

- 1.Hipotez Nedir?
- 2.Yalın (Null) & (Alternative) Karşıt Hipotez.
- 3.Hipotez Testinin Aşamaları.
- 4.Değişkenler Arası Farklar ve İlişkiler
- 5.Değişkenler Arası Fark veya İlişkiler Neden Önemlidir?
- 6.Değişkenler Arası Fark veya İlişkide İki Temel Özellik.
- 7.İstatistiksel Anlamlılık Nedir? ( $p$  değeri)
- 8.İstatistiksel Anlamlılık ve Örneklem Sayısı Arasındaki İlişki.
- 9.Değişkenler Arası Güç ve Güvenilirlik İlişkisi
- 10.Fark veya İlişki Yok Sonucu Anlamlı Bir Sonuç mudur?
- 11.Değişkenler Arası İlişkinin Büyüklüğünü (gücünü) Nasıl Ölçebiliriz?
- 12.İstatistiksel Anlamlılık Derecesi Nasıl Ölçülür?

## 1. Hipotez Nedir?

İstatistiğe genelde ulaşılabilen veya bilinen örneklerden daha genel yığın veya evren hakkında genellemeler yapabilmek için başvururuz. Sahip olduğumuz veri veya yargılardan yola çıkarak teoriler geliştirir, istatistik yöntemlerle de teorilerimizi ispat etmeye veya çürütmeye çalışırız. Teorilerimizi ölçülebilir somut ifadeler haline getirdiğimizde hipotezler oluşur.

**Hipotezler teorilerimizi test etmek veya denemek için geliştirdiğimiz/belirlediğimiz somut ölçülebilir ifadelerdir.**

## 2. Yalın (Null) & Karşıt (Alternative) Hipotez

Hipotez testinde iki tür hipotezden bahsedebiliriz. Bunlardan birincisi ve en yaygın tercih edilen yalın hipotez (" $H_0$ ", hipotezi de denir ve  $H_0$  ile sembolize edilir) diğeri ise doğruluğuna yüksek derecede inanılan durumlarda kullanılan karşıt hipotezdir ( $H_1$  ile sembolize edilir). Bu hipotezler birbirlerinin karşıtıdır. Birisi doğru olduğunda diğeri yanlıştır. Biri reddedildiğinde diğeri kabul edilir. Hangi hipotezi test ettiğinizi göstermek için  $H_0=0$  veya  $H_1 \neq 0$  sembolleri kullanılır. Kolayca anlaşılacağı üzere birincisi yalın diğeri karşıt hipotezdir.

**Yalın (tarafsız, fark yoktur) ve Alternatif (Biri veya diğeri daha iyi, daha başarılı veya ilişkili gibi yanlı) olmak üzere iki tür hipotez geliştirilir.**

## 3. Hipotez Testinin Aşamaları.

Hipotezin seçimi ve yazımı.

Bir hipotez çok açık, yalın, somut ve ölçülebilir olmalıdır. Bir hipotezde kişiye, zamana ve yere göreceli değişken ifadeler bulunmamalıdır. Yazdığımız önerme kesinlikle ölçülebilir olmalıdır.

Bir hipotezde ařađıdaki sorular cevabını bulmalıdır.

1. Katılımcı-Kim veya Kime? (Öđrenciler kimlerdir)
2. Durum/řart-Nasıl? (Öđrencinin hangi řartlar altında, ne durumda veya nasıl yapması gerektiđi)
3. Davranıř-Ne veya Neyi? (Öđrencinin ne olması, bilmesi veya ne yapması gerektiđi)

Örnek:

Normal klinik řartlarında bir steteskop verildiđinde (Durum), tıp öđrencileri (Katılımcı) kalpteki ritm bozukluklarını %90 dođrulukla (Ölçüt) teřhis edebilir.

Bir okulda 1.kademede bulunan A ve B řubelerindeki öđrencilerin matematik dersindeki başarıları arasında bir fark olup olmadığı merak edilmektedir. A sınıfında 50 öđrenci bulunmakta matematik dersindeki başarı ortalaması 65 ve standart sapması 6 dır. B sınıfında 60 öđrenci bulunmakta başarı ortalaması 58 ve standart sapması 8 dir.

$H_0$  = A ve B sınıfındaki öđrencilerin Matematik dersinde (Verilen spesifik sınavlar ve başarı testleri dođrultusunda) gösterdikleri başarıları arasında bir fark yoktur. ( $H_0=0$ )

$H_1$  = A ve B sınıfındaki öđrencilerin Matematik dersinde (Verilen spesifik sınavlar ve başarı testleri dođrultusunda) gösterdikleri başarıları farklıdır. ( $H_1\neq 0$ )

" $H_1$ " hipotezide pozitif ve negatif yönlü (directional) olmak üzere geliştirilebilir.

Pozitif  $H_1$ = A sınıfındaki öđrenciler Matematik dersinde (Verilen spesifik sınavlar ve başarı testleri dođrultusunda) B sınıfındaki öđrencilerden daha başarılıdır. ( $H_1: \mu_A > \mu_B$ )

Negatif  $H_1$ = B sınıfındaki Matematik dersinde (Verilen spesifik sınavlar ve başarı testleri dođrultusunda) öđrenciler A sınıfındaki öđrencilerden daha başarısızdır. ( $H_1: \mu_B < \mu_A$ )

Hipotezin testi

Yukarıdaki örnekte araştırma sorusu bir fark olup olmadığıdır. Bu dođrultuda genel öneri yalın hipotezle yola çıkmaktır ( $H_0=0$ ). Ancak eldeki veriler A sınıfının daha başarılı olduğunun güçlü bir göstergesidir. Bu nedenle karşıt ( $H_1\neq 0$ ) veya karşıt hipotezlerden pozitif ( $H_1: \mu_A > \mu_B$ ) veya negatif ( $H_1: \mu_B < \mu_A$ ) yönlü hipotezi seçebiliriz.

Yalın ( $H_0=0$ ) hipotezi seçelim ve "Z" testini uygulayalım.

$$\hat{\sigma}^2 = \frac{n_1 \cdot S_1^2 + n_2 \cdot S_2^2}{n_1 + n_2 - 2}$$

$$\hat{\sigma}^2 = \frac{50 \cdot 36 + 60 \cdot 64}{50 + 60 - 2}$$

$$\hat{\sigma}^2 = \frac{5640}{108}$$

$$\hat{\sigma}^2 = 52,22$$

$$\hat{\sigma} = 7,22$$

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$Z = \frac{65 - 58}{7,22 \cdot \sqrt{\frac{1}{50} + \frac{1}{60}}}$$

$$Z = \frac{7}{7,22 \cdot 0,1915}$$

$$Z = 5,22$$

Hipotezin kabulü veya reddi

$p=0.05$  güvenilirlik seviyesinde ve yaklaşık 100 serbestlik derecesi için "T" tablo değeri  $\pm 1,66$ 'dır. Hesapladığımız "T=5,22" değeri tablo değerinden yüksek olduğundan iki ortalama arasındaki farkın istatistiksel olarak anlamlı bir fark olduğunu söyleyebiliriz. Bu nedenle ( $H_0=0$ ) hipotezini reddediyor ve ( $H_1: \mu_A > \mu_B$ ) hipotezini kabul ediyoruz.

4. Değişkenler Arası Farklar ve İlişkiler. İstatistik analizde genelde iki veya daha fazla değişken arasındaki fark ve ilişki bağıntısı araştırılmaktadır. Az önce verdiğimiz matematik dersinde A ve B şubesi öğrencilerinin başarıları örneğini hatırlayalım. Yaptığımız analizle A şubesindeki öğrencilerin B şubesindeki öğrencilerden daha başarılı olduğunu ortaya koymuştuk. Burada iki şube öğrencilerinin Matematik dersindeki başarıları arasında bir fark olduğu ispat edilmiştir. Arkasından arzu edersek bu farkı yaratan değişkenin ne olduğunu bulmak için araştırmalar yapabiliriz. Sınıfların ışık, ısı, gürültü, kalabalıklık, genişlik, ve diğer değişkenleri ile başarı arasında bir ilişki tespit edersek o zaman hangi değişkenlerin bu başarı farkında etkili olduğunu da ortaya koymuş oluruz. Dolayısıyla istatistik analizlerde esas olan değişkenler arası farklar ve ilişkilidir. Bizde daha çok fark ve ilişki analizleri üzerinde duracağız.

5. Değişkenler arasındaki fark veya ilişki neden önemlidir? Tüm araştırma ve bilimsel analizlerin nihai amacı değişkenler arasında bir fark veya ilişki bulmaktır. Bilim felsefesi bize bazı nitelikler ve nicelikler arasındaki farkların veya ilişkilerin olması veya olmaması dışında bir "anlamlılıktan" bahsetmenin mümkün olmadığını öğretiyor. Böylece bilimin ilerlemesi dediğimiz şey daima değişkenler arasındaki fark veya ilişkiyi anlatmaktadır. Teknoloji, bilim, gelişme dediğimiz her şey yeni bilginin üretilmesi işidir. Bu bilgide bazen doğadan doğrudan elde edilir bazen de doğada farklı nitelik ve nicelikler arasındaki fark ve ilişkilerde gizlidir.

6. Değişkenler arasındaki fark veya ilişkide iki temel özellik. Değişkenler arasındaki her ilişkinin iki temel özellik söz konusudur (a) GÜCÜ (veya "büyüklük") ve (b) GÜVENİLİRLİĞİ (veya "doğruluğu").

- a. GÜÇ (veya "Büüklük"). İlişkinin gücünü veya büyüklüğünü anlamak ve ölçmek güvenilirliğini anlamak ve ölçmekten daha kolaydır. Örneğin, cinsiyet ve beyaz kan hücreleri miktarlarıyla ilgili bir araştırmada tüm erkek deneklerimizde beyaz kan hücreleri miktarları kadınlarda olduğundan daha fazla çıksa, bizim örnekleminiz için cinsiyetle beyaz kan hücreleriyle ilgili değişkenler arasındaki ilişkinin gücü çok yüksek diyebiliriz. En azından kendi örnekleminiz için bir değeri bildiğimizde diğerini de tahmin edebiliriz.
- b. GÜVENİLİRLİĞİ (veya "doğruluğu"). Bir ilişkinin güvenilirliği anlaşılması daha zor bir kavram olmakla birlikte çok önemlidir. Bizim örnekleminizde elde edilen bulgunun popülasyona yani tüm evrene ne kadar genellebileceğini gösterir. Diğer bir değişle, aynı evrenden seçilecek diğer örneklem gruplarında aynı sonucu elde etmemin ne kadar muhtemel olduğunu gösterecektir. Unutmayınız ki araştırmalarda asla örneklem bulguları yegâne amacımız olmamaktadır. Örnekleme ancak evrene genellenebilir olduğu takdirde ilgi gösteririz. Güvenilirliğin sembolü "p" veya "s" veya "sig" dir. "p" olasılık (probability), "s" veya "sig" güvenilirlik (significance) anlamına gelmektedir.

## 7. "İstatistiksel anlamlılık" nedir (p-değeri).

Güvenilirliğin sembolü "p" dir. "p" olasılık (probability) anlamına gelmektedir. Elde edilen sonucun şans eseri elde edilme ihtimalini gösterir. Örneğin, "p=0.01" örnekleminizle elde ettiğimiz bulgunun şans eseri elde edilme ihtimalinin %1 olduğunu yani sonucun %99 güvenilir olduğunu gösterir yani değişkenler arasında %99 ihtimalle bir ilişki den bahsetmek mümkündür.

Veya hata payıdır da denilebilir. Örneğin, "p=0.01" örnekleminizle elde ettiğimiz bulgunun tüm evren için çalışıldığında elde edilen bulgudan farklı olma olasılığının %1 olduğunu anlatır. Yani "p" hata payını gösteriyor da diyebiliriz, hata payı %1 dir.

"p" değerinin ters ifadesi sonucun ne kadar temsil edici olduğunu gösterir. Örneğin, "p=0.01" örnekleminizle elde ettiğimiz bulgunun hatalı olma ihtimalinin %1 olduğunu ve sonucun evreni %99 oranında temsil ettiğini anlatır.

"p" hata payı tersi olan "s" veya "sig" güvenilirliği gösterir. Güvenilirlik "s" veya "sig" significance'dan gelir. "p" 0,05 ise "sig" 0,95 dir. Ancak bazı kaynaklarda "p" veya "s" eşanlamlı kullanılmaktadır.

"p" değeri için genel geçer sınır "0.05" yani "%5"dir. Elde edilen bulgu için "p" değeri %5 den (0.05) büyük bir değerse (p>.05) şeklinde ifade edilebilir, sonuç istatistiksel olarak anlamlı kabul edilemez, yani değişkenler arasında bir ilişki yoktur denilir. Ancak elde edilen bulgu için "p" değeri %5 den küçük veya eşitse (p<0.5) sonuç istatistiksel olarak anlamlı kabul edilir ve değişkenler arasındaki ilişki istatistiksel olarak kabul edilir. Ancak bu durumda örneklem için elde edilen değer evrene genellenebilir.

## 8. İstatistiksel anlamlılık ve örneklem sayısı arasındaki ilişki. Örneklem ve analiz sayısı arttıkça elde edilen değer güvenilir olma sınırı düşer. 100 erkek ve 100 kadın denekle cinsiyet ve beyaz kan hücresi miktarlarını çalıştığımızda, sonucun istatistiksel olarak anlamlı yani "p" değerinin 0.05 olması için erkeklerin %65'inin kadınlardan daha fazla beyaz kan hücresine sahip olması gerekiyorsa 1000 erkek ve 1000 kadın örnekleme çalışıldığında, sonucun istatistiksel olarak anlamlı yani "p" değerinin 0.05 olması için misalen erkeklerin %55'inin kadınlardan daha fazla beyaz kan hücresine sahip olması yeterli olacaktır. (Büyük sayılar kanunu).

Örnek."Erkek ve Kız Çocuk Oranı. Aşağıdaki örnek araştırmayı okuyun. (Nisbett, et al., 1987). İki hastanemiz var: birisinde günde 120 diğesinde 12 doğum gerçekleşiyor. Ortalama olarak, her iki hastanede de her gün doğan erkek bebeklerin kız bebeklere oranı %50'dir. Ancak bir gün, hastanelerden birinde erkek bebeklerin iki katı kız bebek doğmuştur. Hangi hastanede bunun gerçekleşmesi daha olasıdır? Bir istatistikçi için cevap kolay, sıradan birisi için cevap çok belirgin değildir. Ufak hastanede bunun gerçekleşmesi daha muhtemeldir. Çünkü, teknik olarak, örneklem sayısındaki artış, ortalama değerden farklılaşma ihtimalini azaltır. Büyük sayılar kanunu. 10 kez yazı tura attığınızda, 8 tura 2 yazı gelme ihtimali, 100 kez yazı tura attığınızda 80 tura 20 yazı gelme ihtimalinden çok çok daha fazladır.

9. Değişkenler arasındaki güç ve güvenilirlik ilişkisi. Elde edilen fark veya ilişkinin güçlü olması güvenilirlik derecesinin de yüksek olmasını anlatır. Örneğin, cinsiyet ve beyaz kan hücreleri miktarlarıyla ilgili araştırmada erkek deneklerimizin %70'inin beyaz kan hücreleri miktarları kadınlarda olduğundan daha fazla çıksa ve "p" değeride 0.05 olsa, bizim örneklemimiz için cinsiyetle beyaz kan hücreleriyle ilgili değişkenler arasındaki ilişkinin güçlü bir ilişki den bahsedebiliriz. Sonuçlar erkek örneklemimizin %90'ının beyaz kan hücreleri miktarları kadınlarda olduğundan daha fazla çıksa ve "p" değeri kesinlikle 0.05 den daha düşük bir değer çıkacaktır. Çünkü elde edilen değer daha güçlü bir değerdir. Değişkenler arasındaki ilişki artmaktadır ve arttıkçada güvenilir olma derecesi de artacaktır.

10."İlişki veya Fark yoktur" anlamlı bir sonuç mudur? Bir araştırma sonucu iki veya daha fazla değişken arasında her türlü sonuç olası ve değerlidir. İki değişken arasında bir fark veya ilişki olması nasıl bir "bilgi" ise olmaması da aynı değerde bir "bilgi"dir. Bu nedenle iki veya daha fazla değişken arasında bir ilişki olmaması da yadsınamayacak anlamlı bir sonuçtur. Bu nedenle araştırmacı önyargısız bir şekilde ilişki olup olmadığını araştırmalı ve asla ilişki olmalı beklentisine girmemelidir.

11.Değişkenler arasındaki farkın ve ilişkinin gücünün hesaplanması.

Değişkenler arasındaki ilişkinin gücünü ölçen istatistikçiler tarafından geliştirilmiş bir çok yöntem vardır. Belirli bir durumda bir yöntem tercih etmek, araştırmadaki değişken sayısına, kullanılan ölçme aracına, ilişkinin doğasına ve benzeri özelliklere bağlıdır. Her yöntem aslında bir genel kurala uyar: belirli değişkenler arasında gözlenen ilişkiyi "hayal edilebilen maksimum ilişki" ile mukayese etmeye teşebbüs ederler. Teknik olarak açıklamak gerekirse, böyle bir değerlendirmeyi gerçekleştirmenin ortak yolu değişkenlerin değerlerinin ne kadar farklı olduklarına bakmak ve sonra tüm bu farklılığın ne kadarının bahsedilen değişkenler için normal ne kadarının normal olmadığını hesaplamaktır. Daha az teknik anlatmak gerekirse, değişkenlerde gözlenen benzerlikle değişkenler arasındaki ilişki mükemmel olsaydı olması gereken benzerlikler karşılaştırılır. Basit bir örnek vermek gerekirse, farz edin ki cinsiyet ve beyaz kan hücresi miktarı araştırmasında, ortalama beyaz kan hücresi miktarı erkeklerde 100 kadınlarda da 102 dir. Bu durumda, ortalama olarak her bireyin BKHM (beyaz kan hücresi miktarı) değerinin genel ortalamadan (101) sapması (farkı) cinsiyetten kaynaklanan bir parça içerir diyebiliriz (bu araştırma için bu değer ölçüsü 1'dir).

Şu örneklere bir göz atalım;

- a. Tüm erkek örneklemimizin BKHM değerleri tam olarak 100 olsaydı, ve tüm kadınlarımızda 102, o zaman örneğimizde genel ortalamadan olan tüm sapmalar cinsiyete göre hesaplanabilirdi. Bizde örneklemimiz için cinsiyetle BKHM arasında

mükemmel bir ilişki olduğunu söyleyebilirdik, çünkü gözlenen örneklemelerin %100'ünde cinsiyete göre sabit BKHM değeri elde edilmiştir.

- b. Şayet BKHM değeri 0-1000 arasında bir değer alsaydı, araştırmada elde edilen bayan ve erkek örneklemeler arasındaki aynı (2 puanlık) fark örneklemelerin değerlerindeki tüm farklılığın o kadar ufak bir parçasına karşılık olacaktı ki muhtemel umursanmaz bir fark olarak değerlendirilecekti. Bu değer o kadar küçük olacaktı ki, mesela bir tane daha örneklem katılsa sonuçlar değişecek hatta ters dönebilecekti. Bu nedenle değişkenler arasındaki ilişkinin gücünü ölçen her iyi yöntem bireysel skorlardan kaynaklanan tüm farklılığı göz önüne almalı ve bu farklılığın ne kadarının normal (araştırma konusundan) kaynaklandığını hesaplamalıdır. (ne kadarı tesadüf en kadarı değişkenden)

12. "İstatistik anlamlılık" veya "Güvenilirlik derecesi" nasıl hesaplanır? Farzedinki hali hazırda iki değişken arasında bir ilişki değeri elde ettik. İkinci adım bu değer ne kadar anlamlı olduğunu bulmaktır. Örneğin, iki değişken arasındaki %40'luk fark anlamlı bir ilişki olarak kabul edilebilecek bir fark mıdır? Cevap elbette "duruma göre değişir" olacaktır. Özellikle anlamlılık örneklem sayısına bağlıdır. Daha öncede anlatıldığı gibi çok büyük örneklem grupları için ufak ilişki değerleri bile anlamlı olabilirken, çok küçük örneklem grupları için büyük ilişki değerleri bile anlamlı olmayabilir.

Anlamlılık veya güvenilirlik derecesi bilgisayarlarca örneklem genişliğine ve ilişkinin gücüne hesaplanarak bize hazır olarak sunulur. Ancak genel bilgi olarak güvenilirlik derecesinin normal dağılımla ilgili olduğunu ve bu ilişkinin doğasının ise aşağıda "Normal dağılım neden önemlidir?" kısmında anlatıldığı gibi olduğunu biliniz.