

Hazırlayan

Yeşim ÖZTÜRK

Değişkenler

ve

Tanımlayıcı

İstatistikler

1.Değişken ve Değişken Çeşitleri

Değişkenin kelime anlamı; Değişme özelliği gösteren, çok değişken, değişebilir, kararsız, değişici şeklindedir. Matematiksel tanımı ise; Gözlemden gözleme değişik değerler alabilen objelere, özelliklere ya da durumlara “değişken” denir.

İstatistikte değişken(variable) terimi, deneklere ait özellikler anlamında kullanılır. Örneğin araştırmaya katılanların cinsiyeti, medeni hali, gelir düzeyi, eğitim düzeyi, coğrafik konumu gibi demografik özellikleri ve boyu, kilosu, vizeden aldığı not, sorulara verdikleri cevaplar gibi özellikler birer değişkeni ifade ederler.

Değişken aşağıdaki şekilde sıralanabilir:

1. Kantitatif Değişken (Sayısal, Nicel)
 - 1.1.1. Kesikli (Süreksiz) Değişken
 - 1.1.2. Sürekli Değişken
2. Kalitatif Değişken (Sayısal Olmayan, Nitel)

1.1. Nicel (Kantitatif) Değişkenler

Birimlerin ölçüm ve tartım sonucu değerleri saptanan sayısal özelliklerini belirten değişkenlerdir. Bu değişkenler değerleri, mekanik ve elektronik araçlara sayısal olarak aralıklı ölçekli ya da orantılı ölçekli verileridir. Nicel değişkenlerin verilerine nicel veri adı verilir. Örneğin; birimlerin boy uzunluğu, vücut ağırlığı, kilosu, kan basıncı gibi özellikler nicel değişkenlerdir.

1.1.1. Kesikli (Süreksiz) Değişken

Sayılabilen değişkenlerdir. Örneğin; öğrenci sayısı, konut sayısı, tarım arazisi sayısı, kağıt sayısı, çiftçi sayısı vb.

1.1.2. Sürekli Değişken

Ölçülebilen ve tartılabilen değerleri içeren değişkenlerdir. Örneğin; uzunluk, ağırlık, yaş, alan, güç, zeka düzeyi, sıcaklık, basınç, hız vb.

1.2.Nitel (Kalitatif) Değişkenler

Nitel değişkenler; birimlerin kalite, kategorik ya da isimsel olarak belirtilebilen karakteristik özelliklerini, durumlarını ve pozisyonlarını belirten değişkenlerdir. Bu değişkenlerin verileri isimsel ya da sıralı ölçekle edilmişlerdir ve iki ya da daha fazla kategoriye(alt seçenek, sınıf, grup) ayrılarak sayımla elde edilir. Bu değişkenlerin

verilerine nitel veriler adı verilir. Örneğin; birimlerin cinsiyeti, kan grubu, medeni durumu, göz rengi, mesleği, yerleşim yeri, tuttuğu futbol takımı(fanatikler için) gibi nitelik bildiren durumları açıklayan değişkenlerdir.

2.Ölçme ve Ölçek Düzeyleri

Ölçme, herhangi bir niteliği gözlemek ve gözlem sonucunda önceden tespit edilen belirli kurallara göre nesnelere ve kişilerin özelliklerinin sayılar ve semboller atamaktır. Varlık veya olayların belli bir özelliğe sahip oluş derecelerini belirleme işidir. Bir kumaş boyunun cetvel ile ölçülmesi gibi.

Ölçekler ise, deneklerin belirli veriler baz alınarak sınıflandırıldıkları araç veya mekanizmalardır. Nesnelere sayılar vermede ve bu sayıların kullanılmasında uyulması gereken kuralları ve kısıtlamaları belirtmek için kullanılır.

Araştırma yöntemleri bağlamında dört temel ölçüm seviyesi vardır:

- Sözcük (nominal) ölçüm
- Sıralı (ordinal) ölçüm
- Aralı (interval) ölçüm
- Oranlı (rasyo) ölçüm

2.1.Sözcük (Nominal) Ölçüm:

Belirli bir gruba ait olup olmamayı gösterir. Nesnelere sadece gruplandırıldığı bir ölçme seviyesidir. Sayılar ve harfler bir kimlik olmaktan öteye gitmemektedir. Nesnelere verilen sayıların sayısal anlamları yoktur ve miktar belirtmezler. Göz rengi, cinsiyet, illere göre kod numarası verilmesi gibi.

Kullanılan istatistiksel işlemler:

- Parametrik olmayan testler kullanılır.
- Normal dağılım göstermez. Heterojen yapılıdır.
- Frekanslar bulunarak karşılaştırmalar yapılabilir.
- Yüzdeler değeri bulunarak karşılaştırmalar yapılabilir.
- Sınıflarda yığılımın homojenliği test edilebilir.
- Ortalama olarak mod kullanılabilir.

- Nominal ölçeklere uygulanan uygunluk, ilişki, farklılık, etki testleri yapılabilir.
- Çapraz tablolar oluşturularak analizler yapılabilir.

2.2.Sıralı (Ordinal) Ölçüm:

Sıra bildiren ölçüler verir. Bir özelliğe en az sahip olandan en çok sahip olana ya da en çok sahip olandan en az sahip olana doğru bir sıralama yapmak mümkündür. Yani bir nesnenin belirli bir özelliği az mı yoksa çok mu taşıdığını söylemek mümkündür. Sıralananların birbirleri arasındaki farkları konusunda bir şey söylemek imkansızdır. Üç yarışçının sıralamasının bilinmesi fakat ne kadar süre farkı olduğunun bilinmemesi gibi.

Kullanılan istatistiksel işlemler;

- Parametrik olmayan testler kullanılır.
- Normal dağılım göstermez. Heterojen yapılıdır.
- Frekanslar bulunarak karşılaştırmalar yapılabilir.
- Yüzdelerik değerler bulunarak karşılaştırmalar yapılabilir.
- Sınıflarda yığılımın homojenliği test edilebilir.
- Sıralı ölçeklere uygulanan uygunluk, ilişki, farklılık etki testleri yapılabilir.
- Ortalama olarak medyan kullanılabilir.
- Kartil aralığı, değişim aralığı vb. belirtici istatistikler hesaplanabilir.
- Çapraz tablolar oluşturularak analizler yapılabilir.
- Dizilişlerin (art arda gelişlerinin) rassallığı test edilebilir.

2.3.Aralı (Interval) Ölçüm:

Sayısal olarak eşit aralıkların eşit mesafeleri temsil ettiği bir ölçek olup nesnelere sıralanmasında kullanılmaktadır. 1 ile 2 arasındaki mesafe ile 4 ile 5 arasındaki mesafe eşittir. Sıfır noktası sabit olmadığı için başlangıç noktası görecelidir. Burada sıfır noktası sayısallaştırmayı kolaylaştıran bir başlangıç noktasıdır, doğal ya da mutlak sıfır noktası değildir. Oda sıcaklığı sıfır derecedir dendiğinde buradaki sıfır gerçek değildir, çünkü oda sıcaklığının olmadığı söylenemez.

Kullanılan istatistiksel işlemler;

- Parametrik testler kullanılır.

- Normal dağılım gösterir. Homojen yapılıdır.
- Frekanslar ve yüzdelik değerler bulunarak karşılaştırmalar yapılabilir.
- Ortalamalar, standart sapma, hata varyans vb. belirtici istatistikler hesaplanabilir.
- Aralık ölçeklere uygulanan uygunluk, ilişki, farklılık, etki testleri yapılabilir.

2.4.Oranlı (Rasyo) Ölçüm:

Daha önceki ölçeklerin tüm özelliklerini taşımakla beraber mutlak sıfır noktasına da sahiptir. En üst ölçüm seviyesidir ve her türlü istatistiksel ve matematiksel işleme imkan tanımaktadır. Cevaplayıcı sıfırdan herhangi uygun bir sayıya kadar cevap verebilir. Yıllık kazancınız ne kadar, kaç çocuğunuz var gibi.

2.5.Ölçek Türleri

1. Sürekli Ölçekler
2. Tekli Ölçekler
 - 2.1. Kategorili
 - 2.2. Sıralama
 - 2.3. Sabit-Toplam
 - 2.4. Grafikselsel/Şekilsel
 - 2.5. İkili-Karşılaştırmalı
3. Çoklu Ölçekler
 - 3.1. Likert
 - 3.2. Semantik Farklılıklar

2.5.1.Sürekli Ölçekler

Bu ölçeklerde cevaplayıcı iki uç arasında uzanmakta olan bir çizgi üzerinde uygun bir yere işaret koyar. Değerlendirme aşamasında araştırmacı cevap çizgisini gereken sayıda aralığa böler ve cevaplayıcının cevabını rakamsal olarak ortaya koyar. Oluşturulması kolay fakat skorların belirlenmesi zor ve güvenilir değildir.

2.5.2.Tekli Ölçekler

Tek madde veya sorudan ibarettir.

Kategorili Ölçekler: Cevap seçenekleri kategoriler halinde verilmektedir. Kategoriler arasında çakışma olmamalıdır, yani her denek sadece bir seçeneğe ait olmalıdır.

Sıralama Ölçekleri: Deneklere inceleme konusuyla ilgili özellikler listesi verilir ve onlardan listedeki özellikleri tercih, önem sırası veya başka bir ölçüte göre sıralaması istenir.

Sabit-Toplam Ölçekleri: Deneklerden, kendilerine verilen 100 tam puanın önem derecesi veya hoşlanma seviyesine göre listede verilen özelliklere dağıtılması istenir. Verilen puanların toplamı tam olarak 100 olmak durumundadır.

Grafiksel/Şekilsel Ölçekler: Özellikle okuma yazma bilmeyen ve küçük çocuklar üzerinde yapılan bu anketlerde cevaplar kelime halinde değil şekiller ve skeçlerle verilmektedir.

İkili-Karşılaştırmalı Ölçekler: Deneklerden ikililer arasında seçim yapması istenmektedir.

2.5.3.Çoklu Ölçekler

Birden çok maddeden meydana gelen ve maddelerden aynı ölçeğe göre değerlendirildiği ölçeklere denir. Tek boyutlu olarak ölçülmesi mümkün olmayan çok boyutlu kavramların ölçülmesinde kullanılır.

Likert Ölçeği: Sosyal bilimler alanında yaygın olarak kullanılmaktadır. Deneklere çeşitli yargılar ve ifadeler yöneltilir ve bu yargılara katılıp katılmama derecelerini belirtmeleri istenir. Örneğin; “1.Kesinlikle katılıyorum./ 2.Katılmıyorum./ 3.Kararsızım./ 4.Katılmıyorum./ 5.Kesinlikle katılmıyorum.” gibi seçenekler. Daha çok kişilerin tutum ve eğilimlerini ölçme amaçlı kullanılır. En önemli dezavantajı doldurma için gerekli sürenin uzunluğudur. Kategori sayısı olarak 5, 7, 9 ve 11’li olabilir. En sık kullanılanı 5’li olmaktadır.

Değerlendirmede iki alternatif bulunmaktadır:

- Ölçekte yer alan her ifade ayrı ayrı değerlendirilir. İfadeler arası korelasyon ve ilişki incelenebilir.
- Skorların toplamı üzerinden değerlendirme yapılabilir.

Semantik Farklılıklar Ölçeği: Mamul, marka ve firma imaj çalışmalarında yaygın olarak kullanılır. İncelenen nesne/marka/ürüne ait çeşitli özelliklere ilişkin nitelendirmeler zıt kutuplara gelecek şekilde yedi noktalı bir doğru üzerine yerleştirilmiştir. Denek kendine uygun bir yere işaret koyar. Bu ölçek ile aralık seviyesinde ölçüm yapmak amaçlanmıştır. 1-7 şeklinde veya -3-+3 şeklinde ölçeklendirilebilir.

3.Genel Tanımlayıcı İstatistikler

Bir veya birden fazla dağılışı karşılaştırmak için kullanılan veya ayrıca örnek verilerden hareketle frekans dağılımlarını sayısal olarak düzenleyen değerlere tanımlayıcı istatistikler denir.

Ham veriler, sahip oldukları özelliklere ilişkin ipuçlarını yansıtmazlar. Düzenlenmemiş, dağınık haldeki verilere bakarak, betimleyici ifadelerde bulunmak mümkün değildir. Bu yüzden, verilerin genel hatlarıyla resmini çekmemizi sağlayacak şekilde düzenlenmesi gerekir.

Frekans dağılımı, yaygın olarak kullanılan veri düzenleme yöntemidir. Basit olarak, bir frekans dağılımı; verileri gruplara veya sınıflara ayırarak, her bir grup ya da sınıftaki gözlem sayılarının gösterilmesi sürecidir.

Aritmetik Ortalama: Üzerinde inceleme yapılan veri setindeki elemanların toplanıp incelenen eleman sayısına bölünmesiyle elde edilen yer ölçüsüne aritmetik ortalama denir. Serideki bütün değerlerden etkilenir. Halk dilinde ortalama ifadesi kullanıldığında ilk akla gelen kavram aritmetik ortalamadır. Örneğin; sınav notlarının ortalaması, yaz aylarında metrekaresine düşen ortalama yağış miktarı.

Özellikleri:

- Aritmetik dizi şeklinde artış veya azalış gösteren serileri en iyi temsil eden parametrik merkezi eğilim ölçüsüdür.
- Serideki birimlerin aritmetik ortalamadan sapmalarının toplamı sıfırdır.
- Serideki birimlerin aritmetik ortalamadan sapmalarının kareleri toplamı minimumdur.

- Örnek gözlemlerin tümü a gibi bir sabit ile çarpılırsa bu yeni veri setinin aritmetik ortalaması da eski veri setinin aritmetik ortalamasının a ile çarpımı kadar değişir.
- Örnek gözlemlerin tümü a gibi bir sabit ile toplanırsa bu yeni veri setinin aritmetik ortalaması da eski veri setinin aritmetik ortalamasının a ile toplamı kadar değişir.
- Aritmetik ortalama tüm verileri hesaplama fonksiyonu içinde kullanması nedeniyle güçlü bir istatistiktir.
- Aritmetik ortalama verilerdeki uç değerlerden etkilenmesi ise bu istatistiğin zayıf yönünü oluşturur.

Mod: Bir veri setinde en çok gözlenen (en çok tekrar eden) değere veya frekansı en fazla olan şans değişkeni değerine mod adı verilir. Mod, tüm ölçme düzeyleriyle ölçülmüş değişkenler için kullanılabilir. Aşırı değerlerden etkilenmez. Duyarlılık göstermez. Normal dağılım göstermez. Parametrik olmayan ortalamalar içinde yer alır. Merkezi dağılım ölçüsüdür. Kalitatif değişkenler için yani sınıflama ve sıralama ölçeklerde kullanılır. Basit seride mod, en çok tekrar eden değerdir. Sınıflandırılmış seride frekansı en büyük olan değerdir. Gruplandırılmış seride frekansı en büyük olan sınıf içerisinde yer alır.

Özellikleri:

- Ortalamalar arasında mod en temsili olandır. Çünkü verileri en çok tekrar eden değer mod değeridir.
- Mod değeri tamsayı olarak elde edilir.
- Mod, serideki aşırı terimlerin etkisi altında kalmaz. Uç değerlere karşı duyarlı değildir.

Medyan: Bir veri seti büyükten küçüğe veya küçükten büyüğe sıralandığında tam orta noktadan veri setini iki eşit parçaya ayıran değere medyan adı verilir. Veri setinde aşırı uçlu elemanlar olduğunda aritmetik ortalamaya göre daha güvenilirdir. Medyan, veri setindeki tüm elemanlardan etkilenmez. Basit ve sınıflandırılmış serilerde $(n+1)/2$ 'nci veridir. Gruplandırılmış serilerde $n/2$ 'nci veridir. Sıralama ölçeklerde kullanılır. Parametrik olmayan ortalamalar içinde yer alır. Merkezi dağılım ölçüsüdür. Normal dağılım göstermez.

Özellikleri:

- Verilerin medyandan mutlak sapmalarının toplamı minimumdur. Yani $|x - \text{Med}| = \min$ olur.
- Medyan serideki aşırı terimlerden etkilenmez.

Değişim Aralığı: En büyük gözlem değeri ile en küçük gözlem değeri arasındaki farktır. Yani, değişim aralığı " $R = DA = X_{\max} - X_{\min}$ " formülü ile gösterilir. Aşırı değerlere karşı duyarlıdır. Parametrik olmayan dağılım ölçüleri içerisinde yer alır. Değişkenlik ölçüsüdür.

Kartiller (Dörde Bölenler): Bir veri setinin büyükten küçüğe veya küçükten büyüğe sıralandığında dört eşit parçaya ayıran üç değere kartiller adı verilir. İlk %25'lik kısmı içinde bulunduran 1. Kartil (Q_1), %50'lik kısmı içinde bulunduran 2. Kartil (Q_2), %75'lik kısmı içinde bulunduran 3. Kartil (Q_3) olarak adlandırılır. %50'lik kısmı içinde bulunduran 2. Kartil (Q_2) aynı zamanda veri setinin medyanıdır.

Aritmetik Ortalama	Medyan	Mod
Sayısal verilerde kullanılır. Simetrik (normal dağılmış) verilerde kullanılır. Ölçek düzeyi en az aralık (interval) ölçek olmalıdır. Kıyaslama amacıyla kullanılır. Serideki bütün terimler kullanılır.	Sayısal verilerde ve mantıksal sıraya dizilebilen sayısal olmayan verilerde kullanılır. Simetrik (normal dağılmış) ve simetrik olmayan (çarpık) verilerde kullanılır. J ve ters J serilerinde kullanılmaz. Ölçek düzeyi en az sıralı (ordinal) ölçek olmalıdır. Seriye temsil etmek amacıyla kullanılır. Serideki bazı terimler kullanılır.	Sayısal, sayısal olmayan verilerde kullanılır. Simetrik (normal dağılmış) ve simetrik olmayan (çarpık) verilerde kullanılır. U serilerinde kullanılmaz. Her ölçek düzeyinde kullanılır. Seriye temsil etmek amacıyla kullanılır. Serideki bazı terimler kullanılır.

Kartil Aralığı (Kartil Sapma): Üçüncü kartilden birinci kartilin çıkarılmasıyla elde edilir ve “ $K.A=Q_3-Q_1$ ” biçiminde gösterilir. Değeri ne kadar büyük olursa, değişkenliği o kadar fazla demektir. Kartil aralığı, serinin baştan ve sondan %25’lik kısımlarını ihmal ettiğinden genellikle açık sınıflardan etkilenmez. Değişim aralığına nispeten uç değerlerden daha az etkilenir. Serinin yarısı ihmal edildiği ve hesaplama bütün değerlerin katılmaması nedeniyle hatalı sonuçlara yol açmaktadır. O yüzden bu oranın küçülmesinde yarar vardır. Kartil aralığı ikiye bölündüğünde yarı kartil aralığı elde edilir.

Ortalama Sapma: Gözlem değerlerinin ortalama değerinden çıkarılıp mutlak değeri alınıp toplam gözlem sayısına (n) bölünmesiyle elde edilir. Değişkenlik ölçüleri arasında yer alır.

Standart Sapma: Gözlem değerlerinin ortalamalarından olan sapmalarının karelerinin toplamının toplam gözlem sayısına (N veya n-1) oranının kareköküne denir. Ana kütle standart sapması için toplam gözlem sayısı “N” olarak, örneklem standart sapmasında ise “n-1” olarak ele alınır. Dağılım (değişkenlik) ölçüleri arasında yer alır. Parametrik dağılım ölçüleri içerisindedir. Normal dağılım gösterir.

- Bir veri grubundaki bütün birimlerin hesaba katılmasıyla belirlendiği için, uygulamada en fazla kullanılan ve en güvenilir dağılım ölçüsüdür.
- Eşit aralıklı ve oranlı sayısal ölçek düzeylerinde kullanılır.
- Bir veri grubundaki verilerin ortalama değerlerden ne kadar uzaklıkta yoğunlaştığını, yani verilerin ne kadar yaygınlıkta dağıldığını gösterir.
- Standart sapma büyüdükçe değişkenlik artar.
- Örneklem sayısı arttıkça standart sapma küçülür.
- Standart sapmanın küçük olması, ana kütle parametresinin daha güvenilir bir biçimde elde edilmesini sağlar.

Varyans: Gözlem değerlerinin ortalamalardan olan sapmalarının karelerinin toplamının toplam örnek sayısına oranıdır. Yani, standart sapmanın karesidir. Dağılım (değişkenlik) ölçüleri arasında yer alır. Parametrik dağılım ölçüleri içerisindedir. Normal dağılım gösterir.

- Bir veri grubundaki bütün birimlerin hesaba katılmasıyla belirlendiđi için, uygulamada en fazla kullanılan ve en güvenilir dağılım ölçüsüdür.
- Eşit aralıklı ve oranlı sayısal ölçek düzeyinde kullanılır.
- Bir veri grubundaki verilerin ortalama değerlerden ne kadar uzaklıkta yoğunlaştığını, yani verilerin ne kadar yaygınlıkta dağıldığını gösterir.
- Varyans büyüdükçe deđişkenlik de artar.
- Varyansın küçük olması, ana kütle parametresinin daha güvenilir bir biçimde elde edilmesini sağlar.

KAYNAKLAR

AKSARAYLI, M., Dokuz Eylül Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, “Merkezi Eğilim ve Değişkenlik Ölçüleri”, İzmir.

ALTUNEL, M., 2014. SETA Eğitim ve Sosyal Politikalar Direktörlüğü, “Araştırmalarda Ölçme ve Ölçekler”, Ankara.

ÇOBANOĞLU, F., 2016. Aydın Adnan Menderes Üniversitesi, Ziraat Fakültesi, Tarım Ekonomisi Bölümü, “Uygulamalı İstatistik” Ders Notları, Aydın.

ÇOBANOĞLU, F., 2017. Aydın Adnan Menderes Üniversitesi, Ziraat Fakültesi, Tarım Ekonomisi Bölümü, “Tarım Ekonomisi İstatistiği” Ders Notları, Aydın.

ÇOBANOĞLU, F., 2018. Aydın Adnan Menderes Üniversitesi, Ziraat Fakültesi, Tarım Ekonomisi Bölümü, “Çeşitli İstatistiksel Analiz Yöntemleri ve Politik Etki Değerlendirme Eğitimi” Ders Notları, Aydın.

ÖZER, O.O., 2019. Aydın Adnan Menderes Üniversitesi, Ziraat Fakültesi, Tarım Ekonomisi Bölümü, “Ekonometrik Yöntemler ve Modeller Ders Notları” – <https://oorkan.wordpress.com/>, Aydın.