BATCH: 223

DT/NT : DATA SCIENCE

LESSON : STATISTICS -2

SUBJECT : PROBABILITY
PERMUTATION – COMBINATION
BAYES THEOREM

TECHPRO EDUCATION

techproeducation.com   +1 (585) 304 29 59

STATISTICS - 2

Data Science Program

Statistics Session -5

Session - 5 Content

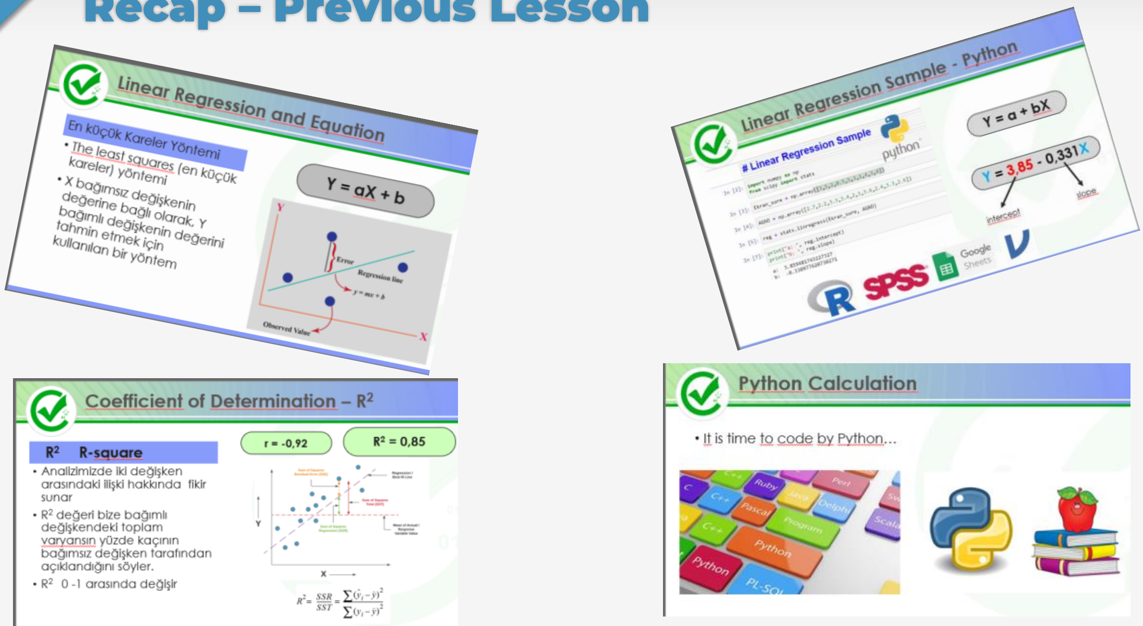
Content

- Probability – Olasılık – İhtimaliyet
- Permütasyon - Kombinasyon
- Koşullu olasılık (Conditional Probability)
- Bayes teoremi

RECAP

Herkes önceki dersten hatırladığı 1 cümle yazabilir mi?

Recap – Previous Lesson



TECHPRO EDUCATION

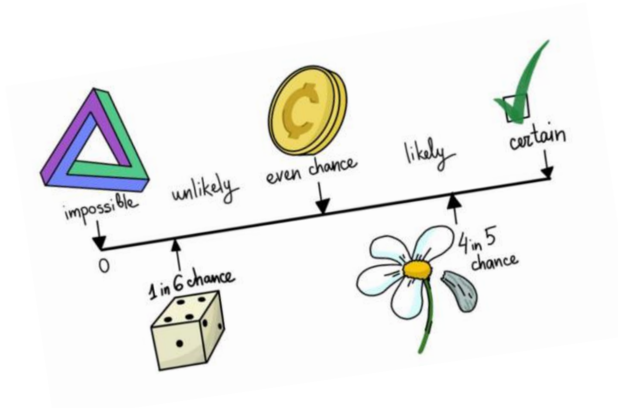
PROBABILITY

Olasılık – İhtimaliyet Prensibi

What is Probability

Olasılık Nedir ?

- Olasılık, 0 ile 1 arasında bir sayı olarak ifade edilen, belirli bir olayın gerçekleşme olasılığını hesaplayan bir istatistik dalıdır
- Rastgele Değişken
- Neden Probabilistik yaklaşım?
- Örneklem Uzayı



What is Probability

Olasılık Nedir ?

- Günlük hayatta olasılık
- Tahmin nasıl yapılır
- Rastgele Olay

A olayının gerçekleşme ihtimali :

P (A) = Olaydaki sonuç sayısı (n) / Toplam Sonuçlar sayısı (N)

Örnek, iskambil destesinde sinek olma ihtimali
P (Sinek) = 13/ 52 = 0,25

	M	T	W	TH	F	S	S
Chance of rainfall	70%	80%	90%	80%	60%	20%	0%

Probability and Statistical Experiment

Statistical Experiment

Eğer

- Bir deney birden fazla sonuç içeriyorsa
- Bu sonuçlar önceden tahmin edilebilir nitelikte ise
- Sonuçlar şansa dayalı ise burada istatistiksel deney şartları sağlanır.

Eğer 3 muhtemel sonuç var ise;
A, B, C gibi…
O zaman P(A) +P(B) + P(C)=1

Bir statistical experiment de olasılıklar toplamı 1'e eşit olur. Örneğin yazı – tura atma olayı da bu şartları sağlar.

Probability Computations

Olasılık Hesaplamaları

Örnek:

Bir torbada 5 kırmızı, 7 siyah ve 3 beyaz bilye bulunmaktadır. Bu torbadan rastgele çekilecek bir bilyenin kırmızı gelme olasılığı nedir?

P (A) = Olaydaki sonuç sayısı (n) / Toplam Sonuçlar sayısı (N)

Probability of a successful outcome (S)

$$P(S) = \frac{Number\ of\ successes}{Total\ number\ of\ outcomes} = \frac{r}{n}$$

$$\mathbf{P(A) = \frac{n}{N} = \frac{5}{5+7+3} = \frac{5}{15} = \frac{1}{3}}$$

Probability Computations

Olasılık Hesaplamaları - 2

Örnek -2 :

Bir önceki örnekteki bilgileri kullanarak;

a) Herhangi bir renkte bilye gelme olasılığını hesaplayınız.

b) Mavi renkte bilye gelme olasılığını hesaplayınız.

c) Siyah renkte bilye gelme olasılığını hesaplayınız

$$\text{a) } P(\text{bilye}) = \frac{5+7+3}{5+7+3} = \frac{15}{15} = 1$$

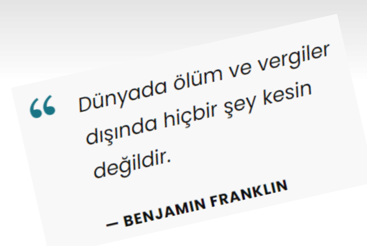
torbadan her durumda bir bilye çekileceğinden bu olay ortaya çıkması kesin bir olaydır

$$\text{b) } P(\text{mavi}) = \frac{0}{5+7+3} = \frac{0}{15} = 0$$

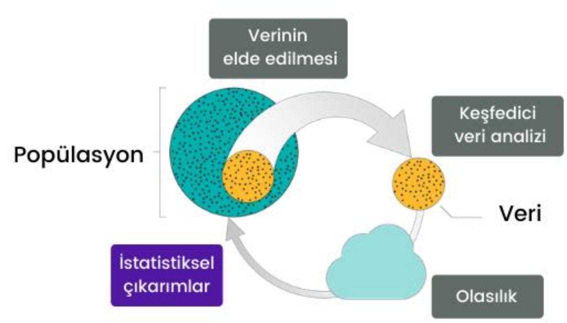
torbada hiç mavi bilye olmadığından, bu olay ortaya çıkması imkansız olan bir olaydır

$$\text{c) } P(\text{siyah}) = \frac{7}{5+7+3} = \frac{7}{15}$$

Olasılık Özet ..



- Belirsizlik kavramı
- Data Science için Olasılık neden gerekli ?



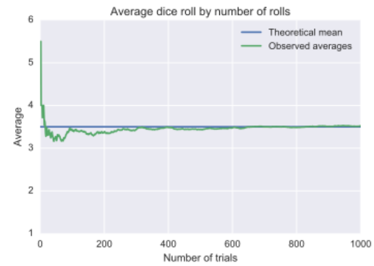
Law of Large Estimates

Bu kanunla alakalı örnek video

Büyük Sayılar Kanunu

- Bir deney tekrar tekrar tekrarlanırken, bir olayın ampirik olasılığı olayın teorik (gerçek) olasılığına yaklaşır
- Law of large numbers olarak da görülebilir

Relative frequency:
- RF = Number of times event occurs / Number of trials



Number of flips	Relative frequency	Percent
10	0.4 to 0.6	66
100	0.49 to 0.51	92
1,000	0.499 to 0.501	97
10,000	0.4999 to 0.5001	99

TECHPRO EDUCATION

INTERVIEW QUESTION

Q20. What Is the Law of Large Numbers?

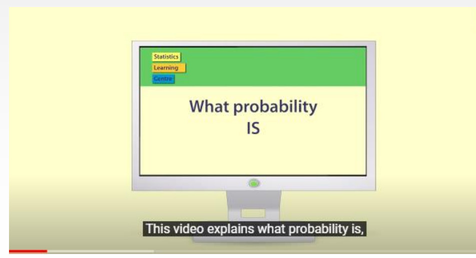
It is a theorem that describes the result of performing the same experiment a large number of times. This theorem forms the basis of **frequency-style** thinking. It says that the sample means, the sample variance and the sample standard deviation converge to what they are trying to estimate.

YOUTUBE VIDEO ONERİLERİ

- Understanding Probability - Probability 1
- Probability, Sample Spaces, and the Complement Rule (6.1)

https://www.youtube.com/watch?v=XHmIRCw5CLY

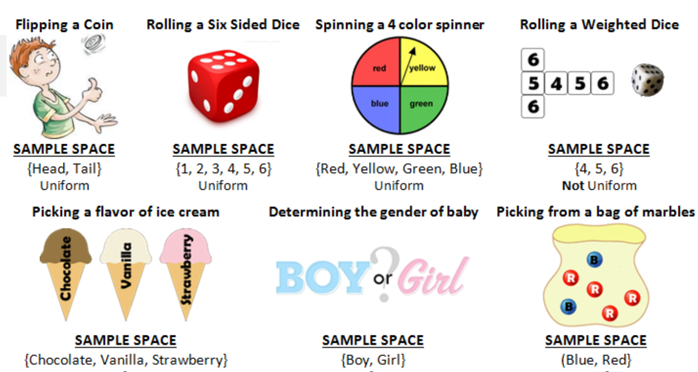
https://www.youtube.com/watch?v=ynjHKBCiGXY



Sample Space

Örnek Uzayı ve Küme

- İlgilenen rastgele olayın alabileceği tüm değerleri içeren uzaydır
- küme teorisi, rastgele olayların tanımlanması kolaylaştıran bir yaklaşımdır

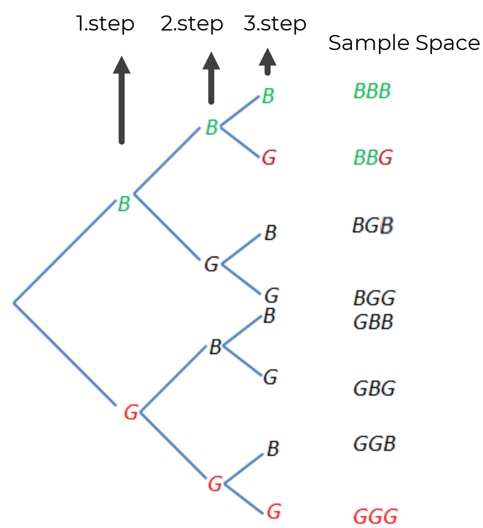


→ { YY, YT, TY, TT } → 2 kez üst üste atılan yazı-tura olayı ihtimalleri

Sample Space - Event

Örnek değerlendirme

- 3 çocuklu bir ailede sırayla 1., 2. ve 3. çocukların boy-girl olma ihtimallerini gösteren sample space
- Bu şekil Tree diagram olarak geçer
- Buradaki her output bir Event olur



Event'lar (Olaylar):

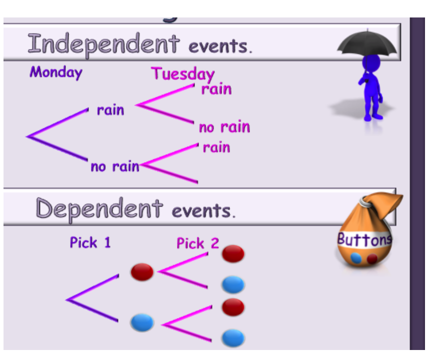
A = Tüm çocuklar boy = { BBB }

B = En az ikisi Girl olanlar = { BGG, GBG, GGB, GGG }

Independent – Dependent Event

Bağımlı / Bağımsız Olay

- Herhangi 2 olay Independent ve Dependent olaylar olarak sınıflandırılır.



A olayının (event) gerçekleşme ihtimali

P (A) = A olayına ait küme içindeki olay sayısı / Kümedeki Toplam olay sayısı

Independent:
Birininin olasılığı diğerini etkilemez

Dependent:
Bir olayın meydana gelmesi diğer bir olayı etkiliyor

Probability of 2 Independent Event

2 Bağımsız olayın Gerçekleşme İhtimali

- 2 olay bağımsız ise, her ikisinin de meydana gelme ihtimali bireysel olayların ihtimallerinin çarpımı olur. → P (A and B) = P (A) x P (B)

1st Question	2nd Question C	2nd Question I
C	0.58	0.05
I	0.11	0.26

A and B

Let A denote {first question correct}

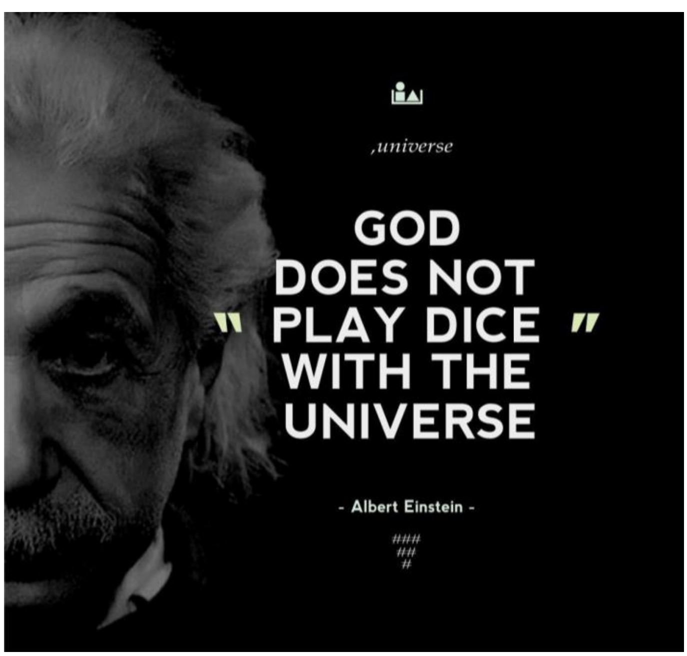
Let B denote {second question correct}

```
P(A) = P({CI, CC}) = 0.05 + 0.58 = 0.63
P(B) = P({IC, CC}) = 0.11 + 0.58 = 0.69
```

```
P(A and B)
P({CC}) = 0.58
```

If A and B were independent, then

```
P(A and B) = P(A) * P(B) = 0.63 * 0.69 = 0.43
```



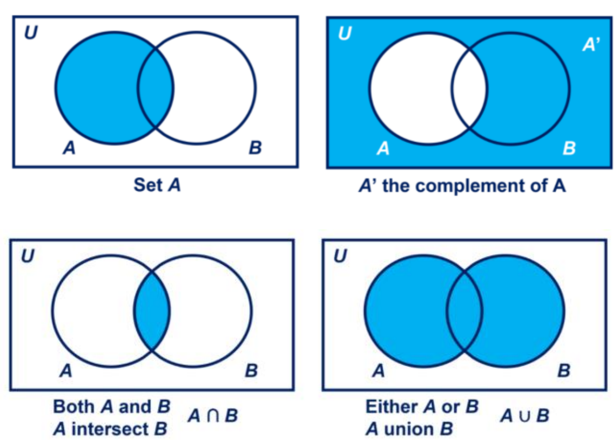
Union, Intersection, Complements

Kesişim, Birleşim, Tümleyen Kümeleri

Intersection, Union, Complement

Birleşim, Kesişim, Tümleyen Kümeleri

- **Union** : "A∪B" ile gösterilir ve A ve B kümesinin tüm öğelerini içeren kümedir. "veya" kullanılarak birleştirilmesine karşılık gelir.
- **Intersection**: A∩B ile sembolize edilir, A ve B kümesinde ortak olan tüm öğeleri içeren kümedir. "ve" kullanılarak birleştirilmesine karşılık gelir.
- **Comploments:** : Bir A kümesinin tümleyeni A' ile gösterilir ve evrensel kümedeki A'da olmayan tüm öğelerin kümesidir



Intersection, Union, Complement

Union Örnek

Set A = {1, 2, 3}
Set B = {3, 6, 8}
AUB = {1, 2, 3, 6, 8}

Intersection Örnek

Set A = {1, 2, 3}
Set B = {2, 4, 5}
A ∩ B = {2}

Complment Örnek

Set U = {1,2,3,4,5,6,7,8,9,10}
Set A = {1, 2, 3}
Set B = {3, 4, 5}
Set A' = {4,5,6,7,8,9,10}

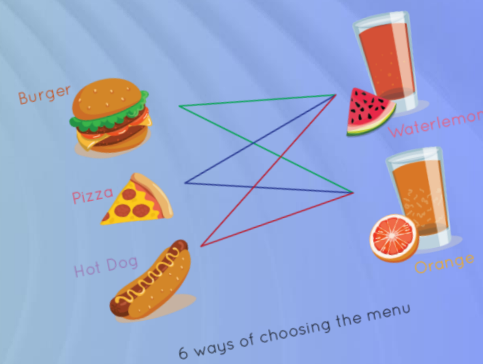
Complement

Complment – Tümleyen Örnek

E olayının tümleyeni, E olayına dahil edilmeyen örneklem uzayındaki tüm sonuçların kümesidir (E′ ile gösterilir).

$$P(E) + P(E') = 1 \qquad P(E) = 1 - P(E') \qquad P(E') = 1 - P(E)$$

Örnek:
Bir sepetin içinde 5 kırmızı pul, 4 mavi pul ve 6 beyaz pul olsun. Rasgele seçilen bir pulun mavi olmaması olasılığını bulun.
P (mavi pul seçilmesi) = 4 /15 = 0.267
P (mavi olmayan pul seçilmesi) = 1 – (4/15) = 0.733



Permutation, Combination

Permütasyon, Kombinasyon

Permutation

Permütasyon

- Permütasyon, nesnelerin belirli bir şekilde veya sırayla düzenlenmesidir.
- bir kümedeki öğelerin sırasını göz önünde bulundurarak düzenlenebileceği tüm olası yollar olarak tanımlanır

Order	1st	2nd
1	blue	pink
2	blue	green
3	blue	white
4	pink	blue
5	pink	green
6	pink	white
7	green	blue
8	green	pink
9	green	white
10	white	blue
11	white	pink
12	white	green

Permutation

Permütasyon

- n farklı nesnelerin farklı permütasyonlarının sayısı n! dir .
- İncelenen **n** bireyden her defasında **r** adedi alınarak, sıra gözetilmek kaydıyla, kaç farklı dizi oluşturulabileceği

Örnek:
Bir ankette 7 soru varsa, olası tüm soru düzenlemelerini
karşılamak için kaç farklı ankete ihtiyaç vardır?
$7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040$ anket

$$P(n,r) = \frac{n!}{(n-r)!}$$

Örnek:
8 listeden 5 kitap okumak için gereklidir. Ne kadar farklı sıralama yapabilirsiniz?

$${}_nP_r = {}_8P_5 = \frac{8!}{(8-5)!} = \frac{8!}{3!} = \frac{8\cdot 7\cdot 6\cdot 5\cdot 4\cdot \cancel{3}\cdot \cancel{2}\cdot \cancel{1}}{\cancel{3}\cdot \cancel{2}\cdot \cancel{1}} = 6720 \text{ yol}$$

Combination

Kombinasyon

- Kombinasyonlar neredeyse permütasyonlara benzer, ancak fark, elemanların sırasının önemli olmamasıdır
- Kombinasyon, sıralamanın önemli olmadığında **n** gruptan **r** nesnenin seçilmesidir

Order	1 st	2 nd
1	blue	pink
2	blue	green
3	blue	white
4	pink	green
5	pink	white
6	green	white

Combination

Kombinasyon

- **n** gruptan seçilen **r** nesnesi kombinasyonlarının sayısı →

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}$$

$$\underset{\text{grup sayı}}{{}_{n}}\underset{\text{Gruptan seçilen sayı}}{C_{r}} = \frac{n!}{(n-r)!r!}.$$

Örnek:
10 profesörün bulunduğu bir gruptan seçilecek 3 kişilik jürinin istenen şahıslardan meydana gelme olasılığı nedir?
Cevap
3 pozisyon için yapılacak seçimde sıra gözetilmeyeceğinden (yani oluşturulan bir XYZ jürisinde heyetinde X jüri üyesi, Y jüri üyesi, Z jüri üyesi iken, YXZ jürisinde de Y jüri üyesi, X jüri üyesi, Z jüri üyesidir) kombinasyon formülü kullanılır

$$10C3 = \frac{10!}{(10-3)!\,3!} = 120$$ farklı jüri oluşturulabilir.

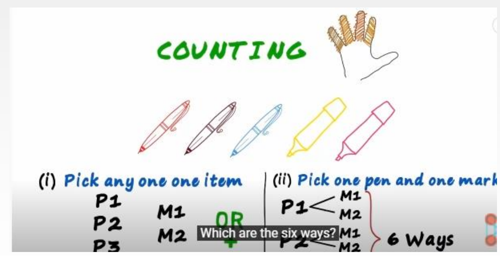
Peardeck Question

Hangisinde Order (Sıralama) önemlidir?

YOUTUBE VIDEO ONERI

https://www.youtube.com/watch?v=0NAASclUm4k

- Permutations and Combinations | Counting | Don't Memorise



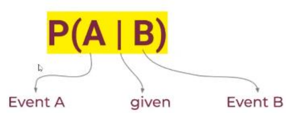
Conditional Probability

Koşullu Olasılık

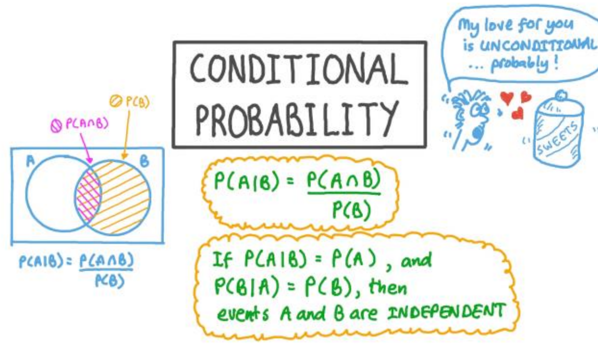
Conditional Probability

Koşullu Olasılık

- başka bir olayın gerçekleşmiş olması koşuluyla, meydana gelen bir olayın olasılığıdır
- **P (A | B)** → "B verildiğinde A nın olması olasılığı



'P(A) given B' diye ifade edilir.



Conditional Probability Formula

Probability of A and B

$$P(A\,|\,B) = \frac{P(A \cap B)}{P(B)}$$

Probability of A given B

Probability of B

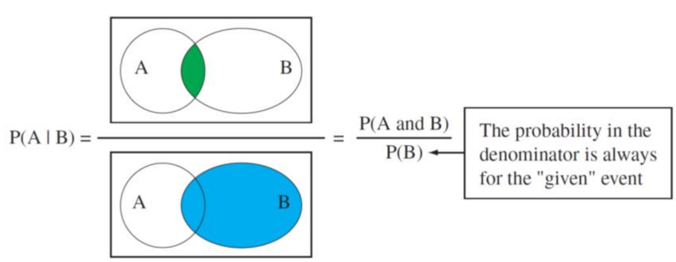
Conditional Probability

Koşullu Olasılık

Conditional Probability Video - StatQuest

- **Örnek**:
  100 üniversite öğrencisine anket uygulanmış ve haftada kaç saat ders çalıştıkları sorulmuştur. Sonuçlar aşağıdaki tablodadır. Bir öğrencinin, bir erkek olması şartıyla 10 saatten fazla ders çalışması olasılığını bulun. (Yandaki tabloya göre)
- **Cevap**:

Örneklem uzayı 49 erkek öğrenciden oluşmaktadır. Bunlardan 16'sı haftada 10 saatten fazla çalışarak geçiriyor



	<5	5 -10	>10	Toplam
Erkek	11	22	16	49
Kadın	13	24	14	51
Toplam	24	46	30	100

P (10 saatten fazla çalışma | erkek) = 16 / 49 = 0,327

Conditional Probability

Koşullu Olasılık

Örnek:

Bir okulun öğrencilerinden **%45'i fizik, %35'i matematik derslerinden ve %25'i hem fizik hem de matematik derslerinden** başarısızdır. Rasgele seçilen bir öğrencinin,

a) Fizikten başarısız olduğu biliniyorsa, matematikten de başarısız olma olasılığını bulunuz

b) Matematikten başarısız olduğu biliniyor ise, fizikten de başarısız olma olasılığını bulunuz

P (M) = 0,35 (Matematikten başarısız olanlar)
P (F) = 0,45 (Fizikten başarısız olanlar)
P (M ∩ F) = 0,25 (Hem M hem F den başarısız olanlar)

a) P (M | F) → (P(M) given F) bulunmalı..
P (M | F) = P (M ∩ F) / P(F)
P (M | F) = 0,25 / 0,45 = **0,55**

b) P (F | M) → (P(F) given M) bulunmalı..
P (F | M) = P (F ∩ M) / P(M)
P (F | M) = 0,25 / 0,35 = **0,71**

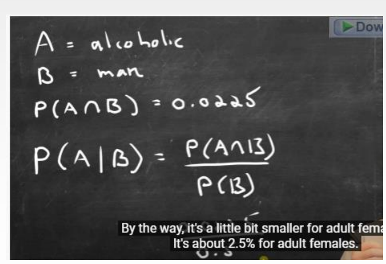
TECHPRO EDUCATION

YOUTUBE VIDEO ONERILERİ

- GCSE Mathematics: Conditional Probability
- Intro to Conditional Probability

https://www.youtube.com/watch?v=_XAqXGJ7fu0

https://www.youtube.com/watch?v=ibINrxJLvlM



Independency check

P (A | B) = P (A) ...?

P (B | A) = P (B) ...?

P (A and B) = P(A) * P(B) ...?

Eğer bu 3 şarttan birisi sağlanıyor ise; A ve B olayları bağımsızdır denilebilir ve diğer 2 şart da gerçekleşir

Independency check

Örnek

Olayların bağımsız mı yoksa bağımlı mı olduğuna karar verin:

Standart bir kart destesinden bir karo (A) ♦ seçilsin ve desteye geri konulsun ve desteden bir maça (B) ♠ seçilsin. Bu durumda

P (B | A) = 13 / 52 = 1 / 4

ve

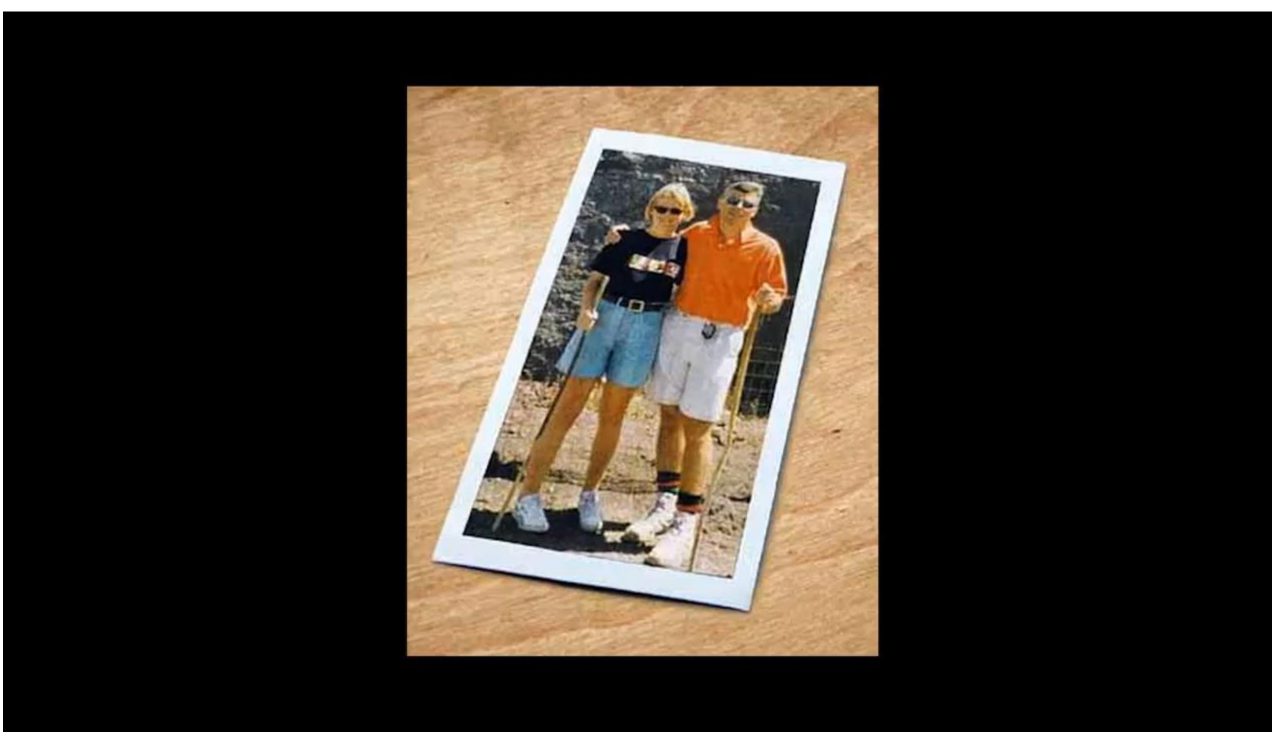
P (B) = 13/52 = 1/4

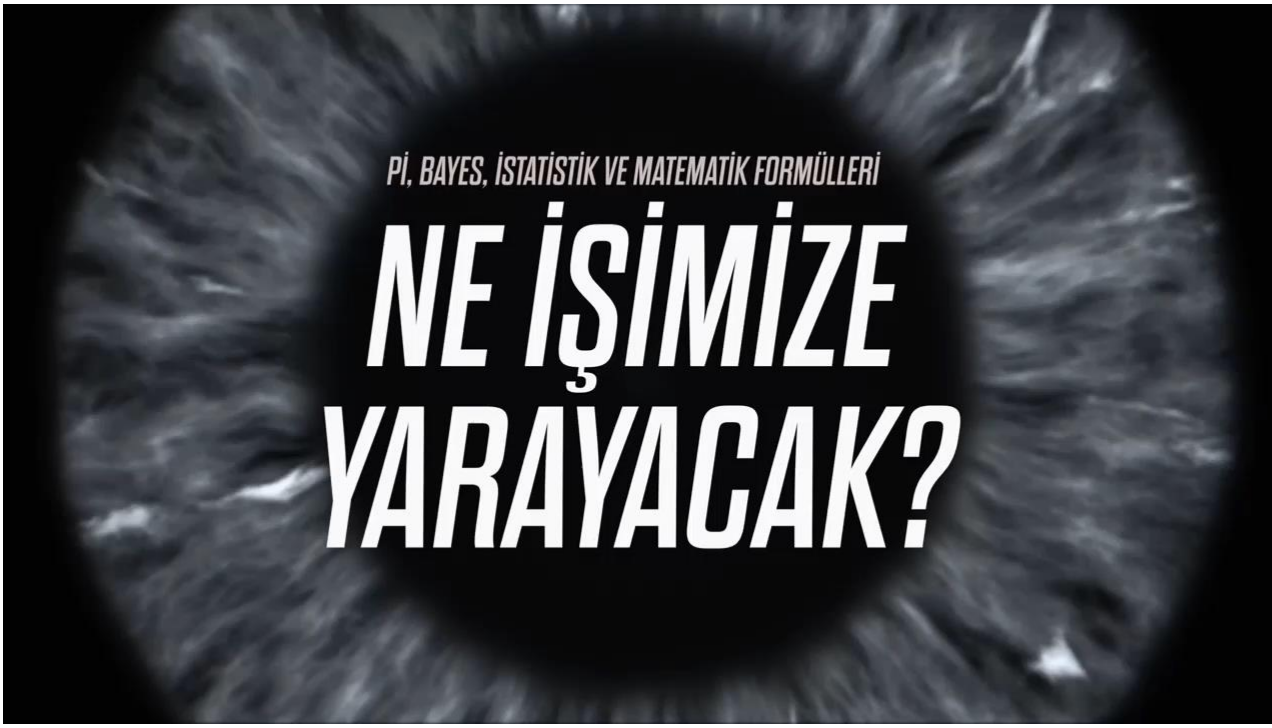
olduğundan

A'nın meydana gelmesi B'nin olasılığını etkilemez, bu nedenle **olaylar bağımsızdır**

BAYES THEOREM

Bayes Teoremi





Bayes Theorem

Bayes Teoremi

- Olayla ilgili olabilecek koşulların ön bilgisine dayalı olarak bir olayın olasılığını tanımlar.
- Bayes Teorem, diğer belirli olasılıkları bildiğimizde, bir olasılık bulmanın yoludur.
- Bu teorem conditional probability den faydalanır

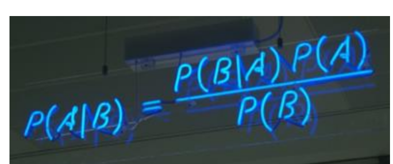
LIKELIHOOD
The probability of "B" being True, given "A" is True

PRIOR
The probability "A" being True. This is the knowledge.

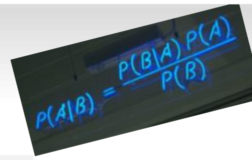
$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

POSTERIOR
The probability of "A" being True, given "B" is True

MARGINALIZATION
The probability "B" being True.



Bayes Theorem



Bayes Teoremi Kurallar

- Sample Space (çözüm Uzayı yani tüm ihtimallerin içinde bulunduğu uzay) birbiriyle kesişmeyen ayrık olaylardan oluşur ($A_1, A_2, \ldots, A_k, \ldots, A_n$)
- Bu örnek uzayındaki P(B)>0 şartını sağlayan (ihtimali 0'dan büyük olan bir B olayı)) B olayı vardır
- Hedefimiz bir koşullu olasılık hesaplamaktır ($P(A_k \mid B)$
- Bazı olasılık değerlerini biliyoruz (her A_k için $P(A_k \cap B)$ değeri ile $P(A_k)$ ve $P(B \mid A_k)$

Bayes Theorem

Bayes Teoremi Formülü

$$P(A \mid B) = \frac{P(A)\, P(B \mid A)}{P(B)}$$

- P(A|B) = B olayı gerçekleştiğinde A olayının gerçekleşme olasılığı
- P(A) = A olayının gerçekleşme olasılığı
- P(B|A) = A olayı gerçekleştiğinde B olayının gerçekleşme olasılığı
- P(B) = B olayının gerçekleşme olasılığı

$$P(A_j \mid B) = \frac{P(B \mid A_j) * P(A_j)}{\Sigma\, P(B \mid A_i) * P(A_i)}$$

Bayes Theorem Example – Basic

Verilenler

- Tehlikeli yangınlar çok az olur - %1
- Duman genelde pikniklerde ve %10 seviyede görülür
- Yangınların %90 ında duman görülür

P(Yangın) = 0,01

P(Duman) = 0,10

P(Duman | Yangın) = 0,90

P (Yangın | Duman) → Bir duman görülünce ne sıklıkla yangın çıkmış olduğu (*Duman verildi, yangın olasılığı)*

P (Duman | Yangın) → Bir yangın olduğu zaman ne sıklıkla duman görüleceği

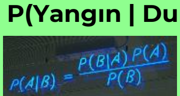
Bayes Theorem Example - Basic

Çözüm

- P(Yangın) = 0,01
- P(Duman) = 0,10
- P(Duman | Yangın) = 0,90

P(Yangın | Duman) = ??

P(Yangın | Duman) = [P(Yangın) * P(Duman|Yangın)] / P(Duman)



P(Yangın | Duman) = (%1 * %90) / %10 = %9

P (Yangın | Duman) → Bir duman görülünce ne sıklıkla yangın çıkmış olduğu (*Duman verildi, yangın olasılığı)*

P (Duman | Yangın) → Bir yangın olduğu zaman ne sıklıkla duman görüleceği

Duman gördüğümüzde %9 ihtimalle bu yangından dolayıdır *(Genelde pikniklerden dolayı idi zaten)*

Bayes Theorem Example – 2

Örnek

- Ali kaşındığını söylüyor. Kedi alerjisi için bir test var, **ancak bu test her zaman doğru değil**:
- Gerçekten alerjisi olan insanlar için, testin "Evet" sonucu vermesi %80 oranında.
- Alerjisi olmayan insanlar için, testin "Evet" sonucu vermesi %10 oranında ("false positive ").
- Nüfusun %1'inde alerji varsa ve test "Evet" çıkıyorsa, **Ali'nin gerçekten alerji olma olasılığı nedir**?

P (A) → Alerji olma ihtimali → **0,01**
P (B) → Testin evet çıkması ihtimali - ?? (Hesaplayacağız)

P (B | A) → Alerji olması durumunda testin evet çıkma olasılığı = **0.80**

P (A | B) → Testin evet çıkması durumunda alerji olasılığı = ?? (bizden istenen olasılık)

Bayes Theorem Example – 2

Çözüm

- Önce P(B) yi bulalım. Testin evet çıkma olasılığını alerjisi olan ve olmayanlar için yani tüm nüfus için hesaplayacağız. (42.slayttaki alttaki formülün paydasına göre hesaplıyoruz)
- P(B) = 0.01 * 0.80 + 0.99 * 0.10 = 0.107 bulunur.

P (A | B) = P(A)* P(B | A) / P(B) = 0,01 *0,80 / 0,107 = 0,075

%7,5 olasılık bulunmuş olur.

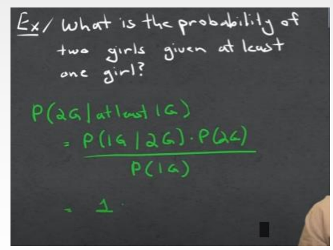
(Testin evet çıkması durumunda alerji olasılığı)

YOUTUBE VIDEO ONERILERİ

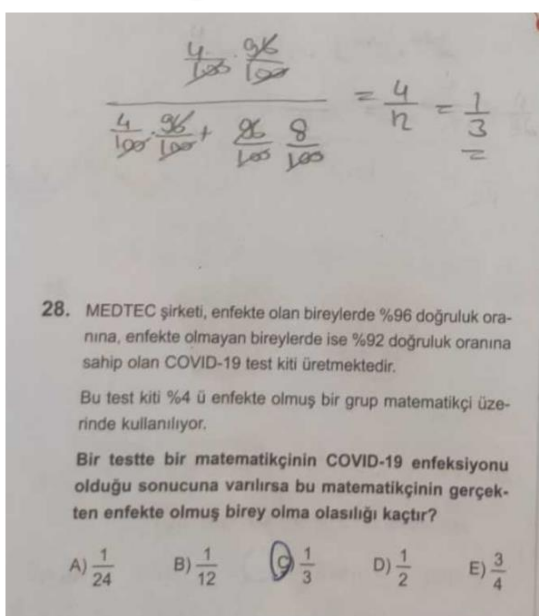
https://www.youtube.com/watch?v=XQoLVl31ZfQ

https://www.youtube.com/watch?v=9wCnvr7Xw4E

- Bayes' Theorem - The Simplest Case
- ayes' Theorem, Clearly Explained!!!!



Güncel Bayes örneği – YKS Sorusu



$$\frac{\frac{4}{100}\cdot\frac{96}{100}}{\frac{4}{100}\cdot\frac{96}{100}+\frac{96}{100}\cdot\frac{8}{100}}=\frac{4}{12}=\frac{1}{3}$$

28. MEDTEC şirketi, enfekte olan bireylerde %96 doğruluk oranına, enfekte olmayan bireylerde ise %92 doğruluk oranına sahip olan COVID-19 test kiti üretmektedir.

Bu test kiti %4 ü enfekte olmuş bir grup matematikçi üzerinde kullanılıyor.

Bir testte bir matematikçinin COVID-19 enfeksiyonu olduğu sonucuna varılırsa bu matematikçinin gerçekten enfekte olmuş birey olma olasılığı kaçtır?

A) $\frac{1}{24}$ B) $\frac{1}{12}$ C) $\frac{1}{3}$ D) $\frac{1}{2}$ E) $\frac{3}{4}$

Bugünkü ders verimli geçti