

# İstatistik Nedir?

İstatistik, ampirik verilerin toplanması, analiz edilmesi, yorumlanması ve sunulmasına yönelik yöntemlerin geliştirilmesi ve incelenmesi ile ilgilenen bir bilim dalıdır. Devletin nüfus sayımı verilerine ve çeşitli ekonomik faaliyetlere ilişkin bilgilere olan ihtiyacı, istatistik alanının ilk dönemlerinde büyük bir itici güç oluşturmuştur. İstatistik son derece disiplinler arası bir alandır; istatistik araştırmaları neredeyse tüm bilimsel alanlarda uygulanabilirlik bulur ve çeşitli bilimsel alanlardaki araştırma soruları yeni istatistiksel yöntemlerin ve teorinin geliştirilmesini motive eder. İstatistikçiler yöntem geliştirirken ve yöntemlerin altında yatan teoriyi incelerken çeşitli matematiksel ve hesaplama araçlarından yararlanırlar.

Prof. Dr. Bahadır Yüzbaşı

İstatistik alanındaki iki temel fikir belirsizlik ve varyasyondur. Bilimde (ya da daha genel olarak hayatta) karşılaştığımız ve sonucun belirsiz olduğu pek çok durum vardır. Bazı durumlarda belirsizlik, söz konusu sonucun henüz belirlenmemiş olmasından kaynaklanırken (örneğin, yarın yağmur yağıp yağmayacağını bilemeyebiliriz), diğer durumlarda belirsizlik, sonucun zaten belirlenmiş olmasına rağmen bunun farkında olmamamızdan kaynaklanır (örneğin, belirli bir sınavı geçip geçmediğimizi bilemeyebiliriz).

# Olasılık

Belirsiz olayları tartıřmak iin kullanılan matematiksel bir dildir ve olasılık istatistikte kilit bir rol oynar. Herhangi bir lm veya veri toplama abası, bir dizi varyasyon kaynađına tabidir. Bununla, aynı lmn tekrarlanması halinde cevabın muhtemelen deđiřeceđini kastediyoruz. İstatistikiler her durumda varyasyon kaynaklarını anlamaya ve kontrol etmeye (mmknse) alıřırlar.

# Anakitle (Popülasyon)

- Belirli bir istatistiksel model biçimine takılıp kalmadan önce, bilim insanlarının genellikle tüm bir varlık popülasyonu için geçerli sonuçlar bulmakla ilgilendiklerini hatırlamakta fayda var.
- Örneğin, psikologlar tüm insanlarda meydana gelen süreçleri keşfetmek isterler, biyologlar tüm hücrelerde meydana gelen süreçlerle ilgilenebilirler, ekonomistler tüm maaşlar için geçerli modeller oluşturmak isterler vb.
- Bir popülasyon çok genel (tüm insanlar) ya da çok dar (pamuk adındaki tüm erkek beyaz kediler) olabilir.
- Genellikle bilim insanları dar popülasyonlardan ziyade genel popülasyonlar hakkında çıkarımlarda bulunmaya çalışırlar. Örneğin, herkesin (ya da çoğu insanın) spor yaralanmalarının masajla düzeldiği sonucuna varabilirsek bu çok daha geniş bir etki yaratacaktır.

Prof. Dr. Bahadır Yüzbaşı

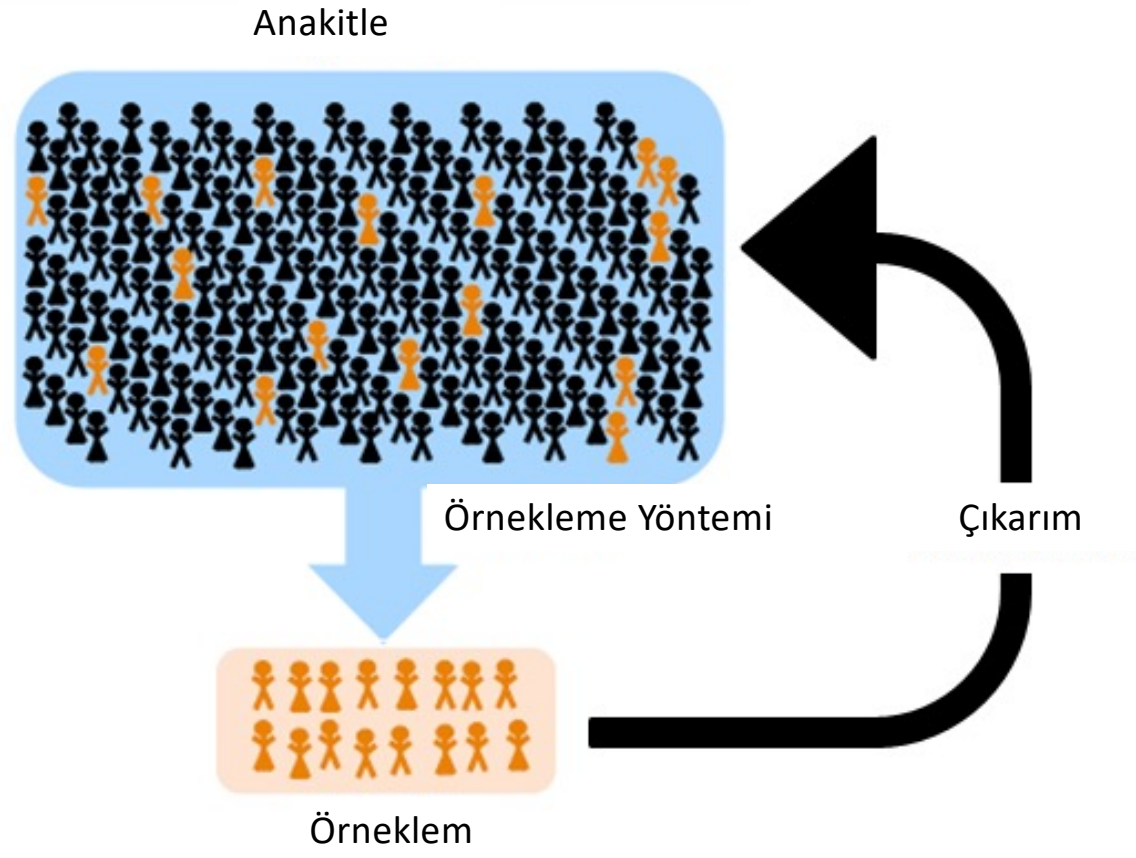
# Örnekleme (Sample)

- Köprü inşa eden bir mühendisin inşa etmek istediği köprünün tam boyutlu bir modelini yapamadığını ve bunun yerine küçük ölçekli bir model inşa ederek çeşitli koşullar altında test ettiğini hatırlayın.
- Küçük ölçekli modelden elde ettiği sonuçlardan, tam boyutlu köprünün nasıl tepki vereceği hakkında çıkarımlarda bulunmuştur.
- Küçük ölçekli model, köprünün tam boyutlu versiyonundan farklı tepki verebilir, ancak model ne kadar büyükse, tam boyutlu köprü ile aynı şekilde davranma olasılığı o kadar yüksektir.

# Örnekleme ...

- Bu metafor bilim insanlarına kadar genişletilebilir: bir popülasyonun (gerçek boyutlu köprü) her üyesine nadiren erişebiliriz. Bu nedenle, popülasyonun örneklem olarak bilinen daha küçük bir alt kümesinden (küçültülmüş köprü) veri toplar ve bu verileri popülasyonun bütünü hakkında çıkarımlarda bulunmak için kullanırız.
- Örneklem ne kadar büyük olursa, popülasyonun tamamını yansıtmaya olasılığı da o kadar yüksek olur. Popülasyondan birkaç rastgele örnek alırsak, bu örneklerin her biri bize biraz farklı sonuçlar verecektir, ancak ortalama olarak, büyük örneklerden elde edilen sonuçlar benzer olmalıdır.

Prof. Dr. Bahadır Yüzbaşı



Prof. Dr. Bahadır Yüzbaşı

# Gerçekten İstatistiğe güvenilir mi





# Örnek Anket Yöntemleri

- Gözlemsel çalışmalardan veri toplamak için örneklem araştırması yöntemleri, deneysel çalışmalardan veri toplamak için ise deneysel tasarım yöntemleri kullanılır. Tanımlayıcı istatistik alanı, öncelikle grafikler, tablolar ve sayısal özetler kullanarak verileri sunma ve yorumlama yöntemleriyle ilgilidir. İstatistikçiler bir örneklemde (yani popülasyonun bir alt kümesinden) elde edilen verileri kullanarak popülasyon hakkında çıkarımlarda bulduklarında istatistiksel çıkarım yapmış olurlar. Tahmin ve hipotez testi, istatistiksel çıkarımlar yapmak için kullanılan prosedürlerdir. Sağlık hizmetleri, biyoloji, kimya, fizik, eğitim, mühendislik, işletme ve ekonomi gibi alanlar istatistiksel çıkarımlardan geniş ölçüde yararlanır

Prof. Dr. Bahadır Yüzbaşı

# Değişkenler

- Bağımsız değişken: Bir etkinin nedeni olduğu düşünülen değişken. Bu terim genellikle deneysel araştırmalarda deneycinin manipüle ettiği bir değişkeni tanımlamak için kullanılır.
- Bağımlı değişken: Bağımsız değişkendeki değişikliklerden etkileneceği düşünülen değişken. Bu değişkeni bir sonuç olarak düşünebilirsiniz.
- Yordayıcı değişken: Sonuç değişkenini tahmin ettiği düşünülen değişken. Bu terim temelde 'bağımsız değişken' demenin başka bir yoludur. (Bazı insanlar bunu söylememden hoşlanmayacak olsa da; sadece yordayıcılar ve sonuçlar hakkında konuşursak hayatın daha kolay olacağını düşünüyorum).
- Sonuç değişkeni: Yordayıcı değişkendeki değişikliklerin bir fonksiyonu olarak değiştiği düşünülen bir değişken. Kolay bir yaşam için bu terim 'bağımlı değişken' ile eş anlamlı olabilir.

Prof. Dr. Bahadır Yüzbaşı

# Ölçüm seviyeleri

- Değişkenler kategorik ve sürekli olarak ayrılabilir ve bu türler içinde farklı ölçüm seviyeleri vardır:
- Kategorik (varlıklar farklı kategorilere ayrılır):
- İkili değişken: Sadece iki kategori vardır (örneğin, ölü veya canlı).
- Nominal değişken: İki'den fazla kategori vardır (örneğin, bir kişinin omnivor, vejetaryen, vegan veya meyveci olup olmadığı).
- Sıralı değişken: Nominal değişkenle aynıdır ancak kategorilerin mantıksal bir sıralaması vardır (örneğin, insanların sınavlarında başarısız mı, başarılı mı, başarılı mı yoksa farklı mı olduğu).
- Sürekli (varlıklar farklı bir puan alır):
- Aralıklı değişken: Değişken üzerindeki eşit aralıklar, ölçülen özellikteki eşit farklılıkları temsil eder (örneğin, 6 ile 8 arasındaki fark, 13 ile 15 arasındaki farka eşittir). Termometre.
- Oran değişkeni: Aralık değişkeni ile aynıdır, ancak ölçekteki puanların oranları da anlamlı olmalıdır (örneğin, anksiyete ölçeğinden 16 puan alan bir kişi gerçekte 8 puan alan birinden iki kat daha endişeli demektir). Bunun doğru olabilmesi için ölçeğin anlamlı bir sıfır noktası olmalıdır. Metre.

Prof. Dr. Bahadır Yüzbaşı

- Kategorik bir deęişken kategorilerden oluşur. Yapmanız gereken kategorik bir deęişken türünüze (örneğin insan, kedi, meyve, vb.) aşına olmanız gerekir. Siz bir insan, bir kedi ya da bir meyve: hem biraz kedi hem de biraz yarası olamazsınız.
- Kategorik bir deęişken, farklı varlıkları adlandıran bir deęişkendir. En basit haliyle erkek veya dişi gibi sadece iki farklı türde şeyi adlandırır. Bu bilinen bir şeydir
- İkili deęişken olarak tanımlanabilir. İkili deęişkenlerin dięer örnekleri canlı veya ölü olmak, hamile olmak ya da deęil, ve bir soruya 'evet' ya da 'hayır' yanıtını vermek. Her durumda sadece iki kategori vardır ve bir varlık iki kategoriden yalnızca birine yerleştirebilir.

Prof. Dr. Bahadır Yüzbaşı

# Merkezi eğilim ölçüleri

- Ortalama, tüm puanların toplamının puan sayısına bölünmesiyle elde edilir. Ortalamanın değeri, uç puanlardan oldukça fazla etkilenebilir.
- Medyan, puanlar artan sırada yerleştirildiğinde ortaya çıkan puandır. Ortalama kadar uç puanlardan etkilenmez.
- Mod, en sık ortaya çıkan puandır.

# Dağılım Ölçüleri

- Sapma veya hata, her bir puanın ortalamadan uzaklığıdır.
- Karesel hataların toplamı, ortalamadaki toplam hata miktarıdır. Hatalar/sapmalar toplanmadan önce kareleri alınır.
- Varyans, puanların ortalamadan ortalama uzaklığıdır. Kareler toplamının puan sayısına bölünmesiyle elde edilir. Bize puanların ortalama etrafında ne kadar dağınık olduğunu gösterir.
- Standart sapma, varyansın kareköküdür. Varyansın, onu hesaplamak için kullanılan puanların orijinal ölçüm birimlerine geri dönüştürülmüş halidir. Ortalamaya göre büyük standart sapmalar verilerin ortalama etrafında geniş bir şekilde yayıldığını gösterirken, küçük standart sapmalar verilerin ortalama etrafında sıkı bir şekilde toplandığını gösterir.
- Aralık (Range), en yüksek ve en düşük puan arasındaki mesafedir.
- Çeyrekler arası aralık, puanların orta %50'sinin aralığıdır.

Prof. Dr. Bahadır Yüzbaşı

# Dağılımlar ve z-skoru

- Frekans dağılımı, bir ölçüm ölçeğindeki her bir olası puanı ve bu puanın verilerde kaç kez oluştuğunu gösteren bir tablo veya grafik olabilir.
- Puanlar bazen z-skorları olarak bilinen standart bir biçimde ifade edilir.
- Bir puanı z-skoruna dönüştürmek için tüm puanların ortalamasını çıkarır ve sonucu tüm puanların standart sapmasına bölersiniz.
- z-skorunun işareti bize orijinal puanın ortalamadan üstünde mi yoksa altında mı olduğunu; z-skorunun değeri ise puanın standart sapma birimleri cinsinden ortalamadan ne kadar uzakta olduğunu söyler.

Prof. Dr. Bahadır Yüzbaşı

# Standart Hata

- Ortalamanın standart hatası, örnek ortalamalarının standart sapmasıdır. Bu nedenle, bir örnek ortalamasının popülasyonu ne kadar temsil edebileceğinin bir ölçüsüdür. Büyük bir standart hata (örnek ortalamasına göre), farklı örneklerin ortalamaları arasında çok fazla değişkenlik olduğu ve bu nedenle elimizdeki örnek ortalamasının popülasyon ortalamasını temsil etmeyebileceği anlamına gelir. Küçük bir standart hata, çoğu örnek ortalamasının popülasyon ortalamasına benzer olduğunu gösterir (yani, örnek ortalamamızın popülasyon ortalamasını doğru bir şekilde yansıtması muhtemeldir).



# Güven aralıkları

- Ortalama için bir güven aralığı, popülasyon ortalamasının örneklerin %95'inde bu aralığa düşeceği şekilde oluşturulmuş bir puan aralığıdır.
- Güven aralığı, popülasyon ortalamasının içinde kalacağından %95 emin olduğumuz bir aralık değildir.

# Boş (Null) hipotez anlamlılık testi (NHAT)

- NHAT, bilimsel teorilerin değerlendirilmesinde kullanılan yaygın bir yöntemdir. Temel fikir, iki rakip hipotezimiz olduğudur: biri bir etkinin var olduğunu (alternatif hipotez), diğeri ise bir etkinin var olmadığını (boş hipotez) söyler. Alternatif hipotezi temsil eden bir test istatistiği hesaplarız ve boş hipotezin doğru olması durumunda elde ettiğimiz değer kadar büyük bir değer elde etme olasılığımızı hesaplarız. Bu olasılık 0,05'ten küçükse, etki olmadığı fikrini reddeder, istatistiksel olarak anlamlı bir bulguya sahip olduğumuzu söyler ve küçük bir parti veririz. Olasılık 0,05'ten büyükse, etki olmadığı fikrini reddetmeyiz, anlamlı olmayan bir bulgumuz olduğunu söyleriz ve üzgün görünürüz.
- İki tür hata yapabiliriz: gerçekte olmadığı halde bir etki olduğuna inanabiliriz (Tip I hata); ve gerçekte olduğu halde bir etki olmadığına inanabiliriz (Tip II hata).
- İstatistiksel bir testin gücü, bir etki mevcut olduğunda bunu bulma olasılığıdır.
- Bir test istatistiğinin önemi örneklem büyüklüğü ile doğrudan bağlantılıdır: aynı etki farklı büyüklükteki örneklerde farklı p-değerlerine sahip olacaktır, küçük farklılıklar büyük örneklerde 'önemli' olarak kabul edilebilir ve büyük etkiler küçük örneklerde 'önemsiz' olarak kabul edilebilir.

# NHAT ile ilgili sorunlar...

- Pek çok bilim insanı NHAT'ı yanlış anlamaktadır. Anlamlılık testi ile ilgili yanlış anlaşılan birkaç örnek şunlardır:
- Önemli bir etki mutlaka önemli bir etki değildir.
- Anlamlı olmayan bir sonuç, sıfır hipotezinin doğru olduğu anlamına gelmez.
- Önemli bir sonuç, sıfır hipotezinin yanlış olduğu anlamına gelmez.
- NHST, p-değeri 0,05'in hemen altında olan bir etkinin önemli, p-değeri 0,05'in hemen üzerinde olan bir etkinin ise önemsiz olarak algılandığı ya hep ya hiç düşüncesini teşvik eder.
- NHST, araştırmacıların başlangıçtaki örnekleme çerçevelerinden sapmaları (örneğin, veri toplamayı planlanandan daha erken durdurmaları) nedeniyle yanlıdır.

Prof. Dr. Bahadır Yüzbaşı

# NHAT ile ilgili sorunlar

- Bilim insanlarının p-değerini etkileyebileceği pek çok yol vardır. Bunlar araştırmacının serbestlik derecesi olarak bilinir ve verilerin seçici olarak hariç tutulmasını, farklı istatistiksel modellerin uygulanmasını ancak yalnızca en olumlu sonuçlara sahip olanın raporlanmasını, veri toplamanın çalışmanın başlangıcında kararlaştırıldıktan farklı bir noktada durdurulmasını ve yalnızca p-değerini olumlu yönde etkileyen kontrol değişkenlerinin dahil edilmesini içerir.
- Bilimde anlamlı sonuçların yayınlanmasını ödüllendiren teşvik yapıları, araştırmacı serbestlik derecelerinin kullanımını da ödüllendirmektedir.
- p-hacking, anlamlı p-değerlerinin seçici bir şekilde raporlanmasına yol açan, çoğunlukla birden fazla analizin denenmesine ve yalnızca anlamlı sonuçlar veren analizin raporlanmasına neden olan uygulamaları ifade eder.
- Sonuçlar bilindikten sonra hipotez kurma (HARKing), bilim insanlarının veri analizinden sonra oluşturdukları bir hipotezi çalışmanın başlangıcında oluşturulmuş gibi sunmaları durumunda ortaya çıkar.

Prof. Dr. Bahadır Yüzbaşı

# Etki büyüklükleri ve meta-analiz

- Etki büyüklüğü, gözlemlenen bir etkinin büyüklüğünü, genellikle arka plan hatasına göre ölçmenin bir yoludur.
- Cohen's d, iki ortalama arasındaki farkın kontrol grubunun ortalamasının standart sapmasına veya her iki grubun standart sapmalarına dayanan birleştirilmiş bir tahmine bölünmesidir.
- Pearson korelasyon katsayısı,  $r$ , iki sürekli değişken arasındaki ilişkinin gücünü (ve yönünü) ölçebilen ve aynı zamanda sürekli bir değişken boyunca gruplar arasındaki farkı ölçebilen çok yönlü bir etki büyüklüğü ölçüsüdür.  $-1$  (mükemmel negatif ilişki) ile  $0$  (hiç ilişki yok) ve  $+1$  (mükemmel pozitif ilişki) arasında değişir.
- Olasılık oranı (odds ratio), bir olayın bir kategoride meydana gelme olasılığının diğerine göre oranıdır. Oranın  $1$  olması, belirli bir sonucun her iki kategoride de gerçekleşme olasılığının eşit olduğunu gösterir.
- Aynı hipotezi test eden farklı çalışmalardan elde edilen etki büyüklüklerini birleştirerek popülasyondaki bir etkinin büyüklüğünü tahmin etmeye meta-analiz denir.

# Bayes süreci

1. Bir hipotez (öncül tek bir değerdir) veya bir parametre (öncül bir olasılık dağılımıdır) hakkındaki öznel inançlarınızı temsil eden bir öncül tanımlayın. Öncül, tamamen bilgilendirici olmayabilir, bu da hemen hemen her şeye inanmaya hazır olduğunuz anlamına gelir, güçlü bir şekilde bilgilendirici olabilir, bu da ilk inançlarınızın oldukça dar ve spesifik olduğu anlamına gelir.
2. İlgili verileri inceleyin. Bizim anlamsız örneğimizde bu, aşık olduğunuz kişinin davranışlarını gözlemlemektir. Bilimde süreç bundan biraz daha resmi olacaktır.
3. Bayes teoremi, önceki dağılımı verilerle güncellemek için kullanılır. Sonuç, bir hipoteze olan yeni inancınızı temsil eden tek bir değer veya verileri gördükten sonra bir parametrenin makul değerlerine olan inancınızı temsil eden bir dağılım olabilen bir sonsal olasılıktır.
4. Bir sonsal dağılım, başlangıçta ilgilendiğiniz parametrenin bir nokta tahminini (belki de dağılımın tepe noktası) veya bir aralık tahminini (sonsal dağılımın belirli bir yüzdesini, örneğin %95'ini içeren bir sınır) elde etmek için kullanılabilir.

Prof. Dr. Bahadır Yüzbaşı

# Bayes Faktörleri

1. Bayes teoremi, gözlemlenen verilere dayanarak bir hipoteze olan ön inancınızı güncellemek için kullanılabilir.
2. Veriler göz önüne alındığında alternatif hipotezin olasılığı, veriler göz önüne alındığında boş hipotezin olasılığına göre posterior oran ile ölçülür.
3. Bayes faktörü, alternatif hipoteze göre verilerin olasılığının boş hipotezin olasılığına oranıdır. Bayes faktörünün 1'den büyük olması, gözlemlenen verilerin alternatif hipoteze göre sıfır hipoteze göre daha olası olduğunu gösterir. 1'den küçük değerler ise tam tersini gösterir. 1 ile 3 arasındaki değerler, alternatif hipotez için 'zar zor bahsedilmeye değer' kanıtları yansıtırken, 1 ile 3 arasındaki değerler 'önemli' kanıtlar, 3 ile 10 arasındaki değerler ise 'güçlü' kanıtlardır (Jeffreys, 1961).

Prof. Dr. Bahadır Yüzbaşı

# Grafikler

- Bir grafiğin dikey eksenini y eksenini (veya ordinat) olarak bilir.
- Bir grafiğin yatay eksenini x eksenini (veya abscissa) olarak bilir.
- İyi bir grafik çizmek istiyorsanız Tufte kültürünü takip edin:
- Y eksenini garip bir şekilde ölçeklendirerek verilerin ne gösterdiğine dair yanlış izlenimler yaratmayın (aynı şekilde etkileri de gizlemeyin).
- Grafik çöplerinden kaçının: Desenler, 3 boyutlu efektler, gölgeler, pembe kediler veya başka herhangi bir şey kullanmayın.



# Çarpıklık (Skewness) ve basıklık (kurtosis)

- Puan dağılımının yaklaşık olarak normal olup olmadığını kontrol etmek için çıktıdaki çarpıklık ve basıklık değerlerine bakın.
- Pozitif çarpıklık değerleri dağılımda çok fazla düşük puan olduğunu gösterirken, negatif değerler yüksek puanların biriktiğini gösterir.
- Basıklığın pozitif değerleri ağır kuyruklu bir dağılıma işaret ederken, negatif değerler hafif kuyruklu bir dağılıma işaret eder.
- Değer sıfırdan ne kadar uzaksa, verilerin normal dağılmama olasılığı o kadar yüksektir.
- Bu skorları standart hatalarına bölerek z skorlarına dönüştürebilirsiniz. Elde edilen puan (eksi işaretini yok saydığınızda) 1,96'dan büyükse anlamlıdır ( $p < 0,05$ ).
- Çarpıklık ve basıklığın anlamlılık testleri büyük örneklerde kullanılmamalıdır (çünkü çarpıklık ve basıklık normalden çok farklı olmasa bile anlamlı olmaları muhtemeldir).

Prof. Dr. Bahadır Yüzbaşı

# Normallik testleri

- Kolmogorov Smirnov (K-S) testi, bir puan dağılımının normal bir dağılımdan önemli ölçüde farklı olup olmadığını görmek için kullanılabilir (ancak kullanılmamalıdır).
- K-S testi anlamlıysa (SPSS tablosundaki Sig. değeri 0,05'ten küçükse) puanlar normal dağılımdan önemli ölçüde farklıdır.
- Aksi takdirde, puanlar yaklaşık olarak normal dağılmaktadır.
- Shapiro-Wilk testi de hemen hemen aynı şeyi yapar, ancak normallikten farklılıkları tespit etmek için daha fazla güce sahiptir (bu nedenle K-S testi anlamlı olmadığında bu test anlamlı olabilir).
- Uyarı: Büyük örneklerde, puanlar normal dağılımdan çok az farklı olsa bile bu testler anlamlı olabilir. Bu nedenle, özellikle tavsiye etmiyorum ve her zaman histogramlar, P-P veya Q-Q grafikleri ve çarpıklık ve basıklık değerleri ile birlikte yorumlanmalıdır.

Prof. Dr. Bahadır Yüzbaşı

# Varyansın homojenliđi

- Varyansın homojenliđi/homoscedasticity, sonu puanlarının yayılımının tahmin edici deđiřken zerindeki farklı noktalarda kabaca eřit olduđu varsayımdır.
- Bu varsayım, modelinizden elde edilen standartlaştırılmıř tahmini deđerlerin standartlaştırılmıř artıklara (zpred vs. zresid) karřı grafiđine bakılarak deđerlendirilebilir.
- Gruplar karřılařtırılırken, bu varsayım Levene testi ve varyans oranı (Hartley'in Fmax) ile test edilebilir.
- Levene testi anlamlıysa (SPSS tablosundaki Sig. 0.05'ten kkse), varyanslar farklı gruplarda nemli lde farklıdır.
- Aksi takdirde, varyans homojenliđi varsayılabilir.
- Varyans oranı, en byk grup varyansının en kđne blnmesiyle elde edilir. Bu deđerin ek materyaldeki kritik deđerlerden daha kk olması gerekir.
- Uyarı: Levene testini veya varyans oranını kullanmamak iin iyi nedenler vardır. Byk rneklemlerde grup varyansları benzer olduđunda anlamlı olabilirler ve kk rneklemlerde grup varyansları ok farklı olduđunda anlamlı olmayabilirler.

Prof. Dr. Bahadır Yzbařı

## Küçük örneklerle parametrik olmayan testler

- Parametrik veya parametrik olmayan bir test seçme kararınız en çok örneklerle küçük olduğunda (örneğin bir düzineden az değer) önemlidir.
- Parametrik bir test seçerseniz ve verileriniz Gauss dağılımından gelmiyorsa, sonuçlar çok anlamlı olmayacaktır. Parametrik testler, örnekler küçük olduğunda Gauss dağılımından sapmalara karşı çok sağlam değildir.
- Parametrik olmayan bir test seçerseniz, ancak gerçekten Gauss verisine sahipseniz, parametrik olmayan testler parametrik testlerden daha az güce sahip olduğundan ve küçük örneklerde fark belirgin olduğundan, çok büyük bir P değeri elde etmeniz muhtemeldir.
- Ne yazık ki normallik testleri, örneklem küçük olduğunda bir örneğin Gauss popülasyonundan gelip gelmediğini tespit etmek için çok az güce sahiptir. Küçük örnekler, tüm popülasyondaki dağılımın şekli hakkında güvenilir çıkarımlar yapmanızı sağlayacak yeterli bilgi içermez.

Prof. Dr. Bahadır Yüzbaşı

# Büyük örneklerle parametrik olmayan testler

- Parametrik veya parametrik olmayan bir test seçme kararı, büyük örneklerde (örneğin 100'den büyük) daha az önemlidir.
- Parametrik bir test seçerseniz ve verileriniz gerçekten Gauss değilse, parametrik testler Gauss varsayımının ihlaline karşı sağlam olduğundan, özellikle örneklem boyutları eşitse (veya neredeyse eşitse) çok fazla bir şey kaybetmezsiniz.
- Parametrik olmayan bir test seçerseniz, ancak gerçekten Gauss verilerine sahipseniz, örneklem büyüklüğü büyük olduğunda parametrik olmayan testler neredeyse parametrik testler kadar güce sahip olduğundan fazla bir şey kaybetmezsiniz.
- Normallik testleri, verilerin alındığı popülasyonun dağılımının şekli hakkında güvenilir çıkarımlar yapmanıza izin verecek kadar veri içeren büyük örneklerde iyi çalışır. Ancak normallik testleri sizin ilgilendiğiniz soruya yanıt vermez. Bilmek istediğiniz şey, dağılımın Gauss'tan parametrik testlerin kullanılabilirliğine şüphe düşürecek kadar farklı olup olmadığıdır. Ancak normallik testleri farklı bir soruya yanıt verir. Normallik testleri, dağılımın Gauss'tan farklı olduğuna dair kanıt olup olmadığı sorusunu sorar. Ancak büyük örneklerde normallik testleri Gauss'tan çok küçük sapmalar tespit edecektir; bu sapmalar parametrik testlerle parametrik olmayan testler arasındaki kararı etkilemeyecek kadar küçük farklılıklardır.

Prof. Dr. Bahadır Yüzbaşı

	Büyük örnekler (>100 ya da daha fazla)	Küçük örnekler (<12 ya da daha fazla)
Nongaussian veriler üzerinde parametrik testler	TAMAM. Testler sağlam.	Yanıltıcı. Sağlam değil.
Gauss verileri üzerinde parametrik olmayan testler	TAMAM. Testlerin gücü iyi.	Yanıltıcı. Çok az güç.
Normallik testinin kullanılabilirliği	Biraz faydalı.	Çok yararlı değil.