

ANKET, GÜVENİLİRLİK -GEÇERLİLİK ANALİZİ



PROF.DR.YÜKSEL TERZİ

ONDOKUZ MAYIS ÜNİVERSİTESİ
FEN-EDEBİYAT FAKÜLTESİ
İSTATİSTİK BÖLÜMÜ
SAMSUN
2019

Anket

Anket, sistematik bir veri toplama yöntemidir. Veriler, önceden belirlenmiş insanlara bir dizi soru sorarak elde edilir. Anket yöntemi ile çok farklı türde veri toplamak mümkündür. Anketten yararlı bilgiler elde edebilmek için anketin hazırlık ve uygulama sürecine özel önem gösterilmelidir. Ayrıca “zaman” ve “maliyet” göz önüne alınması gereken diğer hususlardır.

Ne Zaman Anket Kullanmalıyım?

Anket, en popüler veri toplama yöntemlerinden biri olmasına karşın, ancak uygun şartlar altında kullanılması halinde yarar sağlayacaktır. Anketi, diğer yöntemlere göre daha hızlı, daha kolay veya daha ucuz olduğu zamanlarda kullanın. Bazen diğer veri toplama metotlarının kullanımı daha uygun olabilir. Örneğin, bir kliniğe giden kişi sayısını belirlemek için, basitçe danışmadaki hasta imzalarını ya da günlük kayıtları inceleyebilirsiniz.

Anket nasıl uygulanacak?

Anket uygulamanın çeşitli yolları vardır. Bunlardan en yaygın olanları yüz yüze görüşmeler, telefon anketleri ve posta anketleridir. Ayrıca bir de karma anket uygulamaları olabilir. Bunlar içerisinde istediğiniz verileri en verimli şekilde sağlayacak metodun seçilmesi önemlidir. Yüz yüze görüşme yöntemi kullanılarak çok detaylı veriler elde edilebilir. Ancak bu yöntem özel bir eğitim gerekir; bu da ekstra zaman ve maliyet demektir.

Yazılı anketler daha az detaylı veri sağlamalarına rağmen, kısa zamanda büyük miktarlarda veri toplanmak için uygundur.

POSTA ANKETLERİ

Hazırlanan soru formlarının posta yoluyla cevaplayıcılara ulaştırılmasını öngören bir veri toplama yöntemidir. Cevaplayıcılardan soru formları doldurulup, posta yoluyla araştırmacıya ulaştırılması istenir. Bu yöntem bilimsel araştırmalarda sık kullanılır.

Avantajları :

- ✓ Anketörlere ihtiyaç duyulmaz ve ulaşım, telefon gibi maliyetler olmayacağından az maliyetli bir yöntemdir.
- ✓ Gelir, politik gibi sorulara anketör olmadığı için daha sağlıklı cevap alabilmek mümkündür.
- ✓ Zaman problemi yoktur.

Dezavantajları :

- ✓ Uygulamada ortaya çıkan problemlerin çözümü mümkün değildir.
- ✓ Çok uzun süre alan anket yöntemidir.
- ✓ Sorulara tam olarak kimin cevap verip vermediği belli olmaz.
- ✓ Soruların tamamına cevap verilmeyebilir.
- ✓ Açık uçlu sorular için uygun değildir.

TELEFON ANKETLERİ

Hazırlanan soru formları eğitimli anketörlerin cevaplayıcılarla yaptığı telefon görüşmeleri ile toplanır.

Avantajları :

- ✓ Sonuçlar çok hızlı bir şekilde alınır.
- ✓ Anketör tüm sorulara cevap almak için teşvik edebilir.
- ✓ Şehir içi anket çalışmalarında maliyet uygun olabilir.

Dezavantajları

- ✓ Görsel yardımcı malzemeler bu anket türünde kullanılmaz.
- ✓ Şehirlerarası görüşmelerin olacağı anketlerde ciddi bir maliyet ortaya çıkar.
- ✓ Telefon olmayan veya ulaşılamayan kişiler anket kapsamı dışında kalacaktır. Ayrıca telefon rehberinde kayıt olmama durumları da bir sıkıntıdır.

KARŞILIKLI /YÜZ YÜZE) GÖRÜŞME

Hazırlanan soru formları eğitimli anketörlerin cevaplayıcılarla yaptığı doğrudan görüşmeleri ile toplanır.

Avantajları :

- ✓ Görsel yardımcı malzemeler kullanılarak sorular daha kolay ve anlaşılır hale getirilebilir.
- ✓ Anketör cevaplayıcının anlayamadığı noktaları açıklayabilir.
- ✓ Anketin önemi anlatılarak ve gizlilik konusunda güvence verilerek, cevaplanma oranı yükseltilebilir.
- ✓ Bazı durumlarda cevaplayıcılara ulaşmak diğer yöntemlere göre daha kolaydır. Sinema önünde film ile ilgili görüş almak gibi.
- ✓ Anketin doğru kişi üzerinde yapıldığı tespit edilir.
- ✓ Eksik soru varsa tamamlanması istenir.
- ✓ Evsiz insanlar gibi özel gruplardan bilgi toplanabilir.

Dezavantajları

- ✓ Anketörler topladıkları bilgilerin taraflı olmasına yol açabilir.
- ✓ Personel ve ulaşım maliyeti yönünden harcamalar yüksek olabilir.
- ✓ Cevaplayıcı üzerinde sosyal etki en üst düzeyde olan anket yöntemi türüdür.
- ✓ Sonuçların alınması uzun zaman alabilir.
- ✓ Maliyeti kısma eğilimi hataya yol açar. Örneklem büyüklüğünü küçültmek, araştırma dışındaki kişilerle görüşmek, tecrübesiz anketörler ölçüm hatasına yol açar.
- ✓ Cevaplayıcıyı evde bulmak çoğu zaman güç olabilir.

Anket Maddelerinin Oluřturulması

İlgilenilen her bir konu için anket soruları yazın ve aklınıza gelen her řeyi ama her řeyi not edin. Bařlangıçta daha fazla madde olması daha az madde olmasından iyidir. Mükemmel soru hazırlamaya odaklanmayın. İlk bařta, tüm konulara deęinmek süslü cümlelerle anlatım yapmaktan daha önemlidir.

Anket maddelerini oluşturmak için, beyin fırtınası, ağaç diyagramları ve dięer fikir üreten metotlar yararlı olacaktır.

Bir anket maddesinin taşınması gereken özellikler:

açıkça yazılmalı

İfadeler net olmalı ve kolay anlaşılabilir olmalıdır. Anketi cevaplayan insanların alışık olmadıkları teknik terim, kısaltmalardan kaçınınız.

kısa

Konuya mümkün olduğunca çabuk geliniz. Anketin maddelerini cevaplayan kişinin dikkati dağılmaması ve ankete konsantre olabilmesi için konu dışı kelimeleri ve fikirleri ortadan kaldırınız.

belirli

Her seferinde bir fikir üzerinde odaklanınız. Her madde bir tek davranış, tavır, fikir, olay ya da konu ile ilgili olmalıdır. Bunun için “ve”, “ya da” gibi bağlaçları kullanmayınız.

açık

İnsanları ne sorulduğunu tahmin etmeye zorlamayınız. İfadenizin herkes tarafından aynı şekilde anlaşıldığından emin olun. Eğer gerekliyse ek açıklamalar yapınız.

ANKET FORMU TASARIMI

Anket formu hazırlanırken soruların hedeflenen bilgiye ulaştırarak şekilde hazırlanmasına dikkat edilmelidir. Sorulara başlamadan önce araştırma hakkında kısa bilgi verilmelidir. Daha sonra soruların nasıl cevaplanacağı açık olarak belirtilmelidir. Sorular rahat okunabilecek biçimde olmalıdır ve anket sonunda teşekkür edilmelidir.

Sorulara cevaplama kolay olanlardan başlanmalıdır. Düşünme gerektiren açık uçlu sorular anketin başında sorulmamalıdır. Ayrıca gelir, özel yaşamı gerektiren sorular anketin sonunda sorulmalıdır.

SORU TÜRLERİ

1. KAPALI UÇLU SORULAR

Anket formlarındaki soruların büyük bir kısmı kapalı uçludur. Bu sorular iki seçenekli (evet/hayır), çok seçenekli ve **likert** tipi olabilir.

S1. Amirim çalışanların önemli kararlara katılmalarını teşvik eder.

- | | |
|----------------------------|---|
| A. KESİNLİKLE KATILMIYORUM | 1 |
| B. KATILMIYORUM | 2 |
| C. KARARSIZIM | 3 |
| D. KATILIYORUM | 4 |
| E. KESİNLİKLE KATILIYORUM | 5 |

S2. Lütfen size sağlanan hizmetin kalitesini değerlendirin.

- | | |
|-------------|---|
| A. ÇOK KÖTÜ | 1 |
| B. KÖTÜ | 2 |
| C. ORTA | 3 |
| D. İYİ | 4 |
| E. ÇOK İYİ | 5 |

KAPALI UÇLU SORULARIN AVANTAJLARI

- Cevaplamaları kolaydır.
- Analiz öncesi çalışmalarda tasnif edilmeleri, tablolara aktarılmalar kolaydır.
- Verilerin bilgisayara girilmesinde kolaylık sağlar.
- Verilen cevaplar cevaplayıcıya yardımcı olur.

KAPALI UÇLU SORULARIN DEZAVANTAJLARI

- Evet-Hayır ve Bir fikrim yok şeklindeki sorularda, gerçekten fikir olmadığı için mi, yoksa kararsız kalındığı için mi işaretlendiği belirlenemez.
- Cevap seçeneği çok olduğunda, cevaplayıcı zorlanabilir.
- Cevaplayıcı soru şıklarında kendisine uygun cevabı bulamadığında “Diğer” seçeneği soruya eklenebilir.
- Likert ölçeklerinde kişilerin ilgi duydukları konuda yüksek puan vermeleri ve sevmedikleri konuda düşük puan vermesi bu tekniğin güvenilirliğini etkiler.

2. AÇIK UÇLU SORULAR

Bu tür sorular cevap seçeneği bulunmayan sorulardır. Örneğin A marka hakkında ne düşünüyorsunuz?, yaşıınız kaç? gibi sorular açık uçlu sorulardır.

AÇIK UÇLU SORULARIN AVANTAJLARI

- Cevaplarla cevaplayıcıyı yönlendirmek mümkün değildir.
- Cevaplayıcının düşüncelerini rahatça yazması sağlanır.
- Hiç düşünülmemiş sonuçların alınması sağlanabilir.

AÇIK UÇLU SORULARIN DEZAVANTAJLARI

- Değerlendirilmesi oldukça güç, zaman ve çaba gerektirir.
- İlgisiz ve tekrar niteliğinde cevaplar olabilir.

SORULARIN SIRASI

Giriş Soruları :

- ✓ Açık uçlu soru,
- ✓ Cevaplanması güç sorular,
- ✓ Endişe edici sorular

sorulmaz. Hassas sorular bazı sorular sorulduktan sonra sorulmalıdır.

Ayırma Soruları :

Belirli soruların kimler tarafından cevaplanacağını belirlemek için kullanılır.

Soru. Bilgisayar eğitimi aldınız mı?

1. Evet
2. Hayır (Cevabınız Hayır ise 7. sorudan devam ediniz)

Demografik sorular

Demografik sorular genelde katılımcıları sınıf, yaş ya da eğitim seviyesi gibi kriterlere dayalı olarak daha küçük özellikli gruplara ayırmak için kullanılır. Eğer anketin amaçlarından birisi iki farklı grubun verdiği cevaplar arasındaki önemli farklılıkları belirlemek ise bu ayrıştırma önemlidir.

Demografik soruların cevaplayıcıların kimliğini ortaya çıkartacak olması onların kendilerini rahatsız hissetmelerine ve dürüst cevap vermemelerine yol açabilir. Olası endişeleri azaltmak için, anket cevaplarının nasıl kullanılacağını anlatan bir açıklama ve cevapların gizliliğini içeren bir taahhüt yararlı olabilir.

Ait olduđunuz yař kategorisini iřaretleyiniz.

- 20'NİN ALTINDA
- 20-29
- 30-39
- 40-49
- 50-59
- 59'UN ÜSTÜ

Eđitim durumunuz

- CAHİL
- İLK OKUL
- ORTA OKUL
- LİSE
- ÜNİVERSİTE
- YÜKSEK LİSANS
- DOKTORA

Medeni haliniz

- BEKAR (HİÇ EVLENMEMİŐ)
- EVLİ
- DUL

İYİ SORU YAZMADA DİKKAT EDİLECEK TEMEL KURALLAR

Bir soru sorarken tanımlanmak istenen şeyle ilgili anlamlı bilgiler içeren cevaplar üretmelidir (Geçerlilik).

Ölçüm sürecinin tekrarlanması halinde benzer sonuçlar elde edilmesi önemlidir (Güvenilirlik).

İyi soruların hazırlanması çok zaman almakla birlikte, bu sorulardan elde edilen kazanç büyük olacaktır. İnsanlar iyi tasarlanmış ilginç soruları istekle cevaplayacağından cevap hatası, doğru bilgi vereceğinden de ölçüm hataları azalır.

✓ **Soruların amacınızı karşıladığından emin olun:**

- Kaç yaşındasınız?
- Hangi tarihte doğdunuz?

Burada ilk soruda yalnızca yıl sorulurken, ikinci soruda ise gün/ay/yıl sorulmaktadır. İstenen amaca göre soru sorulmalıdır.

- Aylık maaşınız nedir?
- İkramiye, ekstralara dahil aylık maaşınız nedir?
- Tüm aile fertlerinizin aylık geliri nedir?

✓ **Bir soruda yalnız bir konu seçin:**

Aşağıdakilerden hangisi aile doktorunuzu en iyi biçimde tanımlamaktadır?

- Çok bilgili ve arkadaşça
- Çok bilgili ve resmi
- Bilgili ve arkadaşça
- Bilgili ve resmi
- Bilgisiz ve arkadaşça
- Bilgisiz ve resmi

- ✓ Soruların tüm cevaplayıcılar tarafından aynı şekilde anlaşıldığından emin olun;

Sağlık durumunuzu değerlendirin?

- İyi
 - Orta
 - Kötü
- veya
- Çok iyi
 - İyi
 - Orta
 - Kötü
 - Çok kötü

- ✓ Cevaplayıcıların bilgi düzeyini zorlamayın;
- ✓ Aynı soruda birden fazla olumsuz ifade kullanmayın;
- ✓ Yönlendirici ifadeler kullanmayın;
- ✓ Karmaşık birden fazla anlama gelen sorular sormayın.

Özet olarak bir anketi hazırlarken aşağıdakileri aklımızdan çıkarmamalıyız

Anket yapmak için anlaşılır bir amacınız olsun.

Anketinizi, araştırılan konu ve hedef kitlenizi göz önüne alarak hazırlayın.

Anketlerin kısa ve basit olmasına özen gösterin.

Sonuçları yöneticilere, müdürlere ve cevaplayıcılara duyurun.

Anket Sonularının Sunulması

Anket verilerini toplayıp, analizini yaptıktan sonra yapılması gereken en önemli faaliyet sonuçların sunulmasıdır. Anket sonuçlarının bir özeti, cevaplayıcılara mümkün olduğu durumlarda belirlenen problemlerin ve endişelerin ele alındığı bir tartışma ortamında sunulmalıdır.

- ✓ Başlık sayfası
- ✓ Özet
- ✓ Giriş
- ✓ Tasarım
- ✓ Örneklem
- ✓ Güvenilirlik
- ✓ Analiz
- ✓ Sonuçlar
- ✓ Yorum
- ✓ Tavsiyeler
- ✓ Kaynaklar
- ✓ Ekler

Güvenirlilik

Bir ölçme aracında (testte veya ankette) bütün soruların birbirleriyle tutarlılığını, ele alınan oluşumu ölçmede türdeşliğini ortaya koyan bir kavramdır.



- Ölçme araçlarının güvenilirliğini değerlendirmek amacıyla geliştirilmiş yöntemlere Güvenirlik Analizi ve bu araçta yer alan soruların irdelenmesine ise soru analizi (Reliability and Item Analysis) denilmektedir.



- Testlerin güvenilirliğini analiz etmek amacıyla güvenilirlik katsayıları hesaplanmaktadır. Bugüne kadar güvenilirlik analizi için bir çok yöntem geliştirilmiştir.



Güvenilirlik (reliability) ve geçerlilik (validity) analizleri yapılmadan bir arařtırmanın analiz sonuçlarını yorumlamak doęru deęildir. Güvenilirlik ve geçerlilik herhangi bir řeyin uygun ve saęlam olduęu hakkında bize bilgi verir. Örneęin ayakkabının ayaęımıza hem uygun olması hem ayakkabının saęlam olması gerekir. Burada uygunluk geçerlilięi, saęlamlık ise güvenilirlięi temsil etmektedir.

Geçerlilik, bir testin ölçmek istediği özelliği ölçmesi ile ilgilidir. Bir test ölçmek istediği özelliği doğru ve diğer özelliklerle karıştırmadan ölçüyor ise bu testin geçerli olduğu söylenir. Geçerli bir testte bulunması gereken en önemli özellik testin güvenilir olmasıdır.

Güvenilirlik bir ölçüm sürecinde ölçüm işleminin tekrarlanabilirliği ya da tekrarlardaki tutarlılıktır.

- Güvenilirlik bir özelliğin bağımsız ölçümleri arasındaki kararlılıktır.**
- Güvenilirlik zamana göre değişmezlik ölçüsüdür.**

Güvenilirlik bir testin geçerliliğini etkiler. Geçerli bir test muhakkak güvenilir olmalıdır. **Güvenilir bir test geçerli olmayabilir. Yani güvenilirlik geçerliliği garanti etmemektedir.**

Örneđin kol kuvvetini ölçmek için kullanılan barfix testinin güvenilirliđi çok yüksektir. Bu test yanlışlıkla bacak kuvvetini ölçmede kullanılırsa geçerli olmayacaktır. Yine 20-45 yaş arasındaki bayanlar için geliştirilen ve belli bir özelliđi ölçen güvenilirliđi yüksek bir ölçeđin 55 yaş üzerindekiilere uygulanırsa geçerli olmayacaktır.

Araştırma çalışmaları yapılırken nesnelere deęişik özellikleri ile ilgili ölçüm yapmak için deęişik aletler, laboratuvar aygıtları kullanılır veya sosyal bilimler, davranış bilimleri gibi bilim dallarında kullanılan anket çalışmaları veya psikolojik testlerde de farklı ölçme aygıtları kullanılır.

Bir araştırma sonucu elde edilen verilere sadece verinin iyi toplandığı için güvenilmez, yöntem kısmında araştırmada kullanılan aletlerin detay tanımları verilir, kullanım amaçları açıklanır, soru biçimleri ve ölçme ölçeęi (adlandırma sıralama, aralık veya oran) açıklanır. Bunların hepsi geçerlilik ve güvenilirlięin sağlanabilmesi amacına yönelik gayretlerdir.

Bir arařtırmada geerlilik ve gvenirlilięi artırmak iin ařaęıdaki dzenlemelerin yapılmasına dikkat edilir (Thomas RK, Mueller RO.)

Sıra no	Yapılması gerekenler	Hangi Blmde Yapılır
1	alıřmanın hedef ve hipotezleri yalın ve sade bir Őekilde tanımlanmalıdır.	Giriř, Yntem
2	alıřmada kullanılan her aygıt yeterli detayda tanımlanmalıdır. Kullanılan aygıtların alıřmanın amalarına uygunluęu aıklanmalıdır. Uygun geerlilik delilleri toplanmalıdır. alıřmanın amacına uygun mevcut toplanan veri setinin destekledięi geerlilięin eřidi belirlenmelidir.	Yntem ve sonular
3	Gvenirlik ve geerlilięi artırıcı alınan nlemler aıklanmalıdır. nceki alıřmalardaki gvenirlik indeks deęerleri ile bu alıřmanın amalarına uygun Őekilde hesaplanan indeksler deęerler karřılařtırılmalıdır.	Yntem ve sonular
4	Arařtırmada uygulanan gvenirlik ve geerlilik alıřmaları yorumlanmalı ve rapor edilmelidir. alıřmanın sonuları bu mevcut deliller kapsamında tartiřılmalıdır.	Tartıřma

Geçerlilik ve güvenilirlik ölçme aygıtlarının önemli iki özelliğidir.

Güvenirlik ölçümün kararlılığı ile ilgili bir özelliktir. Alınan ölçümlerin kararlı olması veya değişmemesi istenir, bu kararlılık zamandan zamana, mekandan mekana, sorudan soruya, hakemden hakeme değişmeyen bir özelliktir.

Geçerlilik ise ölçmek istenen özelliğin, ölçülmesi niyet edilen hedef değer gerçekten ölçülebilmesi ile ilgili bir özelliktir. Dolayısıyla geçerlilik çalışmanın amacı ve araştırma sorusu (hipotez) ile yakından ilişkili bir husustur.

Araştırmacılar ölçme aracının farklı kültürler, farklı durumlar, farklı koşullarda aynı hassasiyetle ölçüm yapıp yapmadığı ile yakından ilgilidir. Örneğin Asya ırkından çocuklarda oldukça yüksek güvenilirlik ve geçerliliğe sahip bir psikolojik test, Afrika ırkından çocuklarda olmayabilir.

Güvenirlik ve geçerlilik kavramları ve tanımlamaları zaman içerisinde sürekli gelişmekte bunları ifade edebilmek için farklı ölçütler bulunmaktadır.

Geçerlilik mi, Güvenirlilik mi?

Geçerlilik mi daha önemli güvenirlilik mi daha önemli sorusunun yanıtını vermek biraz zordur.

Geçerli olmayan bir test ne kadar güvenilir olursa olsun bir işe yaramaz. Geçerlilik ön plana çıkmaktadır. En azından yapı ve içerik geçerliliği mutlaka sağlanmalıdır. Bir bakıma güvenirlilikte geçerliliğin bir parçası şeklinde düşünmek kavramların ne kadar iç içe olduğunu vurgulamak için yerinde olur. Burada ayrı ayrı konu edilmesinin amacı aralarındaki farklılıkları vurgulamaktır.

Geçerlilik ve Güvenirlik İlişkisi

Güvenirlik, geçerlik için bir ön koşuldur. Bir ölçme aracının güvenilir olması her zaman geçerli olduğu anlamına gelmez. Geçerlilik, güvenilirliği kapsayan bir kavramdır.

- Geçerlilik ile güvenilirlik arasındaki gerçek fark bir tanım meselesidir.
- **Güvenirlik yapılan ölçümün kararlılığını (consistency) veya daha basit olarak aynı şartlarda, aynı nesnelere, değişik zamanlarda veya kişilerce ölçüldüğünde benzer sonuçların bulunması hadisesidir.**
- **Geçerlilik ise ölçmek istenen, amaçlanan şeyin gerçekten ölçülüp ölçülmediği ile ilgilidir yani kesinlik (accuracy) ölçüsüdür.**

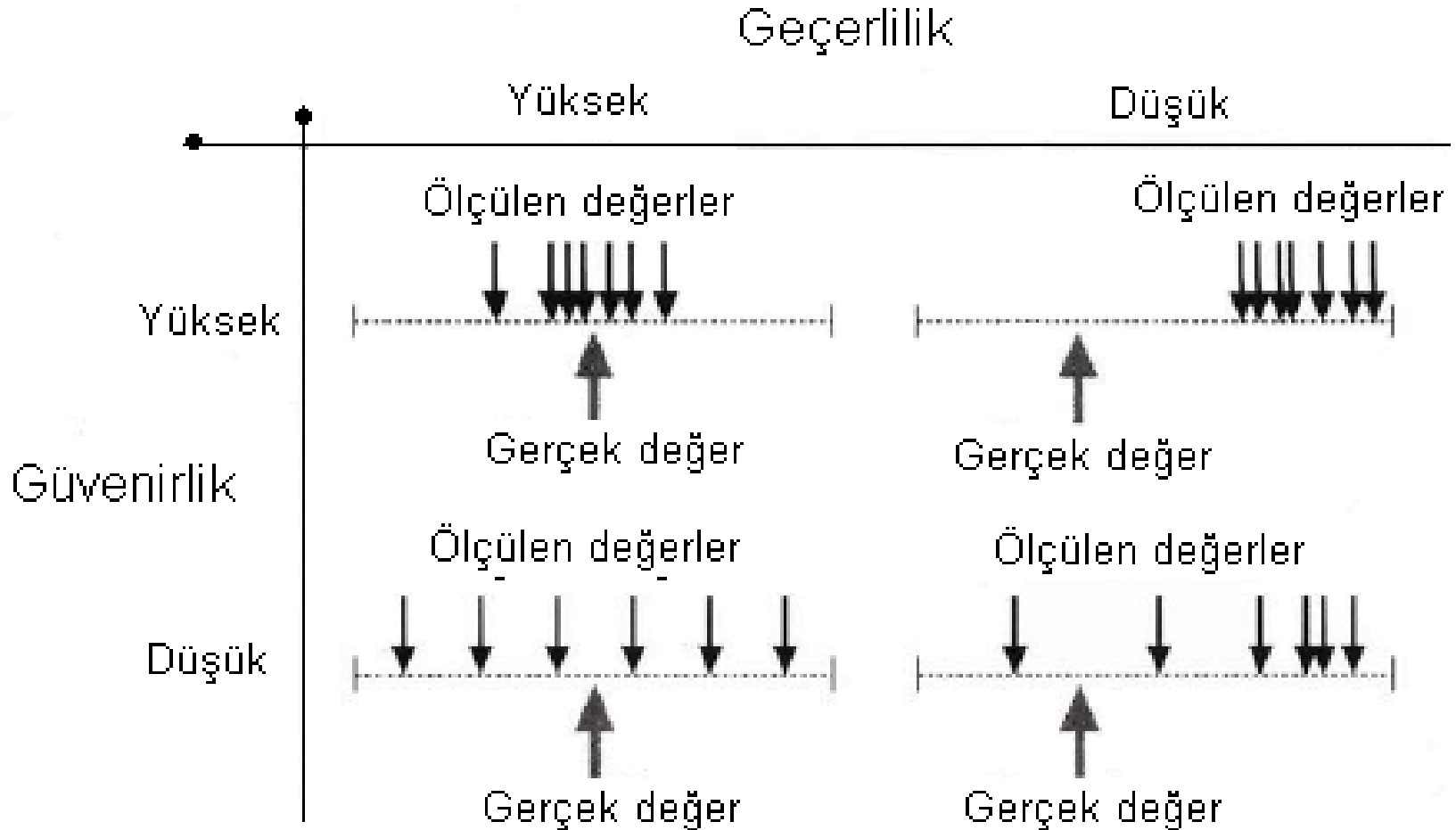
- Amaçlanan hedef için ölçüm yapılmıyorsa ne kadar hassas ölçüm yapılması fazlada bir önem taşımaz, o nedenle geçerlilik daha ön planda düşünülmesi gereken konudur.
- Geçerlilik her türlü hatadan etkilenmesine rağmen, güvenilirlik yalnızca tesadüfî hatadan etkilenir. Bir testin geçerliliği güvenilirlikten etkilenir, bu yüzden testin geçerlilik katsayısı, güvenilirlik katsayısının karekökünden daha büyük olamaz.

Bir testte hata kaynakları kontrol altına alındığı zaman testin güvenilirliği artar. Güvenirliliği artan bir testin geçerliliği de yükselir, ancak bu iki kavram birbirlerinden farklı kavramlardır.

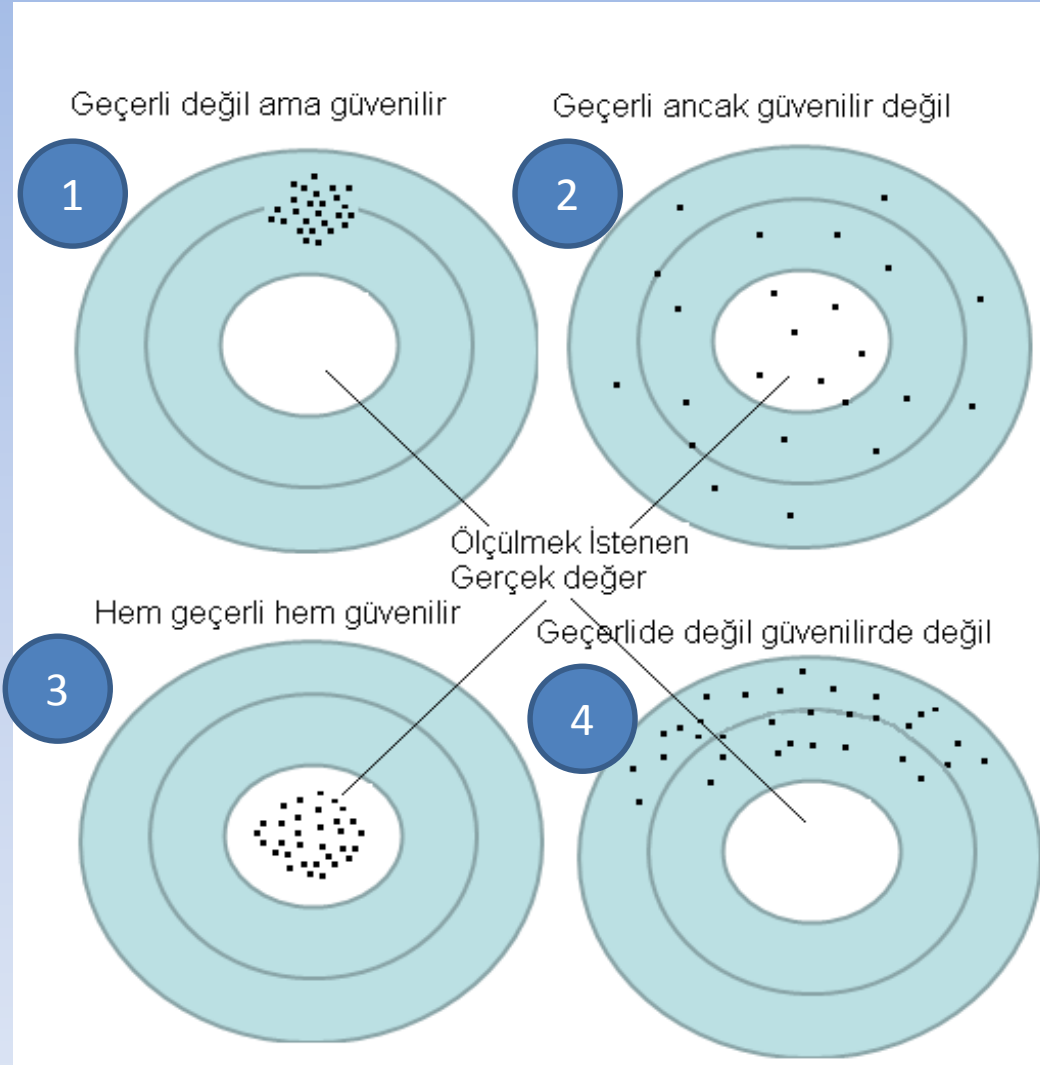
Güvenirlik bir özellik birkaç defa aynı araçla ölçüldüğünde her seferinde benzer sonuçlar elde edilmesi ile ilişkili bir kavramdır.

Geçerlilik ve Güvenirlilik İlişki Şeması

Geçerlilik ve Güvenirlilik



Kavramların şekille ifadesi



1. Her zaman aynı sonucu alıyor, ama istenen sonucu tam kapsamıyor.
2. İstenen sonucun tümünü kapsıyor, ama her zaman farklı sonuç alabiliyor. (Tesadüfî hata yüksek)
3. Her zaman aynı sonucu alıyor ve isteneni tam ve kesin kapsıyor. (Tesadüfî hata düşük)
4. Hem istenen sonucu tam kapsamıyor, hem de her zaman aynı sonucu alamıyor.

GÜVENİLİRLİK (Reliability)

Güvenilirlik, bir testin veya ölçeğin ölçmek istediği şeyi tutarlı ve istikrarlı bir biçimde ölçme derecesidir. Diğer bir deyişle, bireylerin test maddelerine verdikleri cevaplar arasındaki tutarlılıktır.

Güvenilir bir test veya ölçek, benzer şartlarda tekrar uygulandığında aynı benzer sonuçlar verir. Bir test veya ölçek ne derece güvenilir ise ondan elde edilen veriler de o derece güvenilirdir. Güvenilir olmayan bir ölçek yardımıyla elde edilen veriler faydasızdır. Mesela bir zeka testinde bir öğrenci bir gün 100, diğer gün 140 puan alıyorsa yapılan bu testin güvenilirliğinden bahsedilemez.

Güvenirlik;

- **Aynı şeyin bağımsız ölçümleri arasındaki kararlılıktır;**
- **Ölçülmek istenen belli bir şeyin, sürekli olarak aynı sembolleri almasıdır;**
- **Aynı süreçlerin izlenmesi ve aynı ölçütlerin kullanılması ile aynı sonuçların alınmasıdır;**
- **Ölçmenin, tesadüfî yanılılardan arınık olmasıdır;**
- **Bir ölçüm sürecinde ölçüm işleminin tekrarlanabilirliği, kararlılığı ya da tekrarlardaki tutarlılıktır.**

GÜVENİLİRLİK

- ✓ Endüstri mühendisine göre güvenilirlik: "üretimde hata oranının veya başarısızlık oranının düşük çıkmasıdır".
- ✓ Bir sosyologa göre güvenilirlik: "ölçüm sonuçlarının farklı anakütlelerde veya aynı ana kütlede ait farklı örnek kütlelerde benzer sonuçlar vermesidir".
- ✓ Bir insan kaynakları yöneticisine göre güvenilirlik: "psikometrik test sonuçlarının uygulandığı farklı zaman dilimlerinde benzer sonuçlar vermesi ve iş yaşamında personelin göstereceği performansı doğru tahmin etmesidir".

Pilot araştırma verilerinde yapılan güvenilirlik analizleri bir ön yordama niteliğindedir. Güvenilirlik analizleri ise esas araştırma sonuçlarına göre yapılmalıdır.

Güvenilirliği sorgulama :

- **Bir ölçüm aracındaki maddelerin aynı kavramsal yapıyı hatasız bir biçimde ölçmesi**
- **Farklı zamanlarda yapılan ölçüm sonuçlarının benzer çıkması**
- **Bir ölçüm aracına ait sonuçların aynı kavramsal yapıyı ölçen diğer ölçüm araçlarının sonuçlarıyla tutarlı olması**
- **Farklı gözlemciler tarafından yapılan ölçüm sonuçlarının benzer çıkması**

Örnek

Güvenirlilik: (üç senaryo düşünelim)

1. Ağaç masamın boyutları bir zamandan diğerine değişmeyeceğine göre veya bir yerden diğer yere göre değişmeyeceği için ve benim çelik ölçme aletimin boyu değişmeyeceği için ölçüm nerede veya ne zaman yapılırsa yapılsın fark etmez aynı bulunması gerekir. Eğer masa bir yerde 120 cm ölçülmüş ise başka bir yerde de veya başka bir günde ölçümde de 120 cm gelecektir.

Birbiri peşine 3 ölçüm yapsak her 3 ölçümde de 120 cm bulunacaktır.

2. Diğer bir senaryoya göre ölçüm aletinin esneyebilen bezden yapıldığını ve çekme ile biraz uzayabildiğini varsayalım bu durumda 3 farklı ölçümde değişik ölçümler bulunacaktır, 120 cm, 121 cm ve 120,7 cm gibi.

3. Bir başka senaryoya göre ölçüm aleti sıcaklıktan daha fazla etkilenen daha esnek bir materyalden (plastik) yapılmış olsun. Bu aletle aynı masa oda ısısının farklı olduğu zamanlarda alındığında alet sıcaktan çok etkileneceği için çok daha farklı sonuçlar elde edilecektir. 120 cm, 150 cm ve 80 cm bulunmuş olsun.

Bu kavram ölçümün güvenilirliği ile ilgilidir. **Aynı nesne, aynı işlemle ölçüldüğünde neticenin aynı olması gerekir.** 1. durumda olduğu gibi alınan bir ölçüm **güvenilir** bir ölçümdür. 2. durumda çok az değişiklik vardır, dolayısıyla bu ölçümde **nispeten güvenilir**dir, ama 3. durumda çok büyük farklılıklar olduğu için ölçümler **güvenilirlikten uzaktır.**




Nesne: Masa



Ölçüm aracı: Çelik metre
Mezura (Kot)
Mezura (plastik)

ÖRNEK1. 1. ve 2. arařtıřıcılar, altın standartla (gerçek deęere en yakın ölçüm) benzer ölçümler yapmıřtır. **Deęerlendiriciler arası güvenilirlik** tamdır.

	Altın standart Ölçüm	1. Arařtırıcı	2. Arařtırıcı
			
1. Kiři	60 ml.kg ⁻¹ .min ⁻¹	60 ml.kg ⁻¹ .min ⁻¹	60 ml.kg ⁻¹ .min ⁻¹
2. Kiři	55 ml.kg ⁻¹ .min ⁻¹	55 ml.kg ⁻¹ .min ⁻¹	55 ml.kg ⁻¹ .min ⁻¹
3. Kiři	70 ml.kg ⁻¹ .min ⁻¹	70 ml.kg ⁻¹ .min ⁻¹	70 ml.kg ⁻¹ .min ⁻¹

ÖRNEK2. 1. ve 2. arařtıřıcılar, altın standartla (gerçek deęere en yakın ölçüm) benzer ölçümler yapmıřtır. **Deęerlendiriciler arası güvenilirlik** tamdır. Ancak ölçümler hatalıdır.

	Altın standart Ölçüm	1. Arařtıřıcı	2. Arařtıřıcı
			
1. Kiři	60 ml.kg ⁻¹ .min ⁻¹	65 ml.kg ⁻¹ .min ⁻¹	65 ml.kg ⁻¹ .min ⁻¹
2. Kiři	55 ml.kg ⁻¹ .min ⁻¹	60 ml.kg ⁻¹ .min ⁻¹	60 ml.kg ⁻¹ .min ⁻¹
3. Kiři	70 ml.kg ⁻¹ .min ⁻¹	75 ml.kg ⁻¹ .min ⁻¹	75 ml.kg ⁻¹ .min ⁻¹

ÖRNEK3. 1. ve 2. arařtıřıcılar, altın standartla (gerçek deęere en yakın ölçüm) benzer ölçümler yapmıřtır. **Deęerlendiriciler arası güvenilirlik** yoktur, hem de ölçümler hatalıdır.

	Altın standart Ölçüm	1. Arařtırıcı	2. Arařtırıcı
			
1. Kiři	60 ml.kg ⁻¹ .min ⁻¹	72 ml.kg ⁻¹ .min ⁻¹	57 ml.kg ⁻¹ .min ⁻¹
2. Kiři	55 ml.kg ⁻¹ .min ⁻¹	61 ml.kg ⁻¹ .min ⁻¹	52 ml.kg ⁻¹ .min ⁻¹
3. Kiři	70 ml.kg ⁻¹ .min ⁻¹	40 ml.kg ⁻¹ .min ⁻¹	84 ml.kg ⁻¹ .min ⁻¹

✓ KAYNAKLAR

- ✓ Alpar, R. (2016). Uygulamalı İstatistik ve Geçerlik-Güvenilirlik, Detay Yayıncılık, Ankara.
- ✓ Şencan, H. (2005). Sosyal ve Davranışsal Ölçümlerde Güvenilirlik ve Geçerlilik, Seçkin Yayıncılık, Ankara.
- ✓ Altunışık, R. ve ark. (2005). Sosyal Bilimlerde Araştırma Yöntemleri, Sakarya Kitabevi, Adapazarı.
- ✓ Büyüköztürk, Ş. (2005). Sosyal Bilimler İçin Veri Analizi El Kitabı, Pegema yayıncılık, Ankara.
- ✓ Kalaycı, Ş. Ve ark.(2005). SPSS Uygulamalı Çok Değişkenli İstatistiksel Teknikler, Asil Yayın Dağıtım, Ankara.
- ✓ Özdamar, K. (2002). Paket Programlar ile İstatistiksel veri Analizi 1, Kaan Kitabevi, Eskişehir.
- ✓ Baş, T. (2001). Anket, Seçkin Kitabevi, Ankara.
- ✓ Karagöz, Y. (2014). SPSS 21.1 Uygulamalı Biyoistatistik, Nobel Yayın, Ankara.
- ✓ Çokluk Ö., Şekercioğlu G., Büyüköztürk Ş. Sosyal bilimler için çok değişkenli istatistik SPSS ve Lisrel Uygulamaları, Pegem Akademi, Ankara, 2014.
- ✓ Seçer İ. (2015). SPSS ve Lisrel ile pratik veri analizi- Analiz ve raporlaştırma, Anı yayıncılık, Ankara.
- ✓ Çepni Z. (2010). Yapısal eşitlik modellemesi, Hacettepe Ün., <http://yunus.hacettepe.edu.tr/~cepni/mersinsemsunu.ppt>

- ✓ Atkinson, G. and A. M. (1998). Nevill. Statistical methods for assessing measurement error (Reliability) in variables relevant to sports medicine. *Sports Medicine*. 26:217-238.
- ✓ Baykul, Y. (2000). Eğitimde ve Psikolojide Ölçme: Klasik test Teorisi ve Uygulaması. Ankara: ÖSYM yayınları.
- ✓ Crocker, L. ve Algina, J. (1986). Introduction to Classical and Modern Test Theory. New York: Holt. Rinehart and Winston.
- ✓ Erkuş, A. (2003). Psikometri Üzerine Yazılar. Türk Psikologlar Derneği Yayınları no:24, 1. basım: Ankara.
- ✓ Horst, P.(1966). Psychological measurement and prediction. California: Wadsworth.
- ✓ Ercan, İ., Kan, İ. (2004). Ölçeklerde Güvenilirlik ve Geçerlilik, Uludağ Ün. Tıp Fak. Dergisi, 30(3), 211-216.
- ✓ Karagöz, Y. (2014). SPSS 21.1 Uygulamalı Biyoistatistik, Nobel Yayın, Ankara.
- ✓ Çılan, Ç.A. (2013). Sosyal Bilimlerde Kategorik Verilerle İlişki Analizi, Pegem Akademi, Ankara.
- ✓ Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychol Bull*, 107(2),238-46.
- ✓ Tekin, H. (2000). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınları.Yükseköğretim Kurulu (YÖK) (2012). [ttp://www.yok.gov.tr/content/view/471/](http://www.yok.gov.tr/content/view/471/)
- ✓ Şarlak, Y. (2010). Eğitimde Ölçme ve Değerlendirme, İstanbul.
- ✓ Çalık, M. Test Geliştirme ve Madde Analizi, <https://fatihegitim.files.wordpress.com/2011/02/hafta-14-madde-analizi.ppt>
- ✓ Şimşek, Ö. Ve Ceyhun, Ş.Y. Test ve Madde Analizleri, <http://sevimasiroglu.com/wp-content/uploads/2016/12/TEST-VE-MADDE-ANAL%C4%B0ZLER%C4%B0.pdf>
- ✓ Gömleksiz, M. ve Erkan, S. (2016). Eğitimde Ölçme ve Değerlendirme, Nobel Kitabevi, İstanbul.
- ✓ Erkuş, A. (2014). Psikolojide Ölçme ve Ölçek geliştirme-I, Pegem Akademi, Ankara.

- ✓ Ayyıldız, H., Cengiz, ve Ustasüleyman, T. (2006). Üretim ve Pazarlama Bölüm Çalışanları Arası Davranışsal Değişkenlerin Firma Performansı Üzerine Etkisine İlişkin Yapısal Bir Model Önerisi, Muğla Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, Güz 2006, (17,)17.
- ✓ Bollen, K.A. (1989). Structural equations with latent variables. New York: John Wiley & Sons.
- ✓ Schumacker, R.E. & Lomax, R.G. (1996). A beginner's guide to structural equation modeling. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- ✓ Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive Goodness-Of-Fit Measures Of Psychological Research Online, (8), 2, 23-74
- ✓ Özabacı, N. (2011). İlişki Niteliği Ölçeği'nin Türkçe Uyarlaması: Geçerlik ve Güvenirlik Çalışması, Eğitim ve Bilim , 36 (162), 159-167.
- ✓ Munro B.H. (2005). Statistical Methods For Health Care Research. Philadelphia: Lippincott Williams & Wilkins, p.351-76.
- ✓ Waltz C.F. (2010). Strcikland OL, Lenz ER. Measurement in Nursing and Health Research. New York: Springer Publishing Company, p.176-8.
- ✓ Çapık, C. (2014). Geçerlik Ve Güvenirlik Çalışmalarında Doğrulayıcı Faktör Analizinin Kullanımı, Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi, 17(3), 196-205.
- ✓ M.Sözbilir (2010). Ölçme ve Değerlendirme.
- ✓ Büyüköztürk, Ş. (2011). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Yayıncılık.
- ✓ Büyüköztürk, Ş., Çakmak, E.K., Akgün, Ö.E., Karadeniz, Ş. ve Demirel, F. (2010). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Yayıncılık.
- ✓ Güler, N. (2014). Eğitimde Ölçme ve Değerlendirme, Pegem Akademi: Ankara.

KAYNAKLAR

1. Al-Sunduqchi, Mahdi S. 1990. Determining the Appropriate Sample Size for Inferences Based on the Wilcoxon Statistics. Ph.D. dissertation under the direction of William C. Guenther, Dept. of Statistics, University of Wyoming, Laramie, Wyoming.
2. Bonett, Douglas. 2002. 'Sample Size Requirements for Testing and Estimating Coefficient Alpha.' Journal of Educational and Behavioral Statistics, Vol. 27, pages 335-340.
3. Chow, S.C.; Shao, J.; Wang, H. 2003. Sample Size Calculations in Clinical Research. Marcel Dekker. New York.
4. D'Agostino, R.B., Chase, W., Belanger, A. 1988. 'The Appropriateness of Some Common Procedures for Testing the Equality of Two Independent Binomial Populations', The American Statistician, August 1988, Volume 42 Number 3, pages 198-202.
5. Desu, M. M. and Raghavarao, D. 1990. Sample Size Methodology. Academic Press. New York.
6. Fleiss, Joseph L. 1986. The Design and Analysis of Clinical Experiments. John Wiley & Sons. New York.
7. Graybill, Franklin. 1961. An Introduction to Linear Statistical Models. McGraw-Hill. New York, New York.
8. Guenther, William C. 1977. 'Desk Calculation of Probabilities for the Distribution of the Sample Correlation Coefficient', The American Statistician, Volume 31, Number 1, pages 45-48.
9. Kirk, Roger E. 1982. Experimental Design: Procedures for the Behavioral Sciences. Brooks/Cole. Pacific Grove, California.
10. Lachin, John M. 2000. Biostatistical Methods. John Wiley & Sons. New York.
11. Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, MA.
12. <https://www.ncss.com/download/pass/updates/>
13. <https://ncss-wpengine.netdna-ssl.com/wp-content/uploads/2012/09/PASSUG1.pdf>.
14. Spicer, C. C. (1972). Algorithm AS 52: Calculation of power sums of deviations about the mean. Applied Statistics, 21, 226-227.
15. Zar, Jerrold H. 1984. Biostatistical Analysis (Second Edition). Prentice-Hall. Englewood Cliffs, New Jersey.