ayrı analize tabi tutulmuştur. Bağımsız Grup t-Testi kullanabilmek için verilerin eşit aralıklı ölçeklerden toplanması ve kümülatif (toplam puan üzerinden) olması gerekmektedir (Orhunbilge, 1997, s.247). Bu tezlerde olduğu gibi, sorular tek tek ele alındığı zaman veriler kategorik (sınıflamalı) yapıya dönüşür ve parametrik teknikler kullanmamız hatalı olur. Bu durumlarda Kay-kare bağımlılık analizi kullanılması gerekir (Büyüköztürk, 2002, s.139; Ergün, 1995, s.85).
- c. Deneysel araştırmaların çoğunluğunda karşılaşılan başka bir hata; deney veya kontrol gruplarına uygulama öncesi verilen öntestler ile uygulama sonrası verilen sontestler tamamen farklı olmasına rağmen Bağımlı Grup t-Testi ile karşılaştırılmışlardır. Öntest ve sontestler birbirinden farklı olduğu durumlarda, Bağımlı Grup t-Testi kullanmamız hatalı olur.
- d. Kontrol gruplu deneysel araştırmalarda sıklıkla karşılaşılan başka bir hata; her iki gruba da konuyla ilgili bir öntest yapılmış, daha sonra den grubuna alternatif bir öğretim metodu kullanılarak, kontrol grubuna ise klasik metod kullanılarak 3-4 hafta ders anlatılmıştır. Ders anlatımı sonunda sontest tekrar uygulanmıştır. Daha sonra öğrencilerin öntestleri ve sontestleri Bağımlı Grup t-Testi ile karşılaştırılarak, sontest lehine fark tespit edilmiştir. Deney gruplarında, farkın kaynağı alternatif öğretim metoduna bağlanmıştır. Bu durumda, öğretim metodundan daha etkili olan bağımsız değişken 3-4 hafta boyunca anlatılan derstir. Çünkü konu hakkındaki bilgileri sınırlı olan öğrenciler, 3-4 hafta boyunca konuyla ilgili ders almışlar ve doğal olarak sontest başarıları da öntestlerine göre anlamlı olarak yükselmiştir. Bu durumda ise farkın asıl kaynağı kullanılan öğretim metodu değil, verilen eğitimidir.
- e. Homojenlik (Homogeneity) testi, bağımlı değişkene ilişkin varyansların her örneklem için eşit olduğu hipotezini test eder (Büyüköztürk, 2002, s.44). Tek Yönlü Varyans Analizinin (ANOVA) kullanıldığı durumlarda, homojenlik testi sonuçlarına göre kullanılan çoklu karşılaştırmalar da (Post Hotc Tests) değişmektedir. İncelenilen tez çalışmalarının birçoğunda ANOVA kullanılmış, fakat hiçbirisinde homojenlik testi uygulanmamıştır. İncelenen tezlerdeki çoklu karşılaştırmaların tamamı LSD, Tukey veya Scheffe testleri ile gerçekleştirilmiştir. Halbuki, varyansların homojen olmadığı durumlarda Dunnett's C veya Tamhane's T2 gibi testlerin kullanılması gerekirdi. Aynı şekilde t testlerinin hiçbirisinde, varyansların homojenliğini test etmek amacıyla Levene testi kullanılmamıştır.
- f. ANOVA testinin kullanıldığı bazı tezlerde, öncelikle F testi kullanılmış, daha sonra ANOVA tekniği kullanılmış. Dolayısı ile istatistik değerlerinin hemen hemen aynı olduğu 2 tablo alt alta sıralanmıştır. Aslında ANOVA'da da elde edilen değer F değeridir, dolayısı ile ayrıca F analizi yapmak gereksizdir.
- g. Bazı ANOVA testlerinde anlamlı farklılığa ulaşıldığında, farkın kaynağı çoklu karşılaştırmalar yerine, Bağımsız Grup t-Testi ile sınanmıştır. Bazı tezlerde ise ANOVA testi yapılmış ve anlamlı farklılıklar tespit edildiği zaman farkın kaynağını bulabilmek amacıyla ayrı ayrı LSD testleri gerçekleştirilmiştir. Dolayısı ile farkın kaynağını belirlemek için tek tablo yeterli iken, 6-7 farklı LSD tablosu oluşturulmuştur.

- h. Etki büyüklüğü analizi, bağımsız değişkenin bağımlı değişken üzerindeki etki büyüklüğünü belirlemek amacıyla kullanılır (Rosnow ve Rosenthal, 1996, s. 331). Son yıllarda, yaygın olarak kullanılmaya başlayan etki büyüklüğü (η^2 / Effect Size) analizinden hiçbir tezde yararlanılmamıştır.
- i. Gerek ANOVA, gerekse t-Testlerinde hücre başına düşen denek sayısının çok düşük olması durumunda normallik varsayımı reddedilir. Bu durumda ise, parametrik değil, non-parametrik testlerin kullanılması gerekir (Orhunbilge, 1997, s.247; Levin ve Rubin, 1998, s.596; Akgül, 2005, s.164). İncelenen birçok tezde, bu kural göz önünde bulundurulmamış ve bazı kategorilere 2-3 kişi düşmesine rağmen parametrik istatistiğe devam edilmiştir. Hatta bazı tezlerde, kategoriler boş kalmasına rağmen istatistiksel işlemler gerçekleştirilmiştir.
- j. Bağımlı Grup t testlerinin kullanıldığı tezlerin tamamında “r” değeri ya görmezden gelinmiş, ya da yorumlanmamıştır. Bazı öntest-sontest karşılaştırmalarında korelasyonun yorumlanması, farkın yorumlanması kadar önem arz etmektedir. Dolayısı ile “r” değerinin sunulduğu çalışmalarda, bu bulguya yönelik yorum getirilmelidir.
- k. Bazı tezlerde, deney grubunun öntestleri, kontrol grubunun öntestinden anlamlı şekilde yüksek olmasına rağmen uygulamaya geçilmiştir. Uygulama sonucunda yapılan sontestlerde anlamlı fark oluştuğunda, farkın kaynağı öğretim metoduna bağlanmıştır.
- l. Birçok tezde, regresyon analizleri yorumlanırken, sadece regresyon denkleminde odaklanılmıştır. Halbuki regresyon analizlerinin önemli amaçlarından biri de bağımlı değişkenleri anlamlı olarak yordayan bağımsız değişkenleri tespit etmektir (Bluman, 1995, s.402; Levin ve Rubin, 1998, s.538). Bu durum birçok tez çalışmasında ihmal edilmiştir. Bununla birlikte, bazı tezlerde regresyon analizleri yapılmış fakat bağımlı ve bağımsız değişkenler arasındaki korelasyonlar incelenmemiştir.
- m. Kay-kare testinin uygulanabilmesi için çapraz tablodaki hücrelerde beklenen değerlerin sıfır olmaması ve beklenen değeri 5’in altında olan hücre sayısının, toplam hücre sayısının % 20 sini geçmemesi gerekir (Bluman, 1995, s.429; Orhunbilge, 1997, s.251). Bu sağlanmıyorsa çözüm olarak birbirine yakın kategorilerin, anlam bozmayacak şekilde birleştirilmesi yoluna gidilir (Büyüköztürk, 2002, s.142). Araştırma kapsamında incelenen tezlerin çoğunluğunda bu durum dikkate alınmamıştır.

İfade Hataları ve Biçimsel Hatalar

Tezlerde karşılaşılan bazı ifade hataları aşağıda özetlenmiştir;

- a. ANOVA tekniği kullanılan bazı tezlerde, kareler ortalaması (mean square) kavramı, ortalamaların karesi şeklinde ifade edilmiştir. Mann-Whitney U testindeki sıralar ortalaması (mean rank) kavramı, medyan olarak yazılmıştır.
- b. Madde analizi işlemlerinde ise İngilizce ifadeler tercih edilmiştir. Örneğin; madde toplam item-total, madde kalan ise item-reminder şeklinde ifade edilmiştir.
- c. Bazı tezlerde ise Bağımsız Grup t-Testi, “İlişkisiz t Testi” şeklinde ifade edilerek grup kavramı kullanılmamıştır.

- d. İstatistik analizi sonucunda fark/ilişki oluşmadığı durumlarda, “p” değeri ya hiç verilmemiş veya çizgi ile gösterilmiştir. Aslında, “p” değerleri istatistiksel işlemlerin bir sonucudur ve ne olursa olsun mutlaka sunulmalıdır. Bazı tezlerde, “p” değeri anlamlı fark olmaması dolayısı ile (---) şeklinde ifade edilmiş, fakat “t” değerleri incelendiğinde, 6.81 gibi yüksek değerler gözlenmiştir. Serbestlik derecesi dikkate alındığında, bu kadar yüksek bir “t” değeri için fark oluşmaması ihtimal dışıdır.

Tartışma

Bu araştırmada, fen bilimleri ve matematik eğitimi alanlarında yapılan tez çalışmalarındaki istatistiksel hatalar incelenmiş ve birçok hata tespit edilmiştir. Bu durum, yetersiz istatistik eğitiminden kaynaklanmaktadır (Erkuş, 2004, s.180; Ercan ve diğerleri, 2004, s.1). Birçok bölümde istatistik analizler, kişilerin kendi gayretleriyle ve tanıdıkları öğretim elemanlarının yardımıyla gerçekleşmektedir. Bu ise bazı istatistik hatalarının diğer öğretim elemanlarına miras olarak kalmasına neden olmaktadır. Örneğin, incelenen tezlerde de görülmüştür ki, bazı bölümler normal dağılım testlerinin kullanımında çok hassas davranmalarına karşın, geçerlik-güvenirlik analizlerini ihmal etmişler; bazı bölümler ise geçerlik-güvenirlik çalışmalarını önemserken, non-parametrik teknikleri ihmal etmişlerdir. Buradan da anlaşılabileceği üzere, tez çalışmalarındaki istatistik analizler ilgili bölümün öğretim elemanlarının istatistik birikimleriyle sınırlıdır. Bu durum, tez çalışmalarında önemli istatistiksel hatalara neden olmaktadır ve bu hatayı telafi edecek bir mekanizma bulunmamaktadır. İstatistik analizlerinin doğru kullanılması ve tezlerin hatalı kullanımdan arındırılması için; üniversite bünyesinde veya üniversiteler arasında ilgili uzmanlık birimlerin oluşturulması gerekmektedir. Bu birim, istatistik konusunda öğretim elemanlarına danışmanlık verebileceği gibi, istatistik alanında uzman kişilerin yetiştirilmesine de katkıda bulunabilir.

Kaynakça

- Akgül, A. (2005). Tıbbi Araştırmalarda İstatistiksel Analiz Teknikleri: SPSS Uygulamaları, Ankara: Emek Ofset.
- Bluman, A. G. (1995). Elementary Statistic: A Step by Step Approach, US: William C. Brown Publisher.
- Büyükoztürk, Ş. (2002). Sosyal Bilimler İçin Veri Analizi El Kitabı, Ankara: Pegem A Yayıncılık.
- Ercan, İ., Yazıcı, B., Ediz, B., Özkaya, G., ve Kan, İ. (2004). Bilimsel Araştırmalarda Yapılan İstatistiksel Hatalar, VII. Ulusal Biyoistatistik Kongresi, Mersin.
- Ergün, M. (1995). Bilimsel Araştırmalarda Bilgisayarla İstatistik Uygulamaları: SPSS for Windows, Ankara: Ocak Yayınları.
- Erkuş A.(2004). Bazı Tıp Dergilerinin Son Sayılarındaki Makalelerin Yöntemsel ve İstatistiksel Açından İncelenmesi, Mersin Üniversitesi Tıp Fakültesi Dergisi, Cilt: 5, Sayı: 2 (176-181).
- Gronlund, N.E. ve Linn, R.L. (1990). Measurement and Evaluation in Teaching, Sixth Edition, New York: Mac Millian Publishing.
- Karasar, N. (2003). Bilimsel Araştırma Yöntemi, Ankara: Nobel Yayın Dağıtım.
- Levin, R.I. ve Rubin, D.S. (1998). Statistics for Management, Seventh Edition, New Jersey: Englewood Cliffs.
- Orhunbilge N. (1997). Örnekleme Yöntemleri ve Hipotez Testleri, İstanbul: İ.Ü.İşletme Fakültesi Yayını.
- Rosnow, R. L. ve Rosenthal, R. (1996). Computing Contrasts, Effect Sizes, and Counternulls on Other People's Published Data: General Procedures for Research Consumers, Psychological Methods, Cilt :1, Sayı:1.
- Tavşancıl, E.T. (2002). Tutumların Ölçülmesi ve SPSS ile Veri Analizi, Ankara: Nobel Yayınları.
- Yıldırım, A. ve Şimşek, H. (2003). Sosyal Bilimlerde Nitel Araştırma Yöntemleri, Ankara: Seçkin Yayınları.
- Yüzer, A.F., Ağaoğlu, E., Tatlıgil, H., Özmen, A. ve Şıklar, E. (2003). İstatistik, Eskişehir: Anadolu Üniversitesi Yayınları.