Ayrık Matematik
(Ayrık İşlemsel Yapılar)

Fırat İsmailoğlu, PhD

Hafta 10:
Olasılık - II

Hafta 10
Plan

1. Bağımsız ve Bağımlı Olaylar
2. Pozitif Korele ve Negatif Korele Olaylar
3. Şartlı Olasılık
4. Zincir Kuralı
5. Bayes Teoremi

Bağımsız ve Bağımlı Olaylar

A ve B iki olay olsun. Eğer, A'nın olmasi yada olmamasi B'nin olmasi hakkinda bize bir bilgi veriyorsa A ve B olaylari (birbirine) bağımlıdır denir. Bir başka ifadeyle A ve B (birbirleriyle) ilintili (correlated) denir.

Eğer A'nın olmasi yada olmamasi B'nin olma olasılığını etkilemiyorsa A ve B olaylari (birbirinden) bağımızdır denir. Bir başka ifadeyle A ve B (birbirleriyle) ilintisiz (uncorrelated) denir.

Formal olarak A ve B olaylarının bağımsız olması için gerek ve yeter şart:

$$P[A \cap B] = P[A] \cdot P[B]$$

(Yani bu olaylarının ikisinin birden olma olasılığı, ayrı ayrı olma olasılıklarının çarpımı olmalıdır)

Bağımsız ve Bağımlı Olaylar

ör. A olayı bir kişinin çift numarali bir ayda doğma olayı. Olasılık $P[\{2,4,6,8,10,12\}] = \frac{1}{2}$

B olayı bir kişinin üçe bölünebilir numarali bir ayda doğma olayı. Olasılık $P[\{3,6,9,12\}] = \frac{1}{3}$

Bu iki olay birbirinden bağımsızdır.

$$P[A \cap B] = P[\{6,12\}] = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = P[A] \cdot P[B]$$

olup A ve B olayları birbirinden bağımsızdır.

ör. Düzgün karılmış bir desteden bir kart seçiliyor. Bu kartın kupa as olma olasılığı nedir?

As olma olayı $P[\{A♣, A♦, A♥, A♠\}]$. Olasılık: $\frac{4}{52} = \frac{1}{13}$

Kupa olma olayı $P[\{A♥, 2♥, \ldots, K♥\}]$. Olasılık: $\frac{13}{52} = \frac{1}{4}$

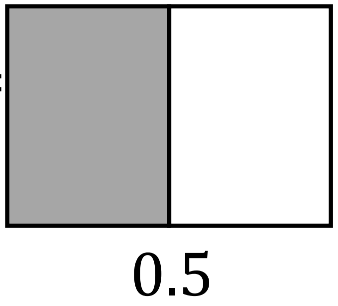
Kupa as olma olayı $P[A♥]$. Olasılık $\frac{1}{13} \cdot \frac{1}{4} = \frac{1}{52}$

Kupa ve as olma birbirinden bağımsız olaylardır.

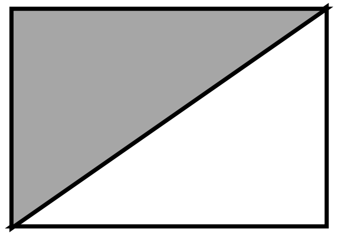
Bağımsız ve Bağımlı Olaylar

örnek uzay, $A =$

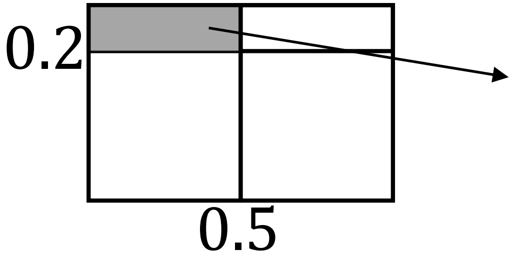
$B =$



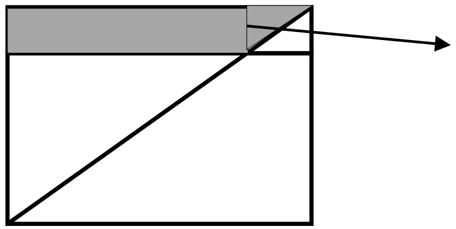
$C =$



olayları verilsin. Bu durumda $P(A) = 0.2, P(B) = 0.5$ ve $P(C) = 0.5$ olur.



$A \cap B$, bu alanın alanı $0.2 \cdot 0.5 = 0.1 = P(A) \cdot P(B)$ olduğundan A ve B olayları bağımsızdır.



$A \cap C$, bu alanın alanı $P(A) \cdot P(C) = 0.5 \cdot 0.5 = 0.25$ den farklı olduğundan A ve C olayları bağımlıdır.

Bağımsız ve Bağımlı Olaylar

ör.Düzgün bir zar atılsın. Zarın

i. tek gelme olayının olasılığı $P[\{1,3,5\}] = 0.5$

ii. asal gelme olayının olaslığı $P[\{2,3,5\}] = 0.5$

tek ve asal gelme olayının olaslığı $P[\{1,3,5\} \cap \{2,3,5\}] = P[\{3,5\}] = 0.33$

Kesişimin olasılığı olan 0.33; $P[\{1,3,5\}] \cdot P[\{2,3,5\}] = 0.5 \cdot 0.5 = 0.25'$ten farklı olduğundan birinci ve ikinci olaylar bağımlı olaylardır.

ör. Bir iskambil destesinden çekilen bir kartın

i. kupa gelme olayının olasılığı $P[\{A♥, 2♥, \ldots, K♥\}] = 0.25$

ii. maça gelme olayının olaslığı $P[\{A♠, 2♠, \ldots, K♠\}] = 0.25$

kupa ve maça gelme olayının olaslığı $P[\emptyset] = 0.$

Kesişimin olaslığı, $P[\{A♥, 2♥ \ldots, K♥\}] \cdot P[\{A♠, 2♠, \ldots, K♠\}] = 0.25 \cdot 0.25 = 0.0625$

den farklı olduğundan olaylar bağımlıdır.

Pozitif Korele ve Negatif Korele Olaylar

A ve B iki olay olsun. Eğer

$$P[A \cap B] > P[A] \cdot P[B]$$

oluyorsa,yanı bırlıkte olma olasiliklari ayri ayri olma olasiliklarindan fazla ise, A ve B olaylarına pozitif korele (positively correlated) olaylar denir.

(Birinin olma olasılığı diğerinin olma olasılığını artırıyor)

Tersine eğer,

$$P[A \cap B] < P[A] \cdot P[B]$$

oluyorsa,yanı birlikte olma olasiliklari ayri ayri olma olasiliklarindan daha az ise, A ve B olaylarına negatif korele (negatively correlated) olaylar denir.

(Birinin olma olasılığı diğerinin olma olasılığını düşürüyor)

Pozitif Korele ve Negatif Korele Olaylar

ör. Sezar şifrelemesinde her bir harf belirli miktarda kaydırılır. Örneğin kaydırma miktarını 3 aldığımızda her harf kendinden sonra gelen 3. harfle yer değiştirir. Bunu geneleşştirirsek her harf farklı miktarlarda kaydırılmış olur. Böylece ortaya çıkan yeni alfabe A,B,C, …. harflerinin yeni bir sıralamasi (permutasyon) olur.

Harfleri bu şekilde rastgele sıralarsak

1. 'D' herfinin yerinde kalması (5.sırada) olasılığı ne olur?
2. 'E' harflerinin yerlerinde kalması ne olur?
3. 'D' ve 'E' harflerinin yerlerinde kalması ne olur?

$$P[D] = \frac{1 \cdot 28 \cdot 27 \cdot \cdots \cdot 1}{29 \cdot 28 \cdot \cdots \cdot 1} = \frac{28!}{29!} = \frac{1}{29}\,; P[E] = \frac{1 \cdot 28 \cdot 27 \cdot \cdots \cdot 1}{29 \cdot 28 \cdot \cdots \cdot 1} = \frac{28!}{29!} = \frac{1}{29}$$

$$P[D \cap E] = \frac{1 \cdot 1 \cdot 27 \cdot \cdots \cdot 1}{29 \cdot 28 \cdot \cdots \cdot 1} = \frac{27!}{29!} = \frac{1}{29 \cdot 28}$$

$P[D \cap E] > P[D] \cdot P[E]$ olduğu için 1. ve 2. olaylar pozitif koreledir. Yani D harfinin yerinde kalmasi, E harfinin yerinde kalmasi olasiligini artirir.

Pozitif Korele ve Negatif Korele Olaylar

ör. PYTHON, MATLAB, JAVA, JUPYTER,PASCAL, PROLOG kelimlerinden biri rastgele seçiliyor. Seçilen kelimede

i. 'A' harfi görme olasılığı nedir?

ii. 'P' görme olasılığı nedir?

iii. 'A' ve 'P' harflerini birlikte görme olasılığı nedir?

$$P[\{MATLAB, JAVA, PASCAL\}] = \frac{3}{6}$$

$$P[\{PYTHON, JUPYTER, PASCAL, PROLOG\}] = \frac{4}{6}$$

$$P[\{MATLAB, JAVA, PASCAL\} \cap \{JUPYTER, PYTHON, PASCAL, PROLOG\}] = P[\{PASCAL\}] = \frac{1}{6}$$

$$P[\{PASCAL\}] < P[\{MATLAB, JAVA, PASCAL\}] \cdot P[\{PYTHON, JUPYTER, PASCAL, PROLOG\}]$$

olduğu için i. ve ii. olaylar negatif koreledir. Bir başka deyişle A'nın görülmesi P'nin görülme olasılığını düşürür.

(P'nin görülmesi A'nın görülme olasılığını düşürür.)

Şartlı Olasılık (Conditional Probability)

Daha önce, A ve B olayları bağımlı ise, B olduğunda A'nin olma olasılığı, B olmadığında A'nin olma olasılığından farklıdır demiştik.

Şimdi B olduğunda A'nin olasılığının nasıl değiştigini sayısal olarak inceliyecegiz.

Şartlı Olasılık

A ve B iki olay olsun. B varken A'nın olma olasılığı (B'nin olmasi şartiyla A'nın olasılığı)

$$\boxed{P[A|B] = \frac{P[A \cap B]}{P[B]}}$$

ör. Bir zar atılsın. Zarın tek geldiği biliniyorsa zarın asal olma olasılığı nedir?

A olayı zarın tek gelmesi $\{1,3,5\}$.

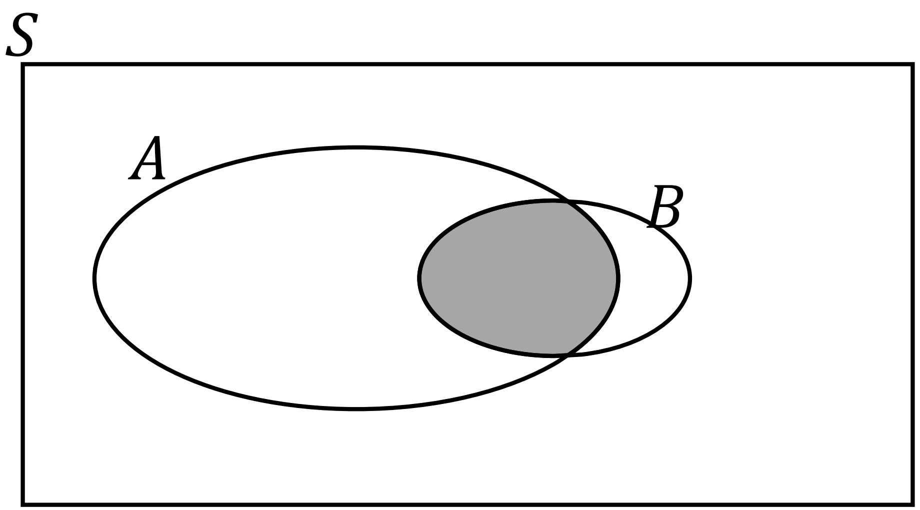
B olayı zarın asal sayı gelmesi $\{2,3,5\}$.

$$P[B|A] = \frac{P[\{3,5\}]}{P[\{1,3,5\}]} = \frac{2/6}{3/6} = \frac{2}{3} \approx 0.66$$

Şartlı Olasılık (Conditional Probability)

Not: $P[B]$ yani zarın asal olma olasılığı 0.5'tir. $P[B|A]$ yani zarın tek olduğu bilinirken zarın asal olma olasılığı 2/3=0.66'dır. Yani zarın asal olma bilgisi olasılığı yükseltir, belirsizliği düşürür!!

ör.



S örnek uzayı, ve A, B olayları verilsin. Varsayalımki $A \cap B$, B'nin %80'nini, A'nın %15'ini oluştursun. Bu durumda $P[A|B] = 0.8$ ve $P[B|A] = 0.15$.

ör. Bozuk bir para 10 kez atılsın. Y en az 9 kez yazı gelme olayı, A ilk atışta yazi gelme olayı, B ilk atışta tura gelme olayı, C ilk 3 atışın yazi gelme olayı olsun.

Bu durumda $P[Y], P[Y|A], P[Y|B], P[Y|C]$ olasılıklarını hesaplayın.

Çözüm.

Y'nin olması için 10 atışın herhangi 9'unun yazı olması yada tamamının yazı olması gerekir.

$$P[Y] = \frac{\binom{10}{9} + \binom{10}{10}}{2^{10}} \approx 0.0107$$

Eğer ilk atışta yazı gelmişse geriye kalan 9 atıştan en az 8'inin yazi gelmesi gerekir. Yani bu 9 atışın ya 8 tanesi yazı olacak; yada hepsi. $Y \cap A$ kümesinin eleman sayısı $\binom{9}{8} + \binom{9}{9} = 10$ olur.

Yada açıkça yazarsak $Y \cap A$ kümesi:

$$\{YYYYYYYYYY, YYYYYYYYYT, YYYYYYYYTY, YYYYYYYTYY, YYYYYYTYYY,$$
$$YYYYYTYYYY, YYYYTYYYYY, YYYTYYYYYY, YYYTYYYYYY, YTYYYYYYYY\}$$

$P[Y \cap A] = \frac{10}{2^{10}}$ olur ve $P[A] = 1/2$ dir.

Bu durumda

$$P[Y|A] = \frac{P[Y \cap A]}{P[A]} = \frac{10/2^{10}}{1/2} \approx 0.019$$

Eğer ilk atışta tura gelmişse geriye kalan 9 atışın tamamının yazı gelmesi gerekir. $Y \cap B$ tek elemanlıdır: $TYYYYYYYYY$. Şu halde $P[Y \cap B] = 1/2^{10}$ olur.

$$P[Y|B] = \frac{P[Y \cap B]}{P[B]} = \frac{1/2^{10}}{1/2} \approx 0.001$$

Eğer ilk 3 atış yazı gelmişse geriye kalan 7 atışın en az 6'sı yazı olmalıdır. $Y \cap C$ kümesinin eleman sayısı $\binom{7}{6} + \binom{7}{7} = 8$ olur.

$$P[Y|C] = \frac{P[Y \cap C]}{P[C]} = \frac{8/2^{10}}{1/8} \approx 0.062$$

Zincir Kuralı

$P[A|B] = \frac{P[A \cap B]}{P[B]}$ idi. Buradan $P[A \cap B] = P[A|B] \cdot P[B]$ elde edilir.

Bu kuralı genelleştirelim:

$A_1, A_2, \dots, A_k$ olaylar olsun. Bu olayların tamamının olma olasılığı

$$P[A_1 \cap A_2 \cap \cdots \cap A_k]$$
$$= P[A_1] \cdot P[A_2|A_1] \cdot P[A_3|A_1 \cap A_2] \cdot \ \dots \cdot P[A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}]$$

Bu kurala zincir kuralı denir.

ör. Pokerde eldeki 5 kartın aynı tür (kupa, sinek,...) olmasına flush denir. Elimize kupa flushın gelme olasılığı nedir?

Zincir Kuralı

K_1 birinci kartin kupa gelme olayı,

K_2 ikinci kartin kupa gelme olayı,

K_1 üçüncü kartin kupa gelme olayı,

K_1 dörduncu kartin kupa gelme olayı,

K_1 beşinci kartin kupa gelme olayı,

olsun. Bu durumda aradığımız olasılık $P[K_1 \cap K_2 \cap K_3 \cap K_4 \cap K_5]$. Zincir kuralı gereği

$$= P[K_1] \cdot P[K_2|K_1] \cdot P[K_3|K_1 \cap K_2] \cdot P[K_4|K_1 \cap K_2 \cap K_3] \cdot P[K_5|K_1 \cap K_2 \cap K_3 \cap K_4]$$

$$= \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} \cdot \frac{10}{49} \cdot \frac{9}{48}$$

$$\approx 0.00049$$

Bayes Teoremi

$P[A|B] = \frac{P[A \cap B]}{P[B]}$ idi. Buaradan $P[A \cap B] = P[A|B] \cdot P[B]$ elde edilir.

Aynı zamanda $P[B|A] = \frac{P[A \cap B]}{P[A]}$. Buradan $P[A \cap B] = P[B|A] \cdot P[A]$ elde edilir.

Bulunan iki eşitligi birbirine esitlersek

$$P[A|B] \cdot P[B] = P[B|A] \cdot P[A]$$

olur. $P[A|B]$ yalnız bırakılırsa:

$$\boxed{P[A|B] = \frac{P[B|A] \cdot P[A]}{P[B]}}$$

olur. Bu teoreme Bayes teoremi denir.

ör. Bir klinikte yapılan kanser testi gerçekte kanser olan hastaların %98'inde pozitif olarak sonuç veriyor. Ayrıca toplumdaki kişilerin 0.008'nin kanser olduğu ve bu testin 0.03 olasılıkla pozitif sonuç verdiği biliniyor. Bu durumda testi pozitif çıkan birinin gerçekte kanser olma olasılığı nedir?

Çözüm.

$$P[test\ pozitif|kanser] = 0.98$$
$$P[kanser] = 0.008$$
$$P[test\ pozitif] = 0.03$$
$$P[kanser|test\ pozitif] = \frac{P[test\ pozitif|kanser] \cdot P[kanser]}{P[test\ pozitif]} = \frac{0.98 \cdot 0.008}{0.03} \approx 0.2$$

Sonuç olarak test pozitif sonuç vermesine rağmen kisinin kanser olma olasılığı düşüktür. Bu anlamda bu test icin basarisiz bir testtir diyebiliriz.

Bayes teoreminin en onemli ozelligi bilgimizi guncellememizi saglamasidir. Önceki ornekte kanser görülme olasılığını normalde 0.008 olarak biliyoruz. Ama bize kisinin testini pozitif çıktığı bilgisi verildiginde bu sonucu kullanarak kanser olma olasılığı ile ilgili bilgimizi günceleyebiliyoruz. Güncellerken edindigimiz yeni bilgiyi hesaba katabiliyoruz.

ör. Bir kampüsteki öğrencilerin %5'nin bilgisayar mühendisliği öğrencisi olduğu biliniyor. Ayrıca bilgisayar mühendisliği öğrencilerinin %50'sinin yurtta kaldığı ve kampüsteki tüm öğrencilerin %10'unun yurtta kaldığı biliniyor. Buna göre yurtta kaldığı bilinen birinin bilgisayar mühendisliği öğrencisi olması olasılığı nedir?

Çözüm.

$$P[bilgisayar] = 0.05, P[yurt|bilgisayar] = 0.5, P[yurt] = 0.1$$

$$P[bilgisayar|yurt] = \frac{P[yurt|bilgisayar] \cdot P[bilgisayar]}{P[yurt]} = \frac{0.5 \cdot 0.05}{0.1} = 0.25$$

Eğer kişinin yurtta kaldığı bilinmeseydi bu kişinin bilgisiyar mühendisliği öğrencisi olma olasılığı 0.05 olacaktı. Fakat bu kisinin yurtta kaldiği bilgisini bilmemiz, bilgisayar muhendisligi ögrencisi olma olasiligini 0.25'e çıkardı.