

ISL-0205

IKT-233

İSTATİSTİK

PROF. DR. FAZIL GÖKGÖZ

GİRİŞ

İstatistiksel bir çalışma yaparken sırasıyla aşağıdaki adımlar izlenir:

- Verinin (data) toplanması
- Analizlerin yapılması
- Analiz sonuçlarının yorumlanıp değerlendirilmesi

DATANIN TOPLANMASI

KAVRAMLAR

- **ÖRNEK:** İnceleme altındaki popülasyonlardan alınan bir parçayı temsil eder.
- **POPÜLASYON:** İnceleme yapılan ve hakkında bilgi toplanmaya çalışılan birimlerin toplamıdır.

İSTATİSTİK TÜRLERİ

- **TARİFSEL (descriptive) İSTATİSTİK:** Nümerik (sayısal) verileri düzenleyip özetlemek için kullanılan istatistik türüdür. Sayısal veriler üzerinden hareketle popülasyon için istatistiki sonuçlara ulaşılır.
- **TÜMEVARIMSAL (inferential) İSTATİSTİK:** Popülasyondan alınan örnekler incelenerek grubun tümü için istatistiksel sonuçlar elde edilir.

KULLANILAN DEĞİŞKEN TÜRLERİ

- KANTİTATİF (Quantitative) DEĞİŞKEN: Sayısal ölçeklerle ifade edilebilen ve miktar ile ilgili bilgileri veren değişkenlerdir.

öğrenci sayısı, hesaptaki bakiye, pilin ömrü vb.

1-süreli kantitatif değişken:

(boy uzunluğu, kargo ağırlığı vb.)

2-süreksiz kantitatif değişken:

(elaman sayısı, satılan ürün adedi vb.)

.....

- KUALİTATİF (qualitative) DEĞİŞKEN: Nümerik olmayan "görsel" değişkenlerdir.

göz rengi, doğum yeri, hisse başına kar düzeyi yüksek/düşük vb.

DATANIN TOPLANMASI

Data toplanmaya başlamadan önce;

- İncelenecek popülasyon **iyi ve açık bir şekilde belirlenmeli**
- Toplanacak **örneklerin popülasyonun tamamını net bir şekilde temsil etmesine** dikkat edilmeli (uygun örnekleme tekniği)

Tüm popülasyon (sayım) yerine neden örneklem üzerinden inceleme yapılır?

- Düşük maliyet sağlar.
- Zaman tasarrufu sağlar.
- Dikkatlice alınmış bir örneklem bazı durumlarda bir sayımdan daha iyi sonuçlar verebilir.
- Bazı durumlarda sayım imkansız olabilir.

ÖRNEKLEME PLANI

ÖZELLİKLERİ:

- **Yüksek kalite**; toplanan verilerin doğruluk ve popülasyona uyumluluğu ile ilişkilidir.
- **Savunabilirlik**; örneklem planını açıklayıcı ve savunucu belgelerin mevcut olması
- **Tekrarlanabilirlik**; örnek planı kullanılarak gerekli verilere tekrar ulaşılabilmesi
- **Temsil edici özelliği**;incelenen popülasyonu tamamen yansıtması
- **Faydalı olması**;toplanan verilerin planın uygulanmasına uygun ve elverişli olması

ÖRNEKLEME ŞEKİLLERİ

Örnekleme 3 farklı şekilde olabilir;



ÖRNEKLEME HATASI

- Örneklerin daimi ortak özelliđi “*tüm popülasyonu her yönüyle karıştılamamasıdır*”. Bu özellik örnekleme ile yapılan tahminlerin popülasyonun gerçek karakteri ile birebir uyuşmamasını sağlar.
- Buna göre örneklem hatası istatistiksel reel bir hata deđil örneklerin birbirine göre doğal deđişkenliklerinden kaynaklanan hataları temsil eder.

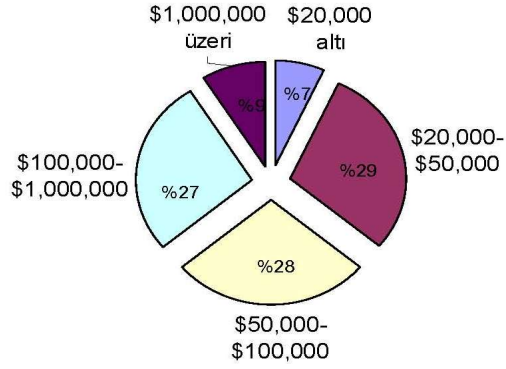
DATA GRAFİKLEME TÜRLERİ

Çalıřmalarda kullanılmak üzere toplanılan verilerin ve elde edilen sonuçların birbirleri ile daha rahat kıyaslanabilmeleri, analizlerin daha sistematik yapılabilmesi için bu bilgiler çeşitli grafikler üzerinde gösterilir. Data grafikleme türleri;

- 1.Pasta diyagramlar
- 2.Bar grafikler
- 3.Kartezyen(x-y) grafikleri
- 4.Frekans dağılım grafikleri
- 5.Histogramlar

1.PASTA DİYAGRAMLAR

Oran veya yüzde kullanılarak kuantitatif (farklı) datanın sunulmasını sağlayan grafiklerdir. Örneğin;



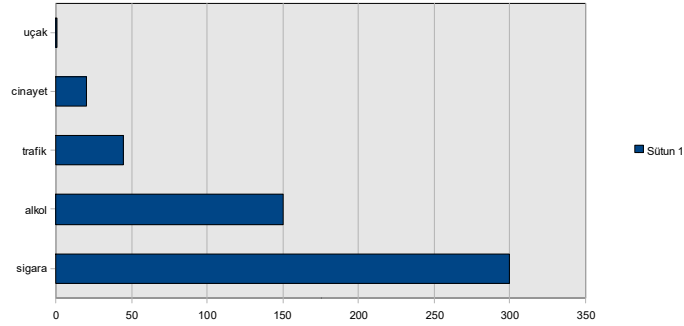
Prof. Dr. Fazıl GÖKGÖZ

13

2. BAR GRAFİKLERİ

Faktör ya da verilerin sıralanmasıyla oluşturulan grafiklerdir. Örneğin;

-yatay bar grafiği-



Prof. Dr. Fazıl GÖKGÖZ

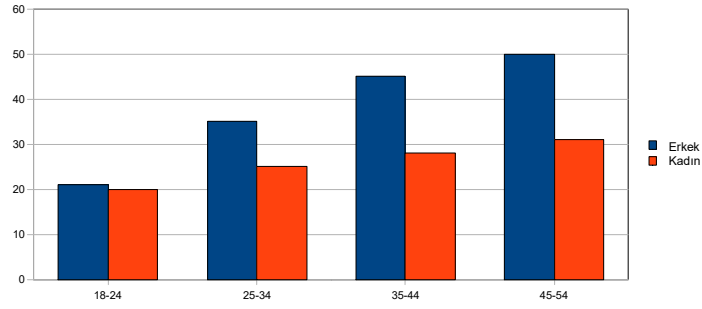
14

.....

-düşey bar grafiği-

x eksenini:yaş grupları

y eksenini:gelir dağılımı



Prof. Dr. Fazıl GÖKGÖZ

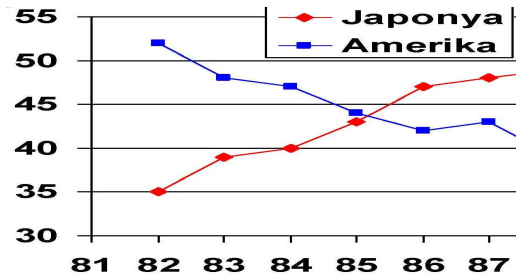
15

3. X-Y GRAFİKLERİ

Eksenler veri özelliklerini göstermek üzere bu özelliklerin birbirlerinin değişiminden nasıl etkilendiğini belirtirler. Örneğin;

x eksenini: yıllar

y eksenini: ülkelerin dünya marketindeki yüzdeleri



Prof. Dr. Fazıl GÖKGÖZ

16

VERİNİN ORGANİZE EDİLİP SUNULMASI

KAVRAMLAR

TASNİF: Özelliklere göre elde edilen grubun sınıflara ayrılması ile yapısal özelliğinin belirlenmesi işlemidir.

Belirlenen özelliklere göre kümeler oluşturulur ve aynı kümeyle ait özellik aynı sınıfta yer alır.

Sınırlı veri sayısının olduğu durumlarda mevcut yapıyı net bir şekilde ortaya koyabilen bir işlemdir.

.....

- Örneğin; 100 kişilik bir grubu yaş dağılımlarına göre tasnif edersek:

YAŞ	FREKANS
18 _____	21
19 _____	25
20 _____	30
21 _____	18
22 _____	6

TOPLAM:100

GRUPLAMA: Tasnif edilecek veri sayısının çok olduđu durumlarda başvurulanan bir yontemdir.

Bu Őekilde veri sayısı yuiksek duizeydeyken aynı vasfi taşıyan birbirine yakın ozellikteki veri aynı başlık altında toplanılarak gruplama yapılır.

Orneđin; duinyadaki en buyuk 29 Őehir nuifuslarına gure bir frekans dađılımına (veya gruplamaya) tabi tutulabilir.

.....

NÜFUS SINIFLARI

(1000 KİŐİ)

3000-4000'den az

4000-5000

5000-6000

6000-7000

7000-8000

8000-9000

9000 ve uzeri

ŐEHİR SAYISI

(FREKANS)

6

3

4

4

4

4

4

TOPLAM:29

GRUPLAMAYA HATALI BİR ÖRNEK;

Bir endüstri dalında faaliyet gösteren işletmelerde çalıştırılan işçi sayısına göre gruplamak istersek;

<u>ÇALIŞAN SAYISI</u>	<u>FREKANS</u>
1-2	315895
3-4	40588
5-9	9508
10-19	2348
20-49	721
50-59	44
100 ve üzeri	68

TOPLAM:369133

Prof. Dr. Fazıl GÖKGÖZ

21

FREKANS DAĞILIMLARI VEYA BÖLÜNMELEİ (Frequency Distributions)

KAVRAMLAR:

- SINIF ARALIĞI: Belirlenen sınıfın alt ve üst sınırları arasındaki farkı gösterir.
- SINIF SINIRLARI: O sınıfa ait maksimum (üst sınır) ve minimum (alt sınır) sınır değerleridir.
- SINIF ORTA NOKTASI : Sınıfın maksimum ve minimum sınır değerlerinin ortalamasıdır.

Prof. Dr. Fazıl GÖKGÖZ

22



O halde frekans bölümleri kavramını şöyle tanımlayabiliriz:

Verilerin her bir sınıf aralığı içindeki gözlem sayısını (frekans) gösterecek şekilde gruplandırılması işlemidir.

FREKANS DAĞILIMININ OLUŞTURULMASI

- Bir taşıtın yıl içerisindeki satış fiyatlarının düzenlenmemiş hali (Ham Veri-Raw Data)

\$20,197	\$20,372	\$17,454	\$20,591	\$23,651	\$24,453	\$14,266	\$15,021	\$25,683
27872	16587	20169	32851	16251	17047	21285	21324	21609
25670	12546	12935	16873	22251	22277	25034	21533	24443
16889	17004	14357	17155	16688	20657	23613	17895	17203
20765	22783	23661	29277	17642	18981	21052	22799	12794
15263	33625	14399	14968	17356	18442	18722	16331	19817
16766	17633	17962	19845	23285	24896	26076	29492	15890
18740	19374	21571	22449	25337	17642	20613	21220	27655
19442	14891	17818	23237	17445	18556	18639	21296	

Minimum fiyat

Maksimum fiyat

Sınıf Sayısının Belirlenmesi

Genel olarak frekans dağılımları oluşturulurken kullanılacak sınıf sayısı 5 ile 15 arasında olmalıdır. Eldeki toplam veri sayısı kullanılarak gerekli sınıf sayısını belirlemek mümkündür.

k:sınıf sayısı

n:toplam veri sayısı



$$2^k \geq n$$

eşitsizliği sağlanmalıdır.

Prof. Dr. Fazıl GÖKGÖZ

25

.....

Örnekte ham veri olarak bize verilen tabloya göre n=80 veri bulunmaktadır.

Yukarıda bize verilen eşitsizliği sağlayan minimum k değeri ise k=7 olarak tespit edilir.

O halde;

tavsiye edilen minimum sınıf sayısı k=7

Prof. Dr. Fazıl GÖKGÖZ

26

Sınıf Aralığının Belirlenmesi

- Sınıf aralığı seçerken yuvarlak rakamlar kullanılmalı
 - Birinci sınıfın alt limiti sınıf aralığının çift bir katı olmalı
 - Sınıf aralıkları birbirleri ile örtüşmemeli
 - Açık sınıf aralıklarından kaçınılmalı
- Tavsiye edilen sınıf aralığı= $\frac{\text{Max. değer} - \text{Min. değer}}{\text{Sınıf sayısı}}$

$$\longrightarrow \frac{33625 - 12546}{7} = \sim 3000 \text{ \$}$$

Frekans Dağılım Tablosu

<u>ARABA SATIŞ FİYATI (Bin \$)</u>	<u>FREKANS</u>
12-15	8
15-18	23
18-21	17
21-24	18
24-27	8
27-30	4
30-33	1
33-36	1
	<u>TOPLAM:80</u>

Frekans Dağılım Grafiği (Histogram)

- Frekans dağılım tablosundaki veriler yardımı ile aşağıdaki grafik oluşturulabilir.



Prof. Dr. Fazıl GÖKGÖZ

29

Frekans Dağılımlarının Oluşturulmasında Dikkat Edilecek Noktalar

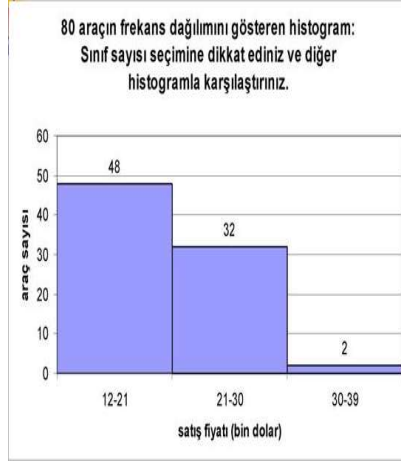
- Mümkün olduğunca eşit sınıf aralıkları seçilmelidir.
 - Çok sayıda boş sınıf oluşması vb. durumlarda eşit olmayan sınıf aralıklı frekans dağılımları oluşturulabilir.
 - Eşit olmayan sınıf aralıkları grafik aşamasında bazı sorunlar doğurabilir.
 - Uygun olarak seçilmemiş sınıf sayısına göre oluşturulmuş frekans dağılımları, verinin frekans dağılımı hakkında faydalı ve net bilgiler sunmayabilir. Örneğin;

<u>ARAÇ SATIŞ FİYATI</u> <u>(FREKANS)</u>	<u>ARAÇ SAYISI</u>
\$12000-21000	48
\$21000-30000	30
\$30000-39000	2
	<u>TOPLAM=80</u>

Prof. Dr. Fazıl GÖKGÖZ

30

••••••••



- Bir önceki histogramda sınıf sayısı 8 iken bu histogramda 3 sınıf kullanılmıştır. Buna karşılık sınıf aralıkları genişlemiş ve araçlar daha yüzeysel bir gruplamaya tabi tutulmuştur.

Bağıl (Nisbi) Frekans Dağılımları (Relative Frequency Distributions)

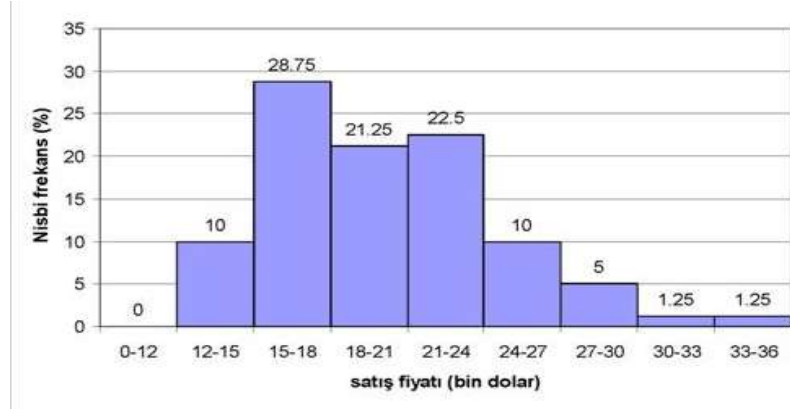
Sınıfın bağıl frekansı: o sınıfın frekansının toplam frekansa oranı ile tanımlanabilir.

Örneğin; \$12000-15000 sınıfının nisbi frekansı $8/80=0.1$ veya yüzde olarak ifadesi ile %10

Genellikle sınıfın mutlak frekansından çok nisbi frekansını bilip buna göre işlem yapmak daha gerekli olmaktadır.

.....

Nisbi (bağıl) frekans dağılım grafiği;



Kümülatif Frekans Dağılımları

- Farklı sınıf aralıklarında bulunan serilerin nisbi frekansları ile kıyaslanmalarını kolaylaştıran bir yöntemdir.
- Kümülatif frekans dağılımlarının en önemli özellikleri belirli bir düzeyin altında ve üstünde bulunan birimlerin frekansını gösterebilmeleridir.

Kümülatif Frekans Dağılımlarının Oluşturulması

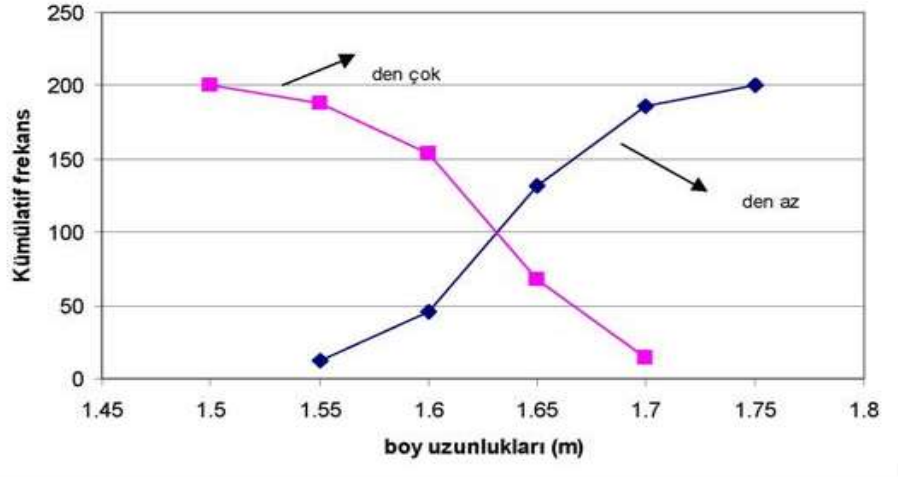
ÖRNEK: Bir sınıftaki öğrencilerin boy uzunluklarının frekans dağılımları aşağıda verilmiştir.

<u>BOY UZUNLUKLARI (CM)</u>	<u>FREKANS</u>
150-155	12
155-160	34
160-165	86
165-170	54
170-175	14
<u>TOPLAM:200</u>	

Kümülatif Frekans Dağılım Tablosu

<u>-den az</u>	<u>frekans</u>	<u>-den çok</u>	<u>frekans</u>
155	12	150	200
160	46	155	188
165	132	160	154
170	186	165	68
175	200	170	14

-den çok ve -den az kümülatif dağılım grafiklerinin oluşturulması



Prof. Dr. Fazıl GÖKGÖZ

37

FREKANS POLİGONLARI

- Sınıf orta noktasının sınıf frekansına göre grafiklendiği yapılardır.
- Frekans dağılımlarının sınıf sayısı ve aralığı eşit olması koşulu ile iki veya daha fazla frekans dağılım grafiğinin kıyaslanması aşamasında oldukça avantajlıdır.
- Bu özellik frekans poligonlarına histogramlara göre avantaj kazandırır.

Prof. Dr. Fazıl GÖKGÖZ

38

Öğrencilerin boy uzunluklarını gösteren frekans poligonu

