

## TEMEL KAVRAMLAR ve VERİLERİN ÖZETLENMESİ

### 1.1. Giriş

Hangi bilim dalı olursa olsun, araştırılan olaylar hakkında bilgi edinebilmek için araştırmaların yapılması ve bu araştırmalardan elde edilen verilerin istatistik yöntemler kullanılarak analiz edilmesi, ulaşılan sonuçların yorumlanması ve genelleştirmelerin yapılması gerekir.

Üzerinde durulan olaylar hakkında ayrıntılı bilgi edinilmesi için mutlaka deneme yapmak zorunludur. Deneme, bilimin ilerlemesi için temeldir. Araştırmacılar, denemelerden elde edilen bilgiler doğrultusunda teorilerini geliştirir, test eder, doğru veya yanlış olduğunu gösterirler.

Deneme, yeni gerçeklerin gözlenmesi veya daha önceki sonuçların doğrulanması veya yanlış olduğunun gösterilmesi için planlanmış ve veri toplama ile sonuçlanan çalışma veya işlemler olup etrafımızda meydana gelen olayların anlaşılır bir şekilde açıklanması için kurulur. Deneme sonuçları önceden kesin olarak tahmin edilemeyen işlemlerdir. Denemeler, en uygun zamanda ve masrafla en güvenilir sonuçların alınacağı şekilde kurulmalı, elde edilen veriler istatistik yöntemler kullanılarak analiz edilmeli ve sonuçlar genelleştirilmelidir.

Bilim dallarının gelişme süreçleri incelendiği zaman, istatistik metot kullanımının yoğunlaşması ile paralellik gösterdiği açık bir şekilde görülür. Çünkü yapılan araştırmalardan elde edilen verilerin uygun istatistik metotlar kullanılarak analiz edilmesi, araştırılan konu ile ilgili hangi noktaların daha fazla araştırılması ve nelere dikkat edilmesi hakkında daha fazla bilgi verecektir. Her bilim dalında kullanılan deney materyali, deneme koşulları ve istatistik metotların gerektirdiği ön koşulların gerçekleşme durumu bir diğerinden farklı olduğundan konulara özgü istatistik yöntemlerin geliştirilmesi zorunlu olmuştur. Bunun sonucu olarak da istatistik biliminde Ekonometri, Psikometri, Sosyometri, Antropometri, Biyoistatistik veya Biyometri ve son yıllarda gelişiminde Türk Bilim adamlarının da rol oynadığı Kemometri gibi değişik bilim dalları ortaya çıkmıştır.

### 1.2. Temel Kavramlar

Yapılan bir araştırmada verilerin doğru olarak toplanması, toplanan verilerin özetlenmesi, tanıtıcı ve/veya tanımlayıcı değerlerin hesaplanması, birden fazla özellik dikkate alınmış ise özellikler arasındaki ilişkilerin hesaplanması, toplanan verilerin araştırmada dikkate alınan faktörler bakımından analizi edilmesi, elde edilen sonuçların değerlendirilmesi, yorumlanması ve tümevarım yaklaşımı ile genelleştirilmesi için kullanılan metotlar **İSTATİSTİK METOTLAR** olarak bilinir.

Araştırmacının üzerinde çalıştığı aynı veya benzer özelliğe sahip bireylerin tamamına **POPULASYON** (kitle = evren = kütle) denir. Araştırmacının çalışma konusuna bağlı olarak söz konusu bireyler öğrenci, hasta, koyun, kuzu, inek, toprak, ağaç vb. gibi her şey olabilir. Populasyon, aslında teorik bir kavram olup, üzerinde durulan özellik bakımından sonsuz sayıda bireyden oluşmuş bir kümedir. Ancak, uygulamada sonsuz birey yerine genellikle sınırlı sayıda bireylerin oluşturdukları topluluklar da populasyon olarak kabul edilmektedirler.

Bir araştırmacı Ziraat Fakültesi öğrencilerinin sorunları ile ilgili bir araştırma yapmaya karar vermiş ise fakültenin bütün öğrencileri araştırmacının populasyonunu oluşturur. Eğer hayvan yetiştirme alanında çalışan bir araştırmacı koyunların kesim ağırlığı ile ilgili bir araştırma yapıyorsa, seçtiği bir ağıldaki bütün koyunlar araştırmacının populasyonunu oluşturur.

Bir çalışma yapılırken üzerinde çalışılan tüm bireylerden veri toplamak çok zor hatta bazen olanaksızdır. Çünkü çoğu zaman araştırmacının zamanı ve maddi olanakları tüm bireylerden veri toplamak için yeterli değildir. Diğer taraftan üzerinde çalışılan özelliğin tipi de çoğu zaman tüm bireylerden veri toplanmasını engeller. Örneğin yukarıdaki örnekte araştırmacı fakültenin tüm öğrencilerinden veri toplayacak kadar zamana sahip olmayabilir. Bu durumda bütün öğrencilerden veri toplaması mümkün değildir. Diğer örnekte kesim ağırlığı ile ilgilenen araştırmacının ağıldaki bütün koyunları keserek kesim ağırlığı verisi toplaması söz konusu değildir.

Bu gibi sebeplerden dolayı popülasyonu oluşturan tüm bireylerden veri toplamak çoğunlukla mümkün olmadığı için araştırmacı, popülasyonu temsil edecek şekilde oluşturduğu daha az sayıdaki birey üzerinde çalışmak durumundadır. Popülasyonu temsil edecek şekilde seçilen belirli sayıdaki bireyden oluşan gruba **ÖRNEK** denir. Oluşturulan örneklerde belirli sayıda birey vardır. Bu belirli sayıdaki birey sayısına ise **ÖRNEK GENİŞLİĞİ** veya **ÖRNEK HACMİ** denir. Yukarıdaki örneklerde araştırma yapacak olan araştırmacıların üzerinde çalıştıkları popülasyonlardan, popülasyonu temsil edecek şekilde ve üzerinde çalıştıkları özelliği etkileyen faktörler bakımından homojen olan belirli sayıdaki bireyleri tamamen tesadüfen seçerek örneklerini oluşturması gerekir.

Eğer popülasyonu oluşturan tüm bireylerden veri toplanmış ise araştırmacı istediği değerleri tüm verileri kullanarak hesaplar. Popülasyondaki tüm bireylerden elde edilen verilerle hesaplanan değerlere **PARAMETRE** denir. Yani parametre, popülasyondan hesaplanan değerdir. Teorik olarak popülasyonlarda sonsuz sayıda birey olduğundan, bu popülasyonlarda her hangi bir değer hesaplanamaz. Ancak, popülasyonlara parametre değerleri atfedilebilir. Parametreler evrensel sabitler olup, hiçbir şekilde değişmezler.

Yapılan çalışmada oluşturulan örnekteki bireylerden veri toplanmış ve bu veriler kullanılarak istenilen değerler hesaplanmış ise bu hesaplanan değerlere **İSTATİSTİK** denir. Yani istatistik, örnekten hesaplanan değerdir. Örnekten hesaplanan değerler parametrelere karşılık gelir ve parametrelerin tahminleridir. Aynı popülasyondan iki farklı örnek alınmış olsa ve bu örneklerin her birinde aynı istatistik hesaplanırsa farklı farklı değerler elde edilebilir. Yani istatistik, örnekten örneğe değişebilmektedir.

Parametreler Yunan (Grek) harfleri, istatistikler ise Latin harfleri ile gösterilir. Parametre ve istatistiklerin gösterimi ile ilgili örnekler Tablo 1.1'de verilmiştir.

Tablo 1.1. Parametre ve istatistiklerin gösterimi ile ilgili örnekler

Hesaplanan değer	Popülasyon	Örnek
Ortalama	$\mu_x, \mu_y$	$\bar{X}, \bar{Y}$
Varyans	$\sigma_x^2, \sigma_y^2$	$S_x^2, S_y^2$
Korelasyon	$\rho, \rho_{xy}, \rho_{ab}$	$r, r_{xy}, r_{ab}$

Örnekten hesaplanan istatistiklerin güvenilir olması yani parametrelerin iyi tahminleri olması veya örneğin popülasyonu iyi temsil etmesi için aşağıdaki noktalara dikkat edilmesi gerekir:

1. Örnekteki bireylerin tamamen tesadüfen seçilmesi gerekir. Genel olarak popülasyona dahil her bireyin örneğe girme şansı aynı tutulmuşsa o örnek bir tesadüf örneğidir. Tesadüfi seçimin en kolay yolu kur'a ile seçimdir. Tesadüfi seçim, tesadüf sayıları kullanılarak da yapılabilir.

2. Örnekteki bireylerin üzerinde çalışılan özelliği etkileyen faktörler bakımından homojen (birbirine benzer) olması gerekir. Örneğin, eğer bir araştırmacı öğrencilerin ağırlıklarıyla ilgili bir araştırma yapıyorsa, homojenliği sağlamak bakımından örneğe seçilecek öğrencilerin aynı yaş veya yaş grubundan, aynı cinsiyetten ve sağlıklı öğrencilerden olmasına dikkat etmelidir. Çünkü bu faktörler biyolojik olarak ağırlığı etkileyen faktörlerdir. Bir başka araştırmacı koyunların canlı ağırlıkları ile ilgili bir araştırma yapıyorsa, bu durumda koyun sürüsünden örneğe seçilecek koyunların ağırlıkları üzerine etki eden ırk, cinsiyet, yaş ve benzeri faktörlerin homojen olmasına dikkat etmelidir.

3. Örnek genişliği arttıkça hesaplanan istatistiklerin güvenilirliği artar. Bu sebeple örnek genişliği mümkün olduğu kadar artırılmalıdır. Fakat örnek genişliği artırılırken hem bireylerin üzerinde durulan özellik veya özellikleri bakımından homojenliğinin bozulmamasına dikkate edilmeli hem de örnek genişliğinin artırılması ile sağlanacak faydanın masraf ve emeği karşılmasına dikkat edilmelidir.

### 1.3. Verilerin Toplanması

Yürütülen bir denemeden güvenilir sonuçların elde edilmesi için araştırmacının verilerini doğru olarak toplaması gerekir. Yanlış olarak toplanan verilerden doğru sonuçları hesaplayabilecek hiçbir yöntem yoktur. Bunun için araştırmacı, verilerini toplarken kullanacağı sayma, ölçme ve analiz yöntemlerini iyi bilmekle birlikte sabırlı ve dürüst olmalıdır. Sayma, ölçme ve analiz yöntemleri ile elde ettiği veriler üzerinde beklentisi doğrultusunda düzeltmeler yapmamalıdır. Doğru olarak kaydettiği verileri değerlendirerek sonuçlara ulaşmalı ve bu sonuçları yorumlamalıdır. Eğer beklediği sonuçlardan farklı değerler elde etmişse üzerinde çalıştığı konu ışığı altında yorumlar getirilmelidir.

### 1.4. Veri Tipleri

Bireylerden toplanan ve üzerinde durulan özelliği belirten rakamlara Gözlem Değeri (Veri) denir. Gözlem değerleri (veriler) 2 gruba ayrılır:

1. Nitel (Qualitative) gözlem değerleri
2. Nicel (Quantitative) gözlem değerleri

Nitel gözlem değerleri, sayısal olmayan verilerdir. Bu tip veriler genellikle anketler ve duyuşsal analizlerde toplanan verilerdir.

Nicel veriler sayısal verilerdir. Eđer veriler saymak sureti ile elde edilmişlerse Kesikli Gözlem Değeri, tartmak, ölçmek veya herhangi bir analiz metodu kullanılarak elde edilmişlerse Sürekli Gözlem Değeri denir. Sürekli veriler tarif aralığında her değeri alabilirler. Örneğin, bitkilerin ağırlıkları ve boyları, yumurta ağırlıkları, herhangi bir gıda maddesinin kurumadde miktarı, yağ oranı vb.

Kesikli veriler saymak sureti ile elde edilir. Örneğin, bir başaktaki dane sayısı, bir tavuğun yumurtladığı yumurta sayısı vb.

### 1.5. Ölçme Seviyeleri

Ölçme, nesne, birey veya olaylara sayısal değerlerin verilmesidir. Bir araştırmada üzerinde durulan özelliğe ait veriler farklı duyarlılık seviyelerinde ölçülebilir. Farklı ölçme seviyelerinde ölçülmüş verilerin değerlendirilmesi için farklı istatistik ve istatistik metotlar geliştirilmiştir. Elde edilen verilerden doğru ve güvenilir sonuçların elde edilebilmesi için istatistik metodun doğru seçilmesi gerekir. Bunun için de araştırmacının verilerin hangi seviyede ölçüldüğünü belirleyebilmesi gerekir.

Değişkenler ölçme seviyelerine göre 4 gruba ayrılırlar:

1. Nominal seviye (İsmlendirilmiş)
2. Ordinal seviye (Sıralanmış)
3. Interval seviye (Aralıklandırılmış)
4. Ratio seviye (Oranlanmış)

Ölçme seviyeleri ile istatistik metotların uygunluğu arasında bir ilişki vardır. Örneğin, nominal ölçümler için ortalamanın hesaplanması bir anlam ifade etmediği gibi hiçbir bilgi de vermez. Buna karşın ortalamayı içeren istatistik metotların ordinal veriler için kullanımı tartışmalıdır. Ortalama ve ortalamaya içeren istatistik yöntemlerin kullanılması ve elde edilen sonuçların güvenilir olması için verilerin interval ve ratio seviyelerde ölçülmüş veriler olması gerekir. Bu sebeple veriler ordinal seviyede elde edilmiş veriler ise en uygun merkezi eğilim ölçüsü ortanca (medyan) olacaktır.

**1. Nominal Seviye (İsmlendirilmiş, Kategorik):** En az duyarlı, en kaba ölçme seviyesi nominal seviyedir. Burada nominal “isim ile” anlamında kullanılmaktadır. Nominal seviye değişkenlere cinsiyet (erkek/kadın), etnik gruplar (Afrikalı, Türk, İngiliz vb.), din (Müslüman, Hıristiyan vb.) vb. örnek verilebilir. Bu seviyede kategoriler sadece isimle belirlenmiştir. Araştırmacı, veri toplama aşamasında gerekli olduğu zaman isim vererek yeni kategoriler oluşturabilir. Burada kategorilerin çoktan aza, yüksekte alçağa, iyiden kötüye gibi bir sırası yoktur. Eğer oluşturulacak kategoriler arasında böyle bir sıralama söz konusu değilse, bu değişkenler Nominal Seviye değişkenler olarak kabul edilir.

**2. Ordinal Seviye (Sıralı veya Sıralanmış):** Eğer üzerinde durulan değişkene ait değerler sıralanabilir kategorilerden oluşuyorsa, söz konusu değişkene Ordinal Seviye Değişken denir. Burada ordinal sıralı anlamındadır. Örneğin bir sınav sonucunda öğrenciler A, B, C gibi notlar alabilir. Bu notlar öğrencilerin sınavdaki yeterliliklerine göre verilir ve A alan öğrencinin B alan öğrenciden daha çalışkan olduğu alınan nottan anlaşılır. Diğer bir örnek olarak at yarışlarını verebiliriz. At yarışlarında birinci bitiren at ikinci bitiren attan daha hızlıdır, ikinci bitiren at üçüncü bitiren attan daha hızlıdır. Bu tip verileri nominal verilerden ayıran özellik anlamlı bir sıralamanın olmasıdır. Nominal veriler sadece aynı olup olmadıklarına göre sınıflandırılırlar. Ordinal veriler de ise aynı gruba dahil olup olmadıklarının, yani birbirinin aynı olup olmadıklarının yanı sıra birbirinden daha hızlı, daha iyi, daha çalışkan, daha üstün olup olmadıklarına göre de anlamlı bir şekilde sıralanırlar. Bu tip verilerde sınıflar bireylerin üzerinde durulan özellik bakımından durumlarına göre oluşturulur. Fakat aynı sınıfta yer alan bireyler arasında hiçbir farklılık yoktur. At yarışlarında olduğu gibi bireyler sayılar verilerek sıralandırılabilir. Bu tip veriler sıralandırılmış (rank) veriler olarak adlandırılır. Bu şekilde sıralandırılmış verilere duyusal (organoleptik) analizlerde çok rastlanır. Burada önemli olan ve bilinmesi gereken nokta şudur: Bireyler 1, 2, 3 şeklinde numaralandırılmış ise bunun anlamı 2, 1 den iki kat iyi, kötü, hızlı, çalışkan vb. demek değildir. Bu şekilde sıralandırıldığı zaman 1 ile 2 arasındaki farklılık ile 2 ile 3 arasındaki farklılık aynı demek değildir.

**3. Interval (Eşit birim aralıklarında ölçülmüş) Seviye:** Yapılan bir araştırmada veriler elde edildiği zaman birbirinden ayrılabiliriyorsa, sıralanabiliriyorsa ve aralarındaki aritmetik farklılık anlamlı ise bu tip veriler aralık (interval) seviye verilerdir. Bu tip verilerde hem bireylerin birinin diğerlerine göre ağırlığı veya hafifliği, azlığı veya çokluğu vb. bilinir hem de birbirlerinden ne kadar az veya çok, hafif veya ağır vb. olduğu da bilinir. Bu tip veriler yukarıda açıklanan iki tip veriden de daha açıklayıcı, daha çok bilgi vericidir. Bireyler arasında belirlenen aralık, ölçümlerin sabit birimlerini ve duruma bağlı olarak sıfır (orijin) noktasına sahip olduğunu gösterir. Bu tip veriler gerçek anlama sahiptirler. Aralık

seviyesinde ölçülen veriler için en klasik örnek sıcaklık ölçümleridir. Sıcaklık ölçümlerinde yaygın olarak kullanılan birim derecedir. Fakat orijin Fahrenheit veya Celsius (santigrat) olarak ölçülmesine bağlı olarak değişir. Eğer sıcaklık santigrat olarak ölçülmüş ise suyun donma noktası 0°C'dir. Suyun donma noktası eğer Fahrenheit olarak ölçülmüş ise 32°F'dir. Verilerin birimine bağlı olarak orijin değişse de ölçülen veriler arasındaki farklılık aynıdır. İnterval seviyenin en önemli özelliği ölçeklendirmede birimler arası bilinen miktarlardır ve aralıklar birbirine eşittir.

4. Ratio Seviye (Oransal): Veri seviyelerinin en yüksek oran (ratio) seviyesidir. İnterval seviyeye benzer. Fakat ölçülen özelliğin olmadığı seviyeyi gösteren gerçek bir sıfır değeri içerir. Sıcaklık (Kelvin), para, yükseklik, ağırlık, yaş, reaksiyon zamanı, tepki vb.

Bu tip verilerde sadece sırasal özellik, ölçümlerin birimi ve aralarında anlamlı aritmetik farklılıklar değil aynı zamanda sabit bir orijin veya orijin olarak seçilmiş bir sıfır noktası vardır. Sıfır noktası sabit orijin olarak alınsa da yükseklik, ağırlık ve uzunluk oran seviyesi ölçümlerdir. Örneğin, sıfır orijin olarak seçilmiş ise 160 kg ağırlık 80 kg'ın tam iki katıdır.

Birbirinden ayrılması zor olan interval ve ratio seviyeleridir. Örneğin, iki ağırlık ölçüldüğü zaman, ağırlıkların biri 40 g diğeri 80 g ise, 80 g 40 g'ın tam iki katıdır. Bu sebeple ağırlık ratio seviye bir değişkendir. Buna karşın iki sıcaklıktan biri 40° diğeri 80° ise 80°, 40°'nin iki katı sıcaklıkta değildir. Bu sebeple de sıcaklık interval seviye bir değişkendir. Ratio seviye değişkenlerde sıfır değeri bir miktarın olmadığını gösterir. Bir ağırlık sıfır ölçülmüşse ağırlık yok demektir. Bu sebeple ağırlık ratio değişkendir. Buna karşın sıcaklık 0° ölçülmüş ise bu sıcaklığın olmadığı anlamına gelmez ve bir interval değişkendir.

## 1.6. Verilerin Özetlenmesi

Araştırmacı üzerinde çalıştığı populasyondan örneğini oluşturarak verilerini topladıktan sonra toplanan verilerin özetlenmesi gerekir. Çünkü veriler toplanırken her hangi bir sıra veya düzene göre toplanamazlar. Tamamen gelişigüzel bir sırada elde edilirler. Mesela bir ağıldaki koyunların önce en ağır olanı tartılamaz. Eğer tartmadan en ağır olan bilinseydi zaten tartmaya gerek kalmazdı. Keza kanındaki hemoglobinin miktarı en düşük olan hastalardan kan örneği önceden, daha yüksek olanlarda daha sonra alınmaz. Dolayısıyla bireylerden gelişigüzel olarak her hangi bir sıra ve düzen gözetilmeden karmaşık bir sıra ile elde edilen veriler, özetlendiği zaman okuyucu bu özetlenmiş verilerden istediği birçok bilgiyi kolayca edinebilecektir. Yapılan bir çalışmadan elde edilen verilerin özetlenmesi frekans dağılım tabloları düzenleyerek veya bu frekans tablolarını grafikler halinde çizerek yapılır.

### 1.6.1. Frekans Dağılım Tablolarının Oluşturulması

Örnekten elde edilen gözlem değerlerinden aynı veya birbirine benzer olanlarını bir sınıfa dahil ederek oluşturulan tablolara frekans dağılım tablosu denir. Bir frekans dağılım tablosu oluşturulurken izlenmesi gereken işlem basamakları aşağıdaki gibi sıralanmıştır:

1. Frekans dağılım tablosu oluşturmanın ilk adımı tablodaki sınıf sayısının belirlenmesidir. Frekans dağılım tablosunda genellikle 8-10 sınıf yapılır. Sınıf sayısının 8'den az, 15'den çok olması istenmez. Çünkü sınıf sayısı 8'den az veya 15'den çok olduğu zaman özetlemenin amacından uzaklaşmış olur ve bilgi kayıplarına yol açar.

Sınıf sayısı (1.1) numaralı eşitlikte verilen Sturges formülü kullanılarak da hesaplanabilir.

$$\text{Sınıf sayısı} = 1 + \log(n) \quad \dots(1.1)$$

(1.1) numaralı eşitlikte  $n$ , örnek genişliğidir.

2. Sınıf sayısı belirlendikten sonra (1.2) numaralı eşitlik kullanılarak sınıf aralığının hesaplanması gerekir.

$$\text{Sınıf aralığı (c)} = \frac{\text{Değişim genişliği}}{\text{Sınıf sayısı}} \quad \dots(1.2)$$

(1.2) numaralı eşitlikte değişim genişliği bir değişim ölçüsü olup (1.3) numaralı eşitlik kullanılarak hesaplanır.

$$\text{Değişim genişliği} = \text{En büyük gözlem değeri} - \text{en küçük gözlem değeri} \dots(1.3)$$

3. Sınıf sayısı belirlenip sınıf aralığı hesaplandıktan sonra frekans dağılım tablosundaki sınıflar oluşturulur. Her bir sınıfın alt ve üst sınırı vardır. Gözlem değerlerinin en küçüğü birinci sınıfın alt sınırı olarak alınır. İkinci sınıfın alt sınırı, birinci sınıfın alt sınırına sınıf aralığı eklenerek hesaplanır. Alt sınır belirlendikten sonra üst sınırın belirlenmesi gerekir. Üst sınır belirlenirken, bir sınıfın üst sınırı ile bir sonraki sınıfın alt sınırının birbiri ile çakışmamasına ve aralarında boşluk kalmamasına dikkat edilmelidir. Bunun için birinci sınıfın üst sınırı belirlenirken ikinci sınıfın alt sınırından gözlem değerlerinin ölçüm hassasiyeti kadar birim çıkarılır. Diğer sınıfların alt ve üst sınırları, elde edilen alt ve üst sınırlara sınıf aralığı eklenerek bulunur. En büyük gözlem değeri, bir sınıfın sınırları içinde yer alana kadar sınıfların oluşturulmasına devam edilir.

4. Frekans dağılım tablosundaki sınıfların alt ve üst sınırları hesaplandıktan sonra her bir sınıf için **sınıf değeri** hesaplanır. Sınıf değeri bir sınıfın alt sınırı ile üst sınırının orta noktası, yani alt ve üst sınırın ortalamasıdır. Dolayısı ile her hangi bir sınıf için hesaplanan sınıf değeri, bu sınıfı en iyi temsil eden değerdir. Çünkü her hangi bir sınıfta yer alan gözlem değerleri o sınıfın sınıf değeri etrafında toplanma eğilimi gösterir.

5. Frekans dağılım tablosundaki sınıflar düzenlenip sınıf değerleri hesaplandıktan sonra her bir sınıfın frekansı belirlenir. Frekans, bir sınıfın yoğunluğunu yani sınıftaki birey sayısını gösterir. Bunun için örnekteki gözlem değerlerinin, frekans dağılım tablosundaki sınıflardan hangisinin sınırları arasında yer aldığı kontrol edilir ve gözlem değerinin dahil olduğu sınıftaki birey sayısı yani sınıfın frekansı belirlenir. Sınıfların frekansları önce işaretle sonra da sayısal olarak gösterilir.

6. Her bir sınıfın frekansı mutlak olarak belirlendikten sonra “**Nispi (%)**” frekanslar hesaplanabilir.

7. Frekans dağılım tablosu oluşturulduktan sonra istenirse eklemeli frekanslar bulunabilir. Eklemeli frekanslar, “**...den daha az**” veya “**...den daha fazla**” eklemeli frekanslar olarak düzenlenebilir. Eklemeli frekanslar, belirli bir değerden daha büyük veya daha küçük olan gözlem değerlerinin sayısı hakkında bilgi verir. Mutlak eklemeli frekanslar hesaplandıktan sonra nispi eklemeli frekanslar da hesaplanabilir.

Eklemeli frekanslar düzenlenirken, sınıfların “**Alt Gerçek Sınırları**” kullanılır. **Gerçek sınır**, bir sınıfın üst sınırı ile bir sonraki sınıfın alt sınırının orta noktasıdır. Bir sınıfın **üst gerçek sınırı** bir sonraki sınıfın **alt gerçek sınır**ıdır. Bir sınıf üst gerçek sınırdan biter ve bu sınırdan bir sonraki sınıf başlar.

**1.6.1. Sürekli Gözlem Değerleri İçin Frekans Dağılım Tablosu****ÖRNEK 1:**

Yapılan bir araştırmada 80 adet mısır bitkisinin boyları cm olarak Tablo 1.2'deki gibi ölçülmüş olsun. Mısır bitkisinin boyları ölçülerek elde edilmiş sürekli gözlem değerleri olmasına rağmen kullanılan ölçü aletinin duyarlılığına bağlı olarak virgülden sonra duyarlılık belirtilmemiş ve tamsayı olarak kaydedilmiştir. Yani ölçümler 1 cm duyarlılık ile yapılmıştır. Bu durumda çalışma ile ilgili sonuçlar yazılırken araştırmacının elde ettiği bu verileri özetlemesi yani bir frekans dağılım tablosu oluşturması gerekir.

Tablo 1.2. 80 adet mısır bitkisinin boyları (cm)

88	89	103	85	60	69	100	73
107	84	79	100	84	55	93	91
70	61	78	81	85	90	68	96
99	87	102	66	74	85	96	92
64	83	93	69	85	73	82	68
111	107	76	88	93	76	98	65
77	102	64	58	96	74	82	81
78	102	50	110	87	89	82	78
63	97	109	62	73	66	76	88
80	100	69	81	98	87	80	73

Frekans dağılımı tablosu hazırlamanın ilk adımı sınıf sayısının belirlenmesidir. Araştırmacı frekans dağılım tablosunda 10 sınıf olmasına karar vermiş olsun. Sınıf sayısı belirlendikten sonra (1.2) ve (1.3) numaralı eşitlikler kullanılarak sınıf aralığı aşağıdaki şekilde hesaplanır.

$$\text{Değişim genişliği} = 111 - 50 = 61$$

$$\text{Sınıf aralığı} = \frac{61}{10} = 6.1 \cong 6$$

Sınıf aralığı hesaplandıktan sonra örneğimizdeki verilerin duyarlılığı oranında sınıf aralığının yuvarlaklaştırılması gerekir. Böylece örneğimizdeki veriler tamsayı olarak elde edildiklerinden sınıf aralığının da tam sayı olması sağlanır. Yani 6.1 olarak hesaplanan sınıf aralığı 6 olarak yuvarlaklaştırılır.

Sınıf sayısı ve sınıf aralığı belirlendikten sonra sınıflar oluşturulmaya başlanır. Birinci sınıfın alt sınırı en küçük gözlem değeri olup 50'dir. Buna sınıf aralığı eklenerek diğer sınıfların alt sınırı bulunur. Bu durumda ikinci sınıfın alt sınırı 56'dır. Birinci sınıfın üst sınırı, ikinci sınıfın alt sınırından 1 çıkarılarak 55 olarak hesaplanır. 56'dan 1 çıkarılmasının sebebi ölçme duyarlılığının 1 cm olmasından dolayı 56 cm'den önceki 55 cm'yi elde edebilmektir. Böylece alt ve üst sınırlar hem çakışmamış hem de aralarında tamsayı olarak boşluk kalmamıştır. Daha sonra en büyük gözlem değerini bir sınıfın sınırları içine dahil oluncaya kadar sınıflar Tablo 1.3'te görüldüğü gibi oluşturulur. Tablo 1.3'te görüldüğü gibi 10 sınıf olarak kararlaştırılmış olmasına rağmen 11 tane sınıf oluşturulmuştur. Bunun sebebi sınıf aralığının yuvarlaklaştırılmış olmasıdır ve bir hata değildir.

## ZZT201 BİYOMETRİ DERSİ I. HAFTA DERS NOTLARI

### BÖLÜM 1

Kaynak Kitap: TEMEL BİYOMETRİ (2013) Kocabaş Z., Özkan, MM ve Başpınar E.  
Ankara Üniversitesi Ziraat Fakültesi Yayın No:1606 Ders Kitabı 558

Sınıflar oluşturulduktan sonra birinci sınıf için sınıf değeri  $\frac{50+55}{2} = 52.5$  olarak hesaplanır. Bu

değere sınıf aralığı eklenerek diğer sınıfların da sınıf değeri bulunur.

Daha sonra elde edilen gözlem değerleri tek tek kontrol edilerek hangi sınıfa dahil olduğu belirlenir. Sınıfların frekansları önce işaret daha sonra da sayı ile gösterilir. Mutlak frekanslar bu şekilde hesaplandıktan sonra nispi (%) frekanslar hesaplanabilir. Örneğin birinci sınıfa ait % frekans;

80 bireyden 2 tanesi 1. sınıfta ise  
100 bireyden X tanesi 1. sınıftadır.

orantısından  $X = \frac{2 \times 100}{80} = 2.50$  olarak hesaplanır. Diğer sınıflara ait % frekanslarda aynı şekilde hesaplanır (Tablo 1.3). Bir frekans dağılım tablosunda mutlak frekansların toplamı örnek genişliğine, % frekansların toplamı ise 100'e eşittir.

Tablo 1.3. 80 mısır bitkisinin uzunluğu için düzenlenen frekans dağılım tablosu.

Sınıflar		Sınıf değeri	İşaretle frekans	Mutlak frekans	% frekans
Alt sınır	Üst sınır				
50	- 55	52.5	//	2	2.50
56	- 61	58.5	///	3	3.75
62	- 67	64.5	//// - //	7	8.75
68	- 73	70.5	//// - ////	10	12.50
74	- 79	76.5	//// - ////	10	12.50
80	- 85	82.5	//// - //// - ////	15	18.75
86	- 91	88.5	//// - ////	10	12.50
92	- 97	94.5	//// - ///	8	10.00
98	- 103	100.5	//// - ////	10	12.50
104	- 109	106.5	///	3	3.75
110	- 115	112.5	//	2	2.50
			Toplam	80	100.00

Frekans dağılım tablosu oluşturulduktan sonra istenirse eklemeli frekanslar bulunabilir. Eklemeli frekanslar, “...den daha az” veya “...den daha fazla” eklemeli frekanslar olarak düzenlenebilir. Eklemeli frekanslar, belirli bir değerden daha büyük veya daha küçük olan gözlem değerlerinin sayısı hakkında bilgi verir. Bunun için sınıfların gerçek sınırlarının bulunması gerekir. İkinci sınıfın alt gerçek sınırı, birinci sınıfın üst sınırı ile ikinci sınıfın alt sınırının orta noktasıdır ve;

2. sınıfın alt gerçek sınırı =  $\frac{55+56}{2} = 55.5$  şeklinde hesaplanır. Gerçek sınırlarda herhangi bir

yuvarlaklaştırma işlemi yapılmaz. Bu değerden sınıf aralığı çıkarılarak 1. sınıfın alt gerçek sınırı ve buna sınıf aralığı eklenerek diğer sınıfların alt gerçek sınırları bulunur. Bu şekilde hesaplanan gerçek sınırlar Tablo 1.4’de frekans dağılım tablosunda verilmiştir.

Frekans dağılım tablosunda sınıflara ait alt gerçek sınırlar hesaplandıktan sonra “...den daha az” ve “...den daha fazla” eklemeli frekanslar düzenlenebilir. “...den daha az” frekansları şu şekilde düzenlenir: Birinci sınıfın alt gerçek sınırı olan 49.5’den daha az değere sahip gözlem değeri sayısı 0’dır.



## ZZT201 BİYOMETRİ DERSİ I. HAFTA DERS NOTLARI

### BÖLÜM 1

Kaynak Kitap: TEMEL BİYOMETRİ (2013) Kocabaş Z., Özkan, MM ve Başpınar E.  
Ankara Üniversitesi Ziraat Fakültesi Yayın No:1606 Ders Kitabı 558

İkinci sınıfın alt gerçek sınırı olan 55.5'den daha az değere sahip gözlem değeri sayısı 2, yani önceki sınıfın frekansı kadardır. Üçüncü sınıfın alt gerçek sınırı olan 61.5'den daha az değere sahip gözlem değerlerinin sayısı, bundan önceki sınıflarda bulunan frekansların toplamıdır ve 5'dir.

Tablo 1.4. 80 mısır bitkisinin uzunluğu için düzenlenen frekans dağılım tablosundaki sınıflara ait alt gerçek sınırlar

Sınıflar		Alt gerçek sınır	Mutlak frekans	% frekans
Alt sınır	Üst sınır			
50	- 55	49.5	2	2.50
56	- 61	55.5	3	3.75
62	- 67	61.5	7	8.75
68	- 73	67.5	10	12.50
74	- 79	73.5	10	12.50
80	- 85	79.5	15	18.75
86	- 91	85.5	10	12.50
92	- 97	91.5	8	10.00
98	- 103	97.5	10	12.50
104	- 109	103.5	3	3.75
110	- 115	109.5	2	2.50

Diğer sınıflar için de “...den daha az” eklemeli frekanslar aynı şekilde bulunur. “...den daha az” eklemeli frekansları bulunurken frekanslar yukarıdan aşağıya toplanır. “...den daha fazla” eklemeli frekansların düzenlenmesinde ise işlem, “...den daha az” eklemeli frekansları bulunurken uygulanan işlemin tam tersidir. Birinci sınıfın alt gerçek sınırı olan 49.5'den daha fazla değere sahip gözlem değeri sayısı 80'dir. İkinci sınıfın alt gerçek sınırı olan 55.5'den daha fazla değere sahip gözlem değeri sayısı 78, yani toplam frekanstan birinci sınıfın frekansı kadar eksiktir. Üçüncü sınıfın alt gerçek sınırı olan 61.5'den daha fazla değere sahip gözlem değerlerinin sayısı, 75, yani toplam frekanstan önceki sınıflardaki frekansların toplamı kadar azdır. Her sınıf için “---den daha az” ve “...den daha fazla” eklemeli frekanslarının toplamı örnek genişliğine, yani toplam frekansa eşittir. Eklemeli frekanslar düzenlendiği zaman sınıf sayısı frekans dağılım tablosundaki sınıf sayısından bir fazladır. Mutlak eklemeli frekanslar düzenlendikten sonra, bunlar toplam frekansa bölünüp 100 ile çarpılarak % eklemeli frekanslar bulunabilir. Bu şekilde düzenlenen eklemeli frekanslar Tablo 1.5'de verilmiştir.

Tablo 1.5. 80 mısır bitkisinin uzunluğu için düzenlenen frekans dağılım tablosundaki sınıflara ait alt gerçek sınırlar, "...den daha az" ve "...dan daha fazla" eklemli frekanslar

Alt gerçek sınır	Frekans	...den daha az		...den daha fazla	
		Mutlak frekans	% frekans	Mutlak frekans	% frekans
49.5	2	0	0.00	80	100.00
55.5	3	2	2.50	78	97.50
61.5	7	5	6.25	75	93.75
67.5	10	12	15.00	68	85.00
73.5	10	22	27.50	58	72.50
79.5	15	32	40.00	48	60.00
85.5	10	47	58.75	33	41.25
91.5	8	57	71.25	23	28.75
97.5	10	65	81.25	15	18.75
103.5	3	75	93.75	5	6.25
109.5	2	78	97.50	2	2.50
115.5		80	100.00	0	0.00

**ÖRNEK 2:**

Aynı tarladan toplanan fasulye tanelerinden tesadüfen seçilen 60 adet fasulye tanesinin ağırlıkları g olarak Tablo 1.6'da verildiği gibi tartılmıştır. Toplanan verilerin okuyucuya daha çok bilgiyi kolayca sunabilmesi için özetlenmesi yani frekans dağılım tablosunun oluşturulması gerekir.

Tablo 1.6. 60 adet fasulye tanesinin gr olarak ağırlıkları

0.55	0.73	0.61	0.54	0.80	0.80	0.71	0.74	0.72	0.54
0.57	0.78	0.68	0.57	0.59	0.54	0.76	0.68	0.65	0.77
0.61	0.71	0.73	0.65	0.60	0.62	0.60	0.64	0.74	0.63
0.67	0.64	0.77	0.68	0.65	0.66	0.57	0.57	0.62	0.63
0.70	0.53	0.81	0.75	0.70	0.65	0.65	0.66	0.73	0.66
0.50	0.58	0.48	0.78	0.72	0.75	0.67	0.69	0.60	0.65

Frekans dağılımı tablosu hazırlamanın ilk adımı sınıf sayısının belirlenmesidir. Araştırmacı frekans dağılım tablosunda 8 sınıf olmasına karar vermiş olsun. Sınıf sayısı belirlendikten sonra (1.2) ve (1.3) numaralı eşitlikler kullanılarak sınıf aralığı aşağıdaki şekilde hesaplanır.

$$\text{Değişim genişliği} = 0.81 - 0.48 = 0.33$$

$$\text{Sınıf aralığı} = \frac{0.33}{8} = 0.04125 \cong 0.04$$

Sınıf aralığı hesaplandıktan sonra örneğimizdeki verilerin duyarlılığı oranında sınıf aralığının yuvarlaklaştırılması gerekir ki örneğimizdeki veriler 0.01 hassasiyetle elde edildiklerinden sınıf aralığının da aynı hassasiyette olması gerekir. Yani 0.04125 olarak hesaplanan sınıf aralığı 0.04 olarak yuvarlaklaştırılır.

Sınıf sayısı ve sınıf aralığı belirlendikten sonra sınıflar oluşturulmaya başlanır. Birinci sınıfın alt sınırı en küçük gözlem değeri olup 0.48'dir. Buna sınıf aralığı eklenerek diğer sınıfların alt sınırı bulunur. Bu durumda ikinci sınıfın alt sınırı 0.52'dir. Fasulye ağırlıkları 0.01 g duyarlılık ile tartıldığından, birinci sınıfın üst sınırı, ikinci sınıfın alt sınırından 0.01 çıkarılarak 0.51 olarak hesaplanır.

## ZZT201 BİYOMETRİ DERSİ I. HAFTA DERS NOTLARI

### BÖLÜM 1

Kaynak Kitap: TEMEL BİYOMETRİ (2013) Kocabaş Z., Özkan, MM ve Başpınar E.  
Ankara Üniversitesi Ziraat Fakültesi Yayın No:1606 Ders Kitabı 558

Böylece alt ve üst sınırlar; hem çakışmamış hem de aralarında boşluk kalmayacak şekilde belirlenmiş olur. Daha sonra en büyük gözlem değeri bir sınıfın sınırları içine dahil oluncaya kadar sınıflar Tablo 1.7’de görüldüğü gibi oluşturulur.

$$\text{Sınıflar oluşturulduktan sonra birinci sınıf için sınıf değeri } \frac{0.48 + 0.51}{2} = 0.495 \text{ olarak}$$

hesaplanır. Bu değere sınıf aralığı eklenerek diğer sınıfların da sınıf değeri bulunur.

Daha sonra elde edilen gözlem değerleri tek tek kontrol edilerek hangi sınıfa dahil olduğu belirlenir. Sınıfların frekansları önce işaret daha sonra da sayı ile gösterilir. Mutlak frekanslar bu şekilde hesaplandıktan sonra nispi (%) frekanslar hesaplanabilir. Örneğin birinci sınıfa ait % frekans;

$$\begin{array}{ll} 60 \text{ bireyden} & 2 \text{ tanesi 1. sınıfta ise} \\ 100 \text{ bireyden} & X \text{ tanesi 1. sınıftadır.} \end{array}$$

orantısından  $X = \frac{2 \times 100}{60} = 3.33$  olarak hesaplanır. Diğer sınıflara ait % frekanslarda aynı şekilde hesaplanır (Tablo 1.7). Bir frekans dağılım tablosunda mutlak frekansların toplamı örnek genişliğine, % frekansların toplamı ise 100’e eşittir.

Tablo 1.7. 60 fasulye tanesinin ağırlığı için düzenlenen frekans dağılım tablosu.

Sınıflar		Sınıf değeri	İşaretle frekans	Mutlak frekans	% frekans
Alt sınır	Üst sınır				
0.48	0.51	0.495	//	2	3.333
0.52	0.55	0.535	////	5	8.333
0.56	0.59	0.575	//// - /	6	10.000
0.60	0.63	0.615	//// - ////	9	15.000
0.64	0.67	0.655	//// - //// - ///	13	21.667
0.68	0.71	0.695	//// - ///	8	13.333
0.72	0.75	0.735	//// - ////	9	15.00
0.76	0.79	0.775	////	5	8.333
0.80	0.83	0.815	///	3	5.000
			Toplam	60	100.000

Frekans dağılım tablosu oluşturulduktan sonra istenirse eklemeli frekanslar bulunabilir. Eklemeli frekanslar, “...den daha az” veya “...den daha fazla” eklemeli frekanslar olarak düzenlenebilir. Eklemeli frekanslar, belirli bir değerden daha büyük veya daha küçük olan gözlem değerlerinin sayısı hakkında bilgi verir. Bunun için sınıfların alt gerçek sınırlarının bulunması gerekir. İkinci sınıfın alt gerçek sınırı, birinci sınıfın üst sınırı ile ikinci sınıfın alt sınırının orta noktasıdır ve;

$$2. \text{ sınıfın alt gerçek sınırı} = \frac{0.51 + 0.52}{2} = 0.515 \text{ şeklinde hesaplanır. Bu değerden sınıf aralığı}$$

çıkarılarak 1. sınıfın alt gerçek sınırı ve buna sınıf aralığı eklenerek diğer sınıfların alt gerçek sınırları bulunur. Bu şekilde hesaplanan gerçek sınırlar Tablo 1.8’de frekans dağılım tablosunda verilmiştir.

Frekans dağılım tablosunda sınıflara ait alt gerçek sınırlar hesaplandıktan sonra “...den daha az” ve “...den daha fazla” eklemeli frekanslar düzenlenebilir. “...den daha az” frekansları şu şekilde

## ZZT201 BİYOMETRİ DERSİ I. HAFTA DERS NOTLARI

### BÖLÜM 1

Kaynak Kitap: TEMEL BİYOMETRİ (2013) Kocabaş Z., Özkan, MM ve Başpınar E.  
Ankara Üniversitesi Ziraat Fakültesi Yayın No:1606 Ders Kitabı 558

düzenlenir: Birinci sınıfın alt gerçek sınırı olan 0.475'den daha az değere sahip gözlem değeri sayısı 0'dır. İkinci sınıfın alt gerçek sınırı olan 0.515'den daha az değere sahip gözlem değeri sayısı 2, yani önceki sınıfın frekansı kadardır. Üçüncü sınıfın alt gerçek sınırı olan 0.555'den daha az değere sahip gözlem değerlerinin sayısı, bundan önceki sınıflarda bulunan frekansların toplamıdır ve 7'dir.

Tablo 1.8. 60 fasulye tanesinin ağırlığı için düzenlenen frekans dağılım tablosundaki sınıflara ait alt gerçek sınırlar

Sınıflar		Alt gerçek sınır	Mutlak frekans	% frekans
Alt sınır	Üst sınır			
0.48	0.51	0.475	2	3.333
0.52	0.55	0.515	5	8.333
0.56	0.59	0.555	6	10.000
0.60	0.63	0.595	9	15.000
0.64	0.67	0.635	13	21.667
0.68	0.71	0.675	8	13.333
0.72	0.75	0.715	9	15.00
0.76	0.79	0.755	5	8.333
0.80	0.83	0.795	3	5.000

Diğer sınıflar için de "...den daha az" eklemeli frekanslar aynı şekilde bulunur. "...den daha az" eklemeli frekansları bulunurken frekanslar yukarıdan aşağıya toplanır. "...den daha fazla" eklemeli frekansların düzenlenmesinde ise işlem, "...den daha az" eklemeli frekansları bulunurken uygulanan işlemin tam tersidir. Birinci sınıfın alt gerçek sınırı olan 0.475'den daha fazla değere sahip gözlem değeri sayısı 60'dır. İkinci sınıfın alt gerçek sınırı olan 0.515'den daha fazla değere sahip gözlem değeri sayısı 58, yani toplam frekanstan birinci sınıfın frekansı kadar eksiktir. Üçüncü sınıfın alt gerçek sınırı olan 0.555'den daha fazla değere sahip gözlem değerlerinin sayısı, 53, yani toplam frekanstan önceki sınıflardaki frekansların toplamı kadar azdır. Her sınıf için "...den daha az" ve "...den daha fazla" eklemeli frekanslarının toplamı örnek genişliğine, yani toplam frekanssa eşittir. Eklemeli frekanslar düzenlendiği zaman sınıf sayısı frekans dağılım tablosundaki sınıf sayısından bir fazladır. Mutlak eklemeli frekanslar düzenlendikten sonra, bunlar toplam frekansla bölünüp 100 ile çarpılarak % eklemeli frekanslar bulunabilir. Bu şekilde düzenlenen eklemeli frekanslar Tablo 1.9'da verilmiştir.

Tablo 1.9. 60 fasulye tanesinin ağırlığı için düzenlenen frekans dağılım tablosundaki sınıflara ait alt gerçek sınırlar, "...den daha az" ve "...dan daha fazla" eklemli frekanslar

Alt gerçek sınır	Frekans	...den daha az		...den daha fazla	
		Mutlak frekans	% frekans	Mutlak frekans	% frekans
0.475	2	0	0.000	60	100.000
0.515	5	2	3.333	58	96.667
0.555	6	7	11.667	53	88.333
0.595	9	13	21.667	47	78.333
0.635	13	22	36.667	38	63.333
0.675	8	35	58.333	25	41.667
0.715	9	43	71.667	17	28.333
0.755	5	52	86.667	8	13.333
0.795	3	57	95.000	3	5.000
0.835		60	100.000	0	0.000

### 1.6.2. Kesikli Gözlem Değerleri İçin Frekans Dağılım Tabloları

Kesikli gözlem değerleri saymak sureti ile elde edilen verilerdir. Elde edilen gözlem değerlerin arasındaki farklılığa bağlı olarak kesikli veriler için frekans dağılım tablosu iki farklı şekilde oluşturulabilir.

#### ÖRNEK 1:

Bir tütün tarlasından tesadüfen seçilen 40 adet tütün bitkisinin yaprak sayıları Tablo 1.10'daki gibi bulunmuştur.

Tablo 1.10. 40 adet tütün bitkisinin yaprak sayıları

7	7	8	9	6	6	8	8	8	6
6	8	8	9	8	6	10	9	7	7
6	12	8	7	9	11	7	7	6	8
6	6	7	10	7	6	7	8	7	6

Araştırmacının tütün tarlasından tesadüfen seçtiği tütün bitkilerinin yaprak sayıları arasında büyük bir farklılık yoktur. En az yaprak sayısı 6 en çok ise 12'dir. Bu nedenle frekans dağılım tablosundaki sınıfların 6'dan 12'ye kadar belirlenmesi yeterlidir. Bu şekilde düzenlenen frekans dağılım tablosu Tablo 1.11'de verilmiştir.

Tablo 1.11. 40 adet tütün bitkisinin yaprak sayıları için düzenlenen frekans dağılım tablosu

Sınıflar (Yaprak sayıları)	İşaretle frekans	Mutlak frekans (Bitki sayısı)	% frekans
6	//// -//// - /	11	27.50
7	//// -//// - /	11	27.50
8	//// -////	10	25.00
9	////	4	10.00
10	//	2	5.00
11	/	1	2.50
12	/	1	2.50
	Toplam	40	100.00

**ÖRNEK 2:**

Bir buğday tarlasından seçilen 50 adet buğday bitkisinde başaktaki dane sayısı Tablo 1.12'deki gibi bulunmuştur.

Tablo 1.12. 50 adet buğday bitkisinde başaktaki dane sayısı

25	36	34	25	27	25	24	36	33	38
33	45	32	18	22	26	18	17	35	38
24	43	35	23	33	19	32	42	23	42
26	34	42	34	21	21	18	41	24	42
35	34	43	22	27	28	24	35	38	26

Burada elde edilen veriler de kesikli veriler olmasına karşın aralarındaki farklılık daha fazladır ve başak sayısı 17 ile 45 arasında değişmektedir. Bu durumda 17'den 45'e kadar sınıfların oluşturulması özetlemenin amacından uzaklaşılmasına sebep olur. Bu durumda frekans dağılım tablosunun sürekli gözlem değerleri için açıklandığı şekilde aşağıdaki gibi oluşturulması gerekir.

Frekans dağılımı tablosu hazırlamanın ilk adımı sınıf sayısının belirlenmesidir. Araştırmacı frekans dağılım tablosunda 8 sınıf olmasına karar vermiş olsun. Sınıf sayısı belirlendikten sonra (1.2) ve (1.3) numaralı eşitlikler kullanılarak sınıf aralığı aşağıdaki şekilde hesaplanır.

$$\text{Değişim genişliği} = 45 - 17 = 28$$

$$\text{Sınıf aralığı} = \frac{28}{8} = 3.5 \cong 4$$

Sınıf aralığı hesaplandıktan sonra örneğimizdeki verilerin duyarlılığı oranında sınıf aralığının yuvarlaklaştırılması gerekir. Örneğimizdeki veriler tamsayı olarak elde edildiklerinden sınıf aralığının da tam sayı olması gerekir. Yani 3.5 olarak hesaplanan sınıf aralığı 4 olarak yuvarlaklaştırılır.

Sınıf sayısı ve sınıf aralığı belirlendikten sonra sınıflar oluşturulmaya başlanır. Birinci sınıfın alt sınırı en küçük gözlem değeri olup 17'dir. Bu değere sınıf aralığı eklenerek diğer sınıfların alt sınırı bulunur. Bu durumda ikinci sınıfın alt sınırı 21'dir. Başaktaki dane sayısı sayılarak elde edilmiş tamsayıdır. Bu sebeple ikinci sınıfın alt sınırından bir çıkarılarak birinci sınıfın üst sınırı 20 olarak bulunur. Daha sonra en büyük gözlem değeri bir sınıfın sınırları içine dahil oluncaya kadar sınıflar Tablo 1.13'te görüldüğü gibi oluşturulmuştur.

Tablo 1.13. 50 adet buğday bitkisinde başaktaki dane sayısı için düzenlenen frekans dağılım tablosu.

Sınıflar		Sınıf değeri	İşaretle frekans	Mutlak frekans	% frekans
Alt sınır	Üst sınır				
17	20	18.5	////	5	10.00
21	24	22.5	//// - ////	10	20.00
25	28	26.5	//// - ///	9	18.00
29	32	30.5	//	2	4.00
33	36	34.5	//// - //// - ///	13	26.00
37	40	38.5	///	3	6.00
41	44	42.5	//// - //	7	14.00
45	48	46.5	/	1	2.00
			Toplam	50	100.00

Sınıflar oluşturulduktan sonra birinci sınıf için sınıf değeri  $\frac{17+20}{2} = 18.5$  olarak hesaplanır. Bu değere sınıf aralığı eklenerek diğer sınıfların da sınıf değeri bulunur.

Daha sonra elde edilen gözlem değerleri tek tek kontrol edilerek hangi sınıfa dahil olduğu belirlenir. Sınıfların frekansları önce işaret daha sonra da sayı ile gösterilir. Mutlak frekanslar bu şekilde hesaplandıktan sonra nispi (%) frekanslar hesaplanabilir (Tablo 1.13). Bir frekans dağılım tablosunda mutlak frekansların toplamı örnek genişliğine, % frekansların toplamı ise 100'e eşittir.

Frekans dağılım tablosu oluşturulduktan sonra istenirse eklemeli frekanslar bulunabilir. Eklemeli frekanslar, “...den daha az” veya “...den daha fazla” eklemeli frekanslar olarak düzenlenebilir. Eklemeli frekanslar, belirli bir değerden daha büyük veya daha küçük olan gözlem değerlerinin sayısı hakkında bilgi verir.

Bunun için sınıfların alt gerçek sınırlarının bulunması gerekir. İkinci sınıfın alt gerçek sınırı, birinci sınıfın üst sınırı ile ikinci sınıfın alt sınırının orta noktasıdır ve;

2. sınıfın alt gerçek sınırı =  $\frac{20+21}{2} = 20.5$  şeklinde hesaplanır. Bu değerden sınıf aralığı çıkarılarak 1.

sınıfın alt gerçek sınırı ve buna sınıf aralığı eklenerek diğer sınıfların alt gerçek sınırları bulunur. Bu şekilde hesaplanan gerçek sınırlar Tablo 1.14'de frekans dağılım tablosunda verilmiştir.

Frekans dağılım tablosunda sınıflara ait alt gerçek sınırlar hesaplandıktan sonra “...den daha az” ve “...den daha fazla” eklemeli frekanslar düzenlenebilir. “...den daha az” frekansları şu şekilde düzenlenir: Birinci sınıfın alt gerçek sınırı olan 16.5'den daha az değere sahip gözlem değeri sayısı 0'dır. İkinci sınıfın alt gerçek sınırı olan 20.5'den daha az değere sahip gözlem değeri sayısı 5, yani önceki sınıfın frekansı kadardır. Üçüncü sınıfın alt gerçek sınırı olan 24.5'den daha az değere sahip gözlem değerlerinin sayısı, bundan önceki sınıflarda bulunan frekansların toplamıdır ve 15'dir. Diğer sınıflar için de “...den daha az” eklemeli frekanslar aynı şekilde bulunur. “...den daha az” eklemeli frekansları bulunurken frekanslar yukarıdan aşağıya toplanır. “...den daha fazla” eklemeli frekansların düzenlenmesinde ise işlem, “...den daha az” eklemeli frekansları bulunurken uygulanan işlemin tam tersidir. Birinci sınıfın alt gerçek sınırı olan 16.5'den daha fazla değere sahip gözlem değeri sayısı 50'dir. İkinci sınıfın alt gerçek sınırı olan 20.5'den daha fazla değere sahip gözlem değeri sayısı 45, yani toplam frekanstan birinci sınıfın frekansı kadar eksiktir. Üçüncü sınıfın alt gerçek sınırı olan 24.5'den daha fazla değere sahip gözlem değerlerinin sayısı, 35, yani toplam frekanstan önceki sınıflardaki frekansların toplamı kadar azdır.

Tablo 1.14. 50 adet buğday bitkisinde başaktaki dane sayısı için düzenlenen frekans dağılım tablosundaki sınıflara ait alt gerçek sınırlar

Sınıflar		Alt gerçek sınır	Mutlak frekans	% frekans
Alt sınır	Üst sınır			
17	20	16.5	5	10.00
21	24	20.5	10	20.00
25	28	24.5	9	18.00
29	32	28.5	2	4.00
33	36	32.5	13	26.00
37	40	36.5	3	6.00
41	44	40.5	7	14.00
45	48	44.5	1	2.00

Her sınıf için “---den daha az” ve “...den daha fazla” eklemeli frekanslarının toplamı örnek genişliğine, yani toplam frekansa eşittir. Eklemeli frekanslar düzenlendiği zaman sınıf sayısı frekans dağılım tablosundaki sınıf sayısından bir fazladır. Mutlak eklemeli frekanslar düzenlendikten sonra, bunlar toplam frekansa bölünüp 100 ile çarpılarak % eklemeli frekanslar bulunabilir. Bu şekilde düzenlenen eklemeli frekanslar Tablo 1.15’de verilmiştir.

Tablo 1.15. 50 adet buğday bitkisinde başaktaki dane sayısı için düzenlenen frekans dağılım tablosundaki sınıflara ait alt gerçek sınırlar, “...den daha az” ve “...den daha fazla” eklemeli frekanslar

Alt gerçek sınır	Frekans	...den daha az		...den daha fazla	
		Mutlak frekans	% frekans	Mutlak frekans	% frekans
16.5	5	0	0.00	50	100.00
20.5	10	5	10.00	45	90.00
24.5	9	15	30.00	35	70.00
28.5	2	24	48.00	26	52.00
32.5	13	26	52.00	24	48.00
36.5	3	39	78.00	11	22.00
40.5	7	42	84.00	8	16.00
44.5	1	49	98.00	1	2.00
48,5		50	100.00	0	0.00

### 1.6.3. Kalitatif Sınıflı Frekans Dağılım Tabloları

Yapılan bir araştırmada araştırmacı bir bölgedeki 250 işletmeyi büyüklüklerine göre küçük, orta ve büyük işletmeler olarak sınıflandırmış olabilir. Bu durumda frekans dağılım tablosunda sınıfların sözel olarak ifade edilmesi, yani kalitatif sınıflı bir frekans dağılım tablosunun oluşturulması gerekir. 250 işletmenin büyüklüklerine göre küçük, orta ve büyük olarak bilgi toplayan araştırmacı, frekans dağılım tablosunu Tablo 1.16’daki gibi oluşturmalıdır. Tablo 1.16’da kalitatif sınıflı frekans dağılım tablosu görülmektedir.



TABLO 1.16. 250 işletmenin büyüklüklerine ait frekans dağılım tablosu

İşletme büyüklüğü	İşletme sayısı
Küçük	160
Orta	50
Büyük	40

#### 16.4. Açık Uçlu Frekans Dağılım Tabloları

Bazı durumlarda araştırmacının üzerinde durduğu özellik ile ilgili örnekten elde ettiği bir kısım verileri, diğer verilerin çoğunluğuna nazaran çok küçük veya çok büyük değerli olabilir. Bu veriler dikkate alınarak frekans dağılım tablosunun oluşturulması frekans olmayan sınıfların olmasına sebep olabilir. Bu gibi durumlarda açık uçlu frekans dağılım tabloları düzenlenir. Açık uçlu frekans dağılım tablolarının iki ucu açık olabileceği gibi tek ucu da açık olabilir. Örneğin bir araştırmacı uyguladığı anket sonucunda çiftçilik ile uğraşan 250 kişinin yaşlarını tespit etmiş olsun. Bu çiftçilerin yaş gruplarına göre dağılımı incelendiğinde çiftçilik ile uğraşanlar arasında, 22 yaşından küçük ve 59 yaşından büyük bireyler az olduğu için, 22 yaşından küçükler birleştirilerek bir sınıf, 59 yaşından büyükler de birleştirilerek bir sınıf içinde toplanmıştır. Bu şekilde elde edilen iki ucu açık frekans dağılım tablosu, Tablo 1.17'deki gibi oluşturulmuştur.

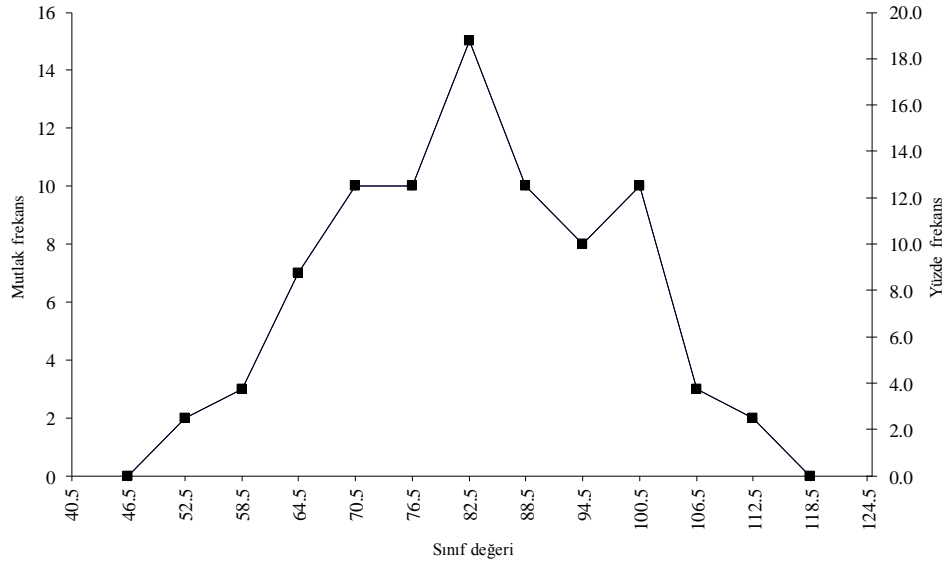
TABLO 1.17. 250 çiftçinin yaşlarına ait açık uçlu frekans dağılım tablosu

Yaş Grupları	Birey sayısı (f)
<25	9
25-29	47
30-34	55
35-39	59
40-44	22
45-49	19
50-54	17
55-59	15
>59	7

#### 1.7. Grafikler

Örnekten toplanan veriler için frekans dağılım tablosu oluşturulduktan sonra söz konusu tablonun görsel ifadesi olan grafikler çizilebilir. Frekans dağılım tablosu sürekli gözlem değerleri için düzenlenmişse grafik türü olarak "POLİGON" veya "HİSTOGRAM" çizilebilir. Düzenlenen frekans dağılım tablosu kesikli (tam) gözlem değerleri kullanılarak hazırlanmışsa bu durumda grafik türü olarak sadece "ÇUBUK DİYAGRAM" çizilebilir.

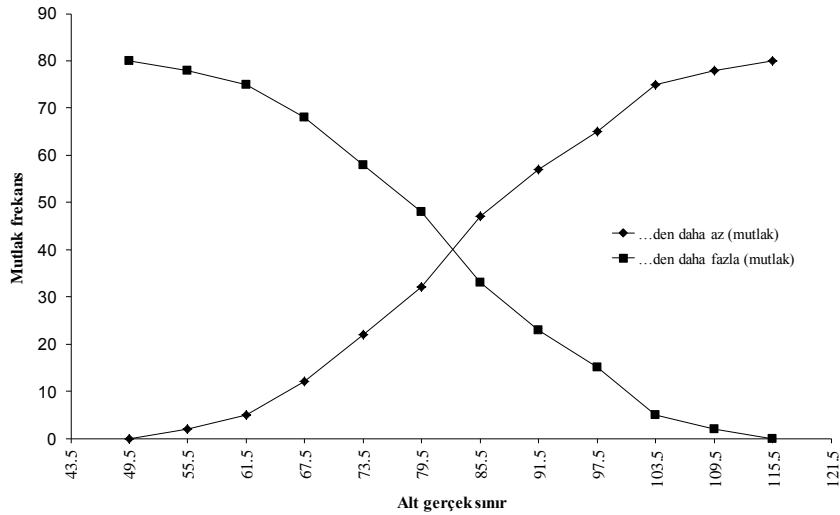
Sürekli veriler için oluşturulan bir frekans dağılım tablosundan poligon grafiği, yatay eksenine "sınıf değerleri" ve düşey eksenine de her sınıfa ait mutlak veya nispi (%) "frekans" koyularak çizilir. Poligon çizilirken, poligon, frekans dağılım tablosundaki birinci sınıftan önce olduğu varsayılan sınıfın sınıf değerinden başlayarak, en son sınıftan sonra olduğu varsayılan sınıfın sınıf değerinde biter. 80 mısır bitkisinin uzunluğu için düzenlenen ve Tablo 1.3 verilen frekans dağılım tablosundan çizilen frekans poligonu Şekil 1.1'deki gibidir.



Şekil 1.1. 80 mısır bitkisinin uzunluğu için düzenlenmiş frekans dağılım tablosundan çizilen mutlak ve yüzde frekans poligonu

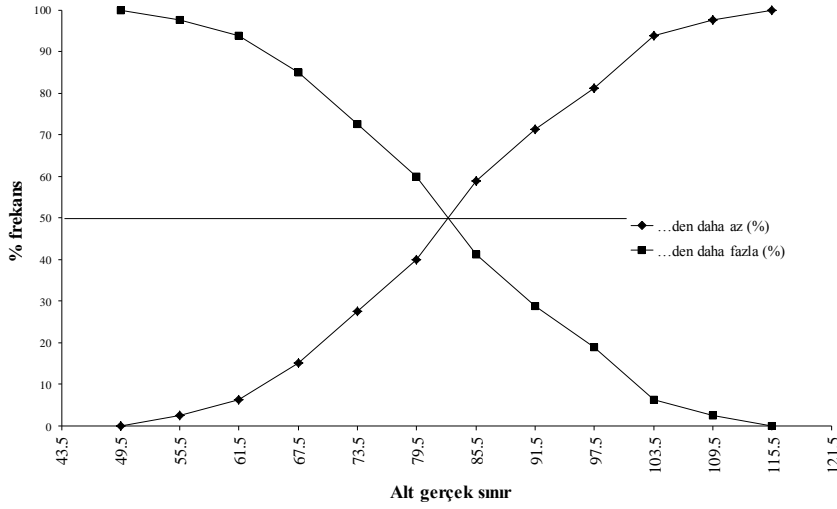
Eğer eklemeli frekanslar düzenlenmiş ise “...den daha az” veya “...den daha fazla” eklemeli frekans poligonları da çizilebilir. Eklemeli frekans poligonları, poligon grafiği yatay eksenine “gerçek sınırlar” ve düşey eksenine de her sınıfa ait mutlak veya nispi (%) “eklemeli frekans” koyularak çizilir. 80 mısır bitkisinin uzunluğu (Tablo 1.5) için düzenlenmiş mutlak ve yüzde eklemeli frekans poligonları sırasıyla Şekil 1.2 (a) ve Şekil 1.2 (b) deki gibidir. Şekil 1.2 (b)’de görüldüğü gibi, yüzde “...den daha az” ve “...den daha fazla” eklemeli frekans poligonları %50 noktasında birbirini keserler.

a.



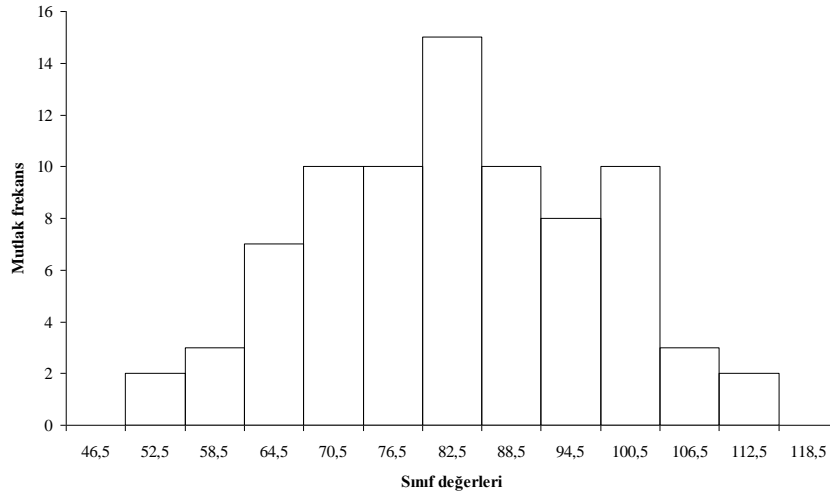
Şekil 1.2 (a) 80 mısır bitkisinin uzunluğu için düzenlenmiş frekans dağılım tablosundan çizilen mutlak eklemeli frekans poligonu

b.



Şekil 1.2 (b). 80 mısır bitkisinin uzunluğu için düzenlenen frekans dağılım tablosundan çizilen yüzde eklemeli frekans poligonu

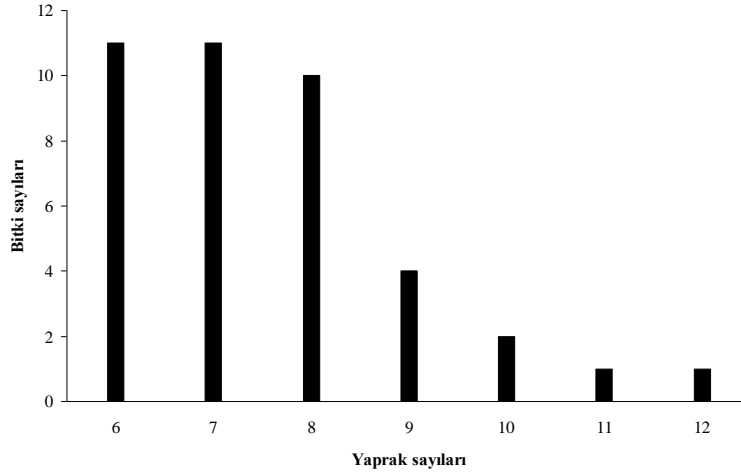
Sürekli veriler için düzenlenen bir frekans dağılım tablosu için grafik olarak histogram da çizilebilir. Histogram, yatay eksenine gerçek sınırlar ve dikey eksenine her sınıfa ait frekanslar koyularak, birbirini takip eden sınıfların gerçek sınırları üzerine frekanslarıyla orantılı olarak dikdörtgenler çizilerek oluşturulur. Histogramda çizilen dikdörtgenlerin orta noktaları sınıf değerlerine karşılık gelir. 80 mısır bitkisinin uzunluğu için düzenlenen ve Tablo 1.3'te verilen frekans dağılım tablosundan çizilen histogram Şekil 1.3'deki gibidir.



Şekil 1.3. 80 mısır bitkisinin uzunluğu için düzenlenen frekans dağılım tablosundan çizilen histogram

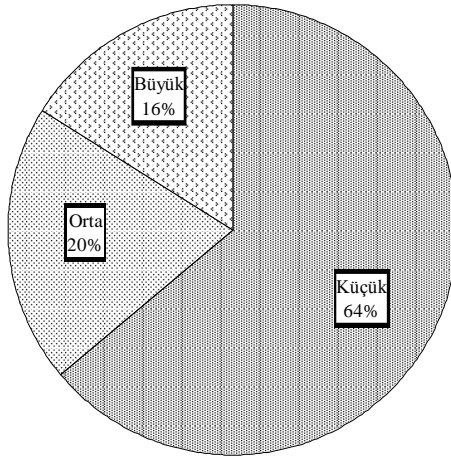
Eğer düzenlenen frekans dağılım tablosu tam sayılı gözlem değerleri içinse, bu frekans dağılım tabloları için grafik olarak çubuk diyagram çizilir. Çubuk diyagramın yatay eksenine sınıflar ve dikey eksenine sınıflara ait frekanslar konur. Çubuk diyagram çizildikten sonra ardışık tam sayılı gözlem değerleri söz konusu iki değer arasında herhangi bir değeri alamayacağından çubuklar çizgi ile

birleştirilmez. 40 adet tütün bitkisinin yaprak sayıları için düzenlenen ve Tablo 1.11’de verilen frekans dağılım tablosu için çizilen çubuk diyagram Şekil 1.4’deki gibidir.



Şekil 1.4. 40 adet tütün bitkisinin yaprak sayıları için düzenlenen frekans dağılım tablosundan çizilen çubuk diyagram

Kesikli ve yüzde ile ifade edilen veriler için uygun olan bir başka grafik türü de Pasta Dilimi grafiklerdir. Pasta dilimi şeklindeki grafikte bir bütünün kısımları sınıflandırılarak yüzde olarak pastanın dilimleri şeklinde gösterilir. 250 işletmenin büyüklüklerine göre bilgiyi küçük, orta ve büyük olarak toplayan araştırmacının Tablo 1.16’da oluşturduğu frekans dağılım tablosu için pasta dilimi grafik Şekil 1.5’deki gibidir.



Şekil 1.5. 250 işletmenin büyüklükleri için çizilen pasta dilimi grafik

**1.8. Sorular**

1. İstatistik ne ifade etmektedir. Açıklayınız.
2. Aşağıdaki kavramları örnek vererek açıklayınız.
  - a. Populasyon
  - b. Örnek
  - c. Parametre
  - d. İstatistik
3. Uygulamalı ve teorik istatistik arasındaki farkları belirterek açıklayınız.
4. Örnek genişliği ne demektir? Açıklayınız.
5. Bir araştırmacı neden bir örnek oluşturarak örnek üzerinde çalışma gereği duyar. Örnek vererek açıklayınız.
6. Bir araştırmacı üzerinde çalıştığı populyasyondan örnek alırken nelere dikkat etmelidir? Açıklayınız.
7. Üzerinde çalışılan örnekte elde edilen istatistiklerin güvenilir olması için nelere dikkat edilmelidir? Açıklayınız.
8. Gözlem değerleri (veriler) kaç gruba ayrılır? Örnek vererek açıklayınız.
9. Elde edilen verilerin ölçme seviyeleri kaç gruba ayrılır. Örnek vererek açıklayınız.
10. Elde edilen verilerin özetlenmesi ne şekillerde yapılabilir? Açıklayınız.
11. Frekans ne demektir? Örnek vererek açıklayınız.
12. Frekans dağılım tablosu ne demektir? Açıklayınız.
13. Bir frekans dağılım tablosu oluşturulurken hangi adımlar izlenir ve nelere dikkat edilmelidir? Açıklayınız.
14. Bir frekans dağılım tablosunda hesaplanan sınıf değeri neden önemlidir? Açıklayınız.
15. Eklemeli frekansların oluşturulmasına neden gerek duyulur? Eklemeli frekanslar nasıl oluşturulur. Açıklayınız.
16. Histogram ve poligon hangi tip veriler için çizilen grafiklerdir? Açıklayınız.
17. Belirli bir çeşide ait elmalardan tesadüfen alınan 50 adet elmanın ağırlıkları aşağıdaki gibidir. Bu veriler için;

118.7	118.0	100.9	109.9	107.5	110.0	114.7	110.1	112.2	116.4
88.7	108.7	106.4	104.0	107.9	105.5	112.1	103.6	120.2	100.4
98.5	111.4	107.4	105.4	104.0	108.7	102.1	107.1	121.9	104.6
103.6	95.4	113.5	106.8	117.3	113.7	97.2	99.4	113.3	119.8
108.2	101.7	103.6	108.0	94.0	108.1	115.9	120.8	112.9	110.2

- a. Frekans dağılım tablosunu oluşturunuz
  - b. Histogramını çiziniz
18. Belirli bir bölgede bulunan Karaçam ağaçlarından tesadüfen seçilen 30 adet Karaçamın göğüs çapları (cm) aşağıdaki gibi tespit edilmiştir. Bu verilere ait frekans dağılım tablosunu oluşturunuz.

21	15	37	38	41	35	24	19	27	23
27	36	57	53	45	56	60	59	20	29
47	40	27	29	31	30	33	24	27	36

19. Bir bahçeden tesadüfen alınan 40 adet elma meyvesindeki çekirdek sayıları aşağıdaki gibi saptanmıştır. Buna göre bu verilere ait frekans dağılım tablosunu oluşturarak uygun grafiği çiziniz

5	6	4	5	7	5	3	6	3	3
3	5	2	6	2	6	6	7	5	4
4	3	5	3	3	1	2	2	3	2
6	4	2	4	1	3	4	1	4	2

20. Örnek genişliğinin artması örnekten hesaplanan istatistikleri nasıl etkiler? Açıklayınız.

21. Aşağıdaki frekans dağılım tablosunda;

- Boş yerleri doldurunuz.
- Eklemeli frekansları bulunuz.
- Histogram grafiğini çiziniz.

Sınıflar	Sınıf Değeri	Frekans
41.5 - ...	...	3
43.5 - ...	...	5
... - ...	...	6
... - ...	...	8
... - ...	...	10
... - ...	...	7
... - ...	...	6
... - ...	...	5

22. Bir araştırmacı 75 adet sağlıklı Tuj koyununun ortalama kulak sayısını bulmak için nasıl bir yol izlemelidir. Açıklayınız.

23. Aşağıda tarif edilen olaylardan hangisi istatistiğin konusu olabilir?

- Sadece 6 yaşındaki çocukların alındığı bir okuldaki, 120 adet çocuğun yaş bakımından incelenmesi.
- Sadece elma ağaçlarının bulunduğu bir bahçede meyve türlerinin dağılımı.
- Sadece buğday ekilmiş bir tarladaki buğdayların başaklarındaki dane sayılarının dağılımı.
- Sadece 60 kg ağırlığına sahip 10 öğrencinin katıldığı bir yarışmada bu öğrencilerin ağırlık bakımından dağılımı.

24. 60 adet Ankara armudu tartıldıktan sonra en hafifinin 58.4 g, en ağırının da 68.4 g olduğu bulunmuştur. Bu armutlar için 10 sınıftan oluşan bir frekans dağılım tablosu yapılmıştır. Bu frekans dağılım tablosunda 3. sınıfın sınıf değeri kaçtır?

25. Aşağıdaki frekans dağılım tablosunda;

a. Üçüncü sınıfın, sınıf değeri kaçtır?

b. Dördüncü sınıfın gerçek alt sınırı kaçtır?

Sınıflar	Frekans
15.25 - .....	1
15.45 - .....	4
.... - .....	7
.... - .....	14
.... - .....	10
.... - .....	8
.... - .....	4
.... - .....	2

26. Frekans dağılım tablosunda gerçek sınırlar hangi amaçla kullanılır?

27. Açık uçlu frekans dağılım tabloları hangi durumlarda oluşturulur? Açıklayınız.