Koşullu Olasılık

IST 108 Olasılık ve İstatistik

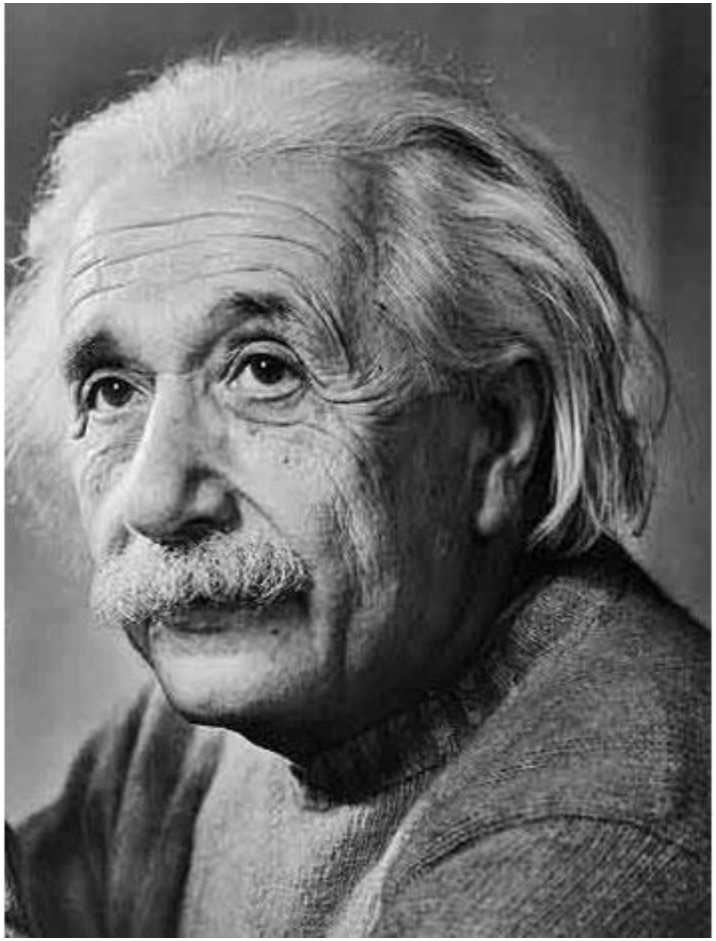
Bahar 2016

Yrd. Doç. Dr. Ferhat Dikbıyık

Önce geçmişten bir alıntı

"Bir kişi için bilgiyi öğrenmek çok önemli değildir. Bunun için de yüksek okul (lisans) eğitimi almasına ihtiyacı yoktur. Bunları ders kitaplarından da öğrenebilir. Yüksek okul eğitiminin değeri, birçok bilgiyi öğrenmek değil, ders kitaplarından öğrenilemeyecek şeyleri düşünmek için zihinlerin eğitilmesidir."

- 1921, Albert Einstein'in, Thomas Edisson'un yüksek okul eğitiminin faydasız olduğu görüşü üzerine söyledikleri, *Einstein: His Life and Times,* Philipp Frank, sayfa 185



Koşullu Olasılık

- Verilen A ve B olayları için, deney defterinde "$A \mid B$" adında başka bir sütun daha düşünelim. Bu sütundaki her bir satırda:
  - Eğer B olayı Hayır ise Uygulanmaz yazacaktır.
  - Eğer bu satırda B olayı için Evet diyorsa, bu durumda bu sütun için A olayının Evet yada Hayır olmasına göre Evet yada Hayır yazacaktır.
- Bu durumda $P(A \mid B)$ bu son sütundaki Evet yazan satırların B sütununa evet yazan satırlara uzun-koşumlu oranı olacaktır , yani B olayı olduğunda A'nın olma olasılığı.

P(AB) ve P($A|B$)

- Genelde P(AB) ve P($A|B$) karıştırılır. İki zar atma örneğimizde X: ilk zarda gelen nokta sayısını ve Y: ikinci zarda gelen nokta sayısını ifade ediyordu. A olayı $X=1$ ve B olayı $X+Y=6$ olsun Bu örnek için P(AB) $=1/36$ iken P($A|B$)$=1/5$ olacaktır. Yani;
  - Deneyin birçok tekrarından sonra yaklaşık olarak deney defterindeki satırların 1/36'sında A (yani $X=1$) ve B (yani $X+Y=6$) olayı için Evet yazacaktır.
  - Deneyin birçok tekrarından sonra, eğer **sadece *B*'nin Evet yazdığı satırlara bakarsak**, *bu satırların* 1/5'inde A olayı için Evet yazacaktır.
- P($A|B$), B olayı olduğunda A'nın koşullu olasılığıdır.

Koşullu Olasılık Formülü

A ve *B*'nin aynı anda olma olasılığı, ama **ne zaman**?

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

B olayı olduğunda!!!

Formüle başka bir açı

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

Yani iki olayın aynı anda olma olasılığı, birinin (diğeri verildiğinde) koşullu olasılığı ile diğer olayın olma olasılığının çarpımıdır.

$$P(AB) = P(B)P(A|B)$$

$$P(AB) = P(A)P(B|A)$$

Örnek 1

- Bir program eşit iki ihtimalle iki fonksiyondan (prosedürden) birini çağırmaktadır. Belirli girdiler (input) için eğer birinci fonksiyonu çağırırsa ½ ihtimalle "bellek dolu" (out of memory) hatası vermektedir. İkinci fonskiyonu çağırdığında ise bu ihtimal 2/3 olmaktadır. Bu durumda bu girdiler için programın "bellek dolu" hatası verme ihtimali nedir?

Örnek 1

- Bir program eşit iki ihtimalle iki fonksiyondan (prosedürden) birini çağırmaktadır. Belirli girdiler (input) için eğer birinci fonksiyonu çağırırsa ½ ihtimalle "bellek dolu" (out of memory) hatası vermektedir. İkinci fonskiyonu çağırdığında ise bu ihtimal 2/3 olmaktadır. Bu durumda bu girdiler için programın "bellek dolu" hatası verme ihtimali nedir?
- A_1: Birinci fonksiyonu çağırması
- A_2: İkinci fonksiyonu çağırması
- H: Programın hata vermesi?
- Sorulan $P(H) = ?$

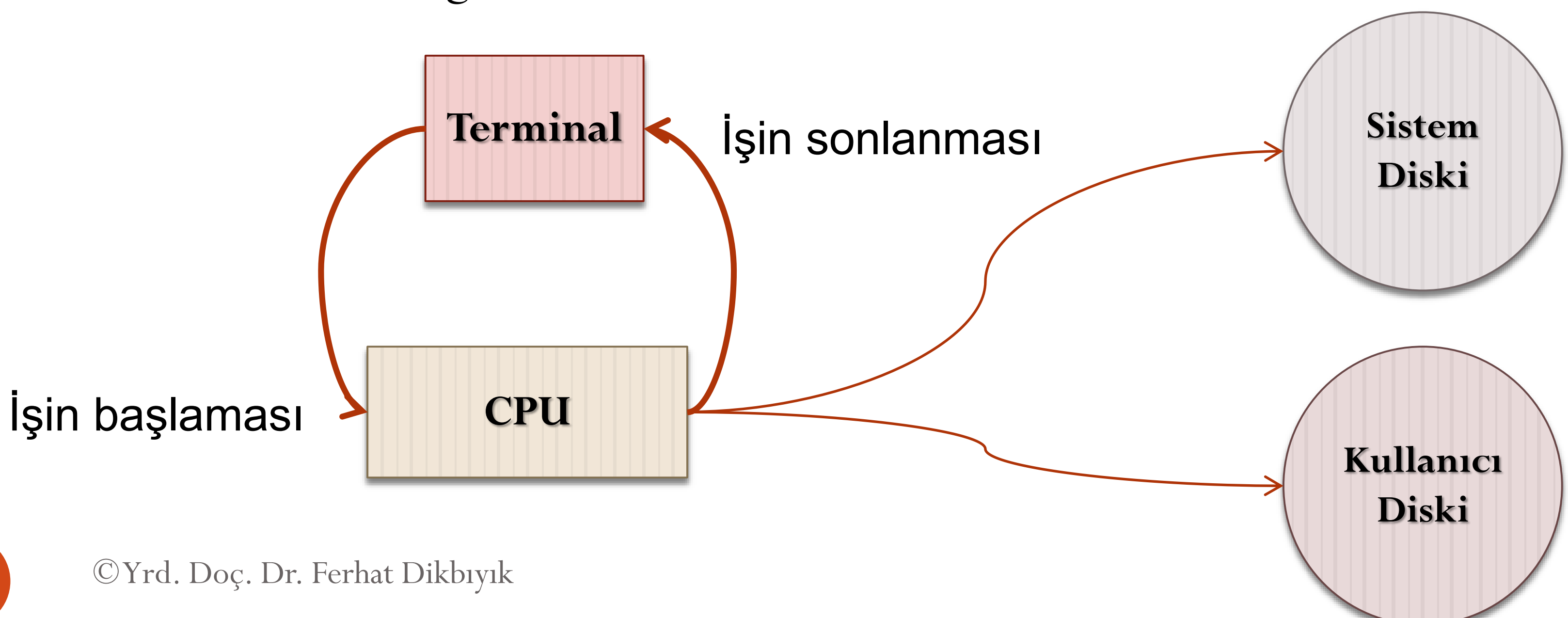
Örnek 1

- Bir program eşit iki ihtimalle iki fonksiyondan (prosedürden) birini çağırmaktadır. Belirli girdiler (input) için eğer birinci fonksiyonu çağırırsa ½ ihtimalle "bellek dolu" (out of memory) hatası vermektedir. İkinci fonskiyonu çağırdığında ise bu ihtimal 2/3 olmaktadır. Bu durumda bu girdiler için programın "bellek dolu" hatası verme ihtimali nedir?

- $P(H) = P(H \text{ ve } A_1 \textit{ veya } H \text{ ve } A_2) = P(HA_1) + P(HA_2)$
- $P(H) = P(A_1)\, P(H|A_1) + P(A_2)\, P(H|A_2)$
- $P(H) = (\,½\,)(\,½\,) + (\,½\,)(\,2/3\,) = 7\,/\,12$

Örnek 2

- Bir bilgisayar sisteminde CPU'ya gelen bir iş sisteme ait bir iş ise sistem diskine, eğer kullanıcıya ait bir iş ise kullanıcı diskine gönderiliyor. Eğer iş tamamlanmışsa CPU işi terminale gönderiyor. Gelen işlerin %40'ı sisteme ait işler ve %60'ı ise kullanıcıya ait işlerden oluşmaktadır. CPU'nun kendisine gelen sisteme ait bir işi sistem diskine göndermeden terminale gönderme ihtimali %2'dir. Bu durumda CPU gelen bir işin sistem diskine gönderilme ihtimali nedir?



Örnek 2

- Bir bilgisayar sisteminde CPU'ya gelen bir iş sisteme ait bir iş ise sistem diskine, eğer kullanıcıya ait bir iş ise kullanıcı diskine gönderiliyor. Eğer iş tamamlanmışsa CPU işi terminale gönderiyor. Gelen işlerin %40'ı sisteme ait işler ve %60'ı ise kullanıcıya ait işlerden oluşmaktadır. CPU'nun kendisine gelen sisteme ait bir işi sistem diskine göndermeden terminale gönderme ihtimali %2'dir. Bu durumda CPU gelen bir işin sistem diskine gönderilme ihtimali nedir?
- S: İşin sisteme ait bir iş olması
- K: İşin kullanıcıya ait bir iş olması
- T: İşin sonlandırılarak terminale gönderilmesi
- A: İşin sistem diskine gönderilmesi
- B: İşin kullanıcı diskine gönderilmesi

Örnek 2

- Bir bilgisayar sisteminde CPU'ya gelen bir iş sisteme ait bir iş ise sistem diskine, eğer kullanıcıya ait bir iş ise kullanıcı diskine gönderiliyor. Eğer iş tamamlanmışsa CPU işi terminale gönderiyor. Gelen işlerin **%40**'ı sisteme ait işler ve **%60**'ı ise kullanıcıya ait işlerden oluşmaktadır. CPU'nun kendisine gelen sisteme ait bir işi sistem diskine göndermeden terminale gönderme ihtimali **%2**'dir. Bu durumda CPU gelen bir işin sistem diskine gönderilme ihtimali nedir?
- $P(S) = 0,4$
- $P(K) = 0,6$
- $P(T \mid S) = 0,02$
- $P(A) = ?$

Örnek 2

- Bir bilgisayar sisteminde CPU'ya gelen bir iş sisteme ait bir iş ise sistem diskine, eğer kullanıcıya ait bir iş ise kullanıcı diskine gönderiliyor. Eğer iş tamamlanmışsa CPU işi terminale gönderiyor. Gelen işlerin **%40**'ı sisteme ait işler ve **%60**'ı ise kullanıcıya ait işlerden oluşmaktadır. CPU'nun kendisine gelen sisteme ait bir işi sistem diskine göndermeden terminale gönderme ihtimali **%2**'dir. Bu durumda CPU gelen bir işin sistem diskine gönderilme ihtimali nedir?
- $P(S) = 0{,}4,\ P(K) = 0{,}6,\ P(T|S) = 0{,}02$
- $P(A) = P(AS)\ /\ P(S|A) = P(AS)$

Bir işin sistem diskine gönderildiği verildiğinde, işin sisteme ait bir iş olması olasılığı 1'dir.

- $P(AS) = P(A|S)P(S) = [1\text{-}P(T|S)]P(S) = (0{,}98)(0{,}4) = 0{,}392$

Ya ikiden fazla olay varsa

- Çarpım kuralı;
  - $P(A_1A_2A_3\ldots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_2A_1)\ldots P(A_n|A_1\ldots A_{n-1})$

$$\cancel{P(A_1)}\frac{\cancel{P(A_1A_2)}}{\cancel{P(A_1)}}\frac{\cancel{P(A_1A_2A_3)}}{\cancel{P(A_1A_2)}}\cdots\frac{P(A_1A_2\ldots A_n)}{\cancel{P(A_1A_2\ldots A_{n-1})}}$$

Örnek 3

- Bir oyun programcısı oyuncunun dört farklı bölgeye eşit olarak yerleştirilmiş 52 parçadan 4 bonus parçayı (yeşil, mavi, sarı, kırmızı) bulacağı bir oyun tasarlamıştır (her bölgede 13 parça vardır).
- Programcı oyunun en zor halinin her bir bonus parçanın 4 farklı bölgede bulunması durumunda olacağını düşünmektedir.
- Programcının bağlı olduğu oyun şirketi oyunun en zor hale gelme olasılığının %10'unun üzerinde olmasını, aksi halde bu durumun oyuncuyu oyundan soğutacağını düşünmektedir.
- Bu durumda programcı oyunda değişiklik yapmalı mıdır, yoksa olduğu gibi bırakmalı mıdır?

Örnek 3

- 52 parça, her bir bölgede 13 parça. Her bölgenin tam olarak 1 bonus parça içerme ihtimalini hesaplayalım.
- E_1: yeşil parça herhangi bir bölgenin içinde
- E_2: yeşil ve mavi parçalar farklı bölgelerin içinde
- E_3: yeşil, mavi ve sarı parçaların hepsi farklı bölgelerin içinde
- E_4: Tüm bonus parçaların hepsi farklı bölgelerin içinde
- Soruda istenen $P(E_1E_2E_3E_4) = ?$

Örnek 3

- 52 parça, her bir bölgede 13 parça. Her bölgenin tam olarak 1 bonus parça içerme ihtimalini hesaplayalım.
- $P(E_1) = 1$
- $P(E_2 | E_1) = 39 / 51$
- $P(E_3 | E_1 E_2) = 26 / 50$
- $P(E_4 | E_1 E_2 E_3) = 13 / 49$
- Çarpım kuralına göre
  - $P(E_1 E_2 E_3 E_4) = 39/51 \text{ x } 26/50 \text{ x } 13/49 = 0{,}105$
  - %10'dan büyük olduğu için programcının bir değişiklik yapmasına gerek yoktur.

Baye's Formülü

$$P(A) = P(AB) + P(A\overline{B})$$

$$= P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$$

$$= P(A|B)P(B) + P(A|\overline{B})\left[1 - P(B)\right]$$

Baye’s Formülü

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})}$$

Örnek 4

- Bir sigorta şirketi insanları kazaya meyilli olanlar ve olmayanlar olarak iki ayrı gruba ayırmıştır. Şirket kazaya meyilli olan insanların 1 yıllık süre içerisinde kaza yapma ihtimallerinin 0,4 olduğunu, kazaya meyilli olmayanlar için aynı ihtimalin ise 0,2 olduğunu hesaplamıştır. Eğer nüfusun %30'unun kazaya meyilli olduğunu varsayarsak yeni kaza poliçesi yaptırmış bir sigortalının 1 yıl içinde kaza yapma ihtimali nedir?

Örnek 4

- Bir sigorta şirketi insanları kazaya meyilli olanlar ve olmayanlar olarak iki ayrı gruba ayırmıştır. Şirket kazaya meyilli olan insanların 1 yıllık süre içerisinde kaza yapma ihtimallerinin 0,4 olduğunu, kazaya meyilli olmayanlar için aynı ihtimalin ise 0,2 olduğunu hesaplamıştır. Eğer nüfusun %30'unun kazaya meyilli olduğunu varsayarsak yeni kaza poliçesi yaptırmış bir sigortalının 1 yıl içinde kaza yapma ihtimali nedir?
  - K: Sigortalının 1 yıl içinde kaza yapması
  - M: Sigortalının kazaya meyilli olması
  - $P(K) = ?$

Örnek 4

- Bir sigorta şirketi insanları kazaya meyilli olanlar ve olmayanlar olarak iki ayrı gruba ayırmıştır. Şirket kazaya meyilli olan insanların 1 yıllık süre içerisinde kaza yapma ihtimallerinin 0,4 olduğunu, kazaya meyilli olmayanlar için aynı ihtimalin ise 0,2 olduğunu hesaplamıştır. Eğer nüfusun %30'unun kazaya meyilli olduğunu varsayarsak yeni kaza poliçesi yaptırmış bir sigortalının 1 yıl içinde kaza yapma ihtimali nedir?
- $P(K) = P(KM) + P(KM') = P(K|M)P(M) + P(K|M')P(M')$

$$= 0{,}4 \times 0{,}3 + 0{,}2 \times 0{,}7 = 0{,}26$$

Örnek 4

- Bir sigorta şirketi insanları kazaya meyilli olanlar ve olmayanlar olarak iki ayrı gruba ayırmıştır. Şirket kazaya meyilli olan insanların 1 yıllık süre içerisinde kaza yapma ihtimallerinin 0,4 olduğunu, kazaya meyilli olmayanlar için aynı ihtimalin ise 0,2 olduğunu hesaplamıştır. Eğer nüfusun %30'unun kazaya meyilli olduğunu varsayarsak yeni kaza poliçesi yaptırmış bir sigortalının 1 yıl içinde kaza yapma ihtimali nedir?
  - Yeni poliçe sahibinin 1 yıl içinde kaza yaptığı bilindiğinde sigortalının kazaya meyilli olma ihtimali nedir?

Örnek 4

- Bir sigorta şirketi insanları kazaya meyilli olanlar ve olmayanlar olarak iki ayrı gruba ayırmıştır. Şirket kazaya meyilli olan insanların 1 yıllık süre içerisinde kaza yapma ihtimallerinin 0,4 olduğunu, kazaya meyilli olmayanlar için aynı ihtimalin ise 0,2 olduğunu hesaplamıştır. Eğer nüfusun %30'unun kazaya meyilli olduğunu varsayarsak yeni kaza poliçesi yaptırmış bir sigortalının 1 yıl içinde kaza yapma ihtimali nedir?
  - Yeni poliçe sahibinin 1 yıl içinde kaza yaptığı bilindiğinde sigortalının kazaya meyilli olma ihtimali nedir?
  - $P(M|K)=?$

Örnek 4

- Bir sigorta şirketi insanları kazaya meyilli olanlar ve olmayanlar olarak iki ayrı gruba ayırmıştır. Şirket kazaya meyilli olan insanların 1 yıllık süre içerisinde kaza yapma ihtimallerinin 0,4 olduğunu, kazaya meyilli olmayanlar için aynı ihtimalin ise 0,2 olduğunu hesaplamıştır. Eğer nüfusun %30'unun kazaya meyilli olduğunu varsayarsak yeni kaza poliçesi yaptırmış bir sigortalının 1 yıl içinde kaza yapma ihtimali nedir?
  - Yeni poliçe sahibinin 1 yıl içinde kaza yaptığı bilindiğinde sigortalının kazaya meyilli olma ihtimali nedir?
  - $P(M|K) = P(MK) \,/\, P(K) = P(K|M)P(M) \,/\, P(K) =$

    $= 0{,}3 \times 0{,}4 \,/\, 0{,}26 = 6/13$

Problem 1

- Bir kan testi belirli bir hastalığı tespit etmek için hastalık gerçekten varsa %95 etkilidir. Fakat test, sağlıklı bir insan için %1 olasılıkla yanlış sonuç vererek hastalık tespit ediyor. Eğer nüfusun %0,5'i bu hastalığa sahipse test sonucu pozitif çıkan birinin gerçekten hasta olma ihtimali nedir?

Problem 1

- Bir kan testi belirli bir hastalığı tespit etmek için hastalık gerçekten varsa %95 etkilidir. Fakat test, sağlıklı bir insan için %1 olasılıkla yanlış sonuç vererek hastalık tespit ediyor. Eğer nüfusun %0,5'i bu hastalığa sahipse test sonucu pozitif çıkan birinin gerçekten hasta olma ihtimali nedir?
- H: Test yapan kişinin hasta olması
- T: Testin pozitif olması
- $P(H|T)=?$

Problem 1

- Bir kan testi belirli bir hastalığı tespit etmek için hastalık gerçekten varsa %95 etkilidir. Fakat test, sağlıklı bir insan için %1 olasılıkla yanlış sonuç vererek hastalık tespit ediyor. Eğer nüfusun % 0,5'i bu hastalığa sahipse test sonucu pozitif çıkan birinin gerçekten hasta olma ihtimali nedir?
- $P(H|T) = P(TH)/P(T)$

$$= P(T|H)P(H) \,/\, [P(TH)+P(TH')]$$

$$= P(T|H)P(H) \,/\, [P(T|H)P(H) + P(T|H')P(H')\,]$$

$$= (0{,}95)(0{,}005) \,/\, [(0{,}95)(0{,}005) + (0{,}01)(0{,}995)\,]$$

$$= 95/294 = 0{,}323$$

Problem 2

- Bir doktor şöyle bir ikilemde kalmıştır.
- "Eğer hastamın belirli bir hastalığa %80 sahip olduğuna inanırsam, o zaman her zaman ameliyat öneririm. Eğer o kadar emin değilsem, yeni testler isterim. Fakat bu testler pahalı ve acı vericidir.
- Başlangıçta Ahmet'in hasta olduğundan %60 emindim, bu yüzden A serisi testlerini yapmasını istedim. Bu testler hasta ise her zaman pozitif sonuç verir ve eğer değilse hiçbir zaman pozitif sonuç vermez.
- Test sonuçları pozitif çıktı ve ben de tam ameliyat olmasını tavsiye edecekken Ahmet diyabet hastası olduğunu söyledi. Bu bilgi sorunu karmaşık hale getirdi, çünkü (hala Ahmet'in %60 ihtimalle o hastalığa sahip olduğunu düşünsem de) bu durum A testinin sonuçlarının okunmasını değiştirecektir.
- Şöyle ki A testi hasta sağlıklı iken hiçbir zaman pozitif sonuç üretmese de diyabet olan ama bu hastalığa sahip olmayan insanlar için %30 ihtimalle pozitif sonuç üretmektedir.
- Bu durumda ne yapmalıyım? Daha fazla test mi yoksa ameliyat mı?"

Problem 2

- "Eğer hastamın belirli bir hastalığa %80 sahip olduğuna inanırsam, o zaman her zaman ameliyat öneririm. Eğer o kadar emin değilsem, yeni testler isterim. Fakat bu testler pahalı ve acı vericidir. Başlangıçta Ahmet'in hasta olduğundan %60 emindim, bu yüzden A serisi testlerini yapmasını istedim. Bu testler hasta ise her zaman pozitif sonuç verir ve eğer değilse hiçbir zaman pozitif sonuç vermez. Test sonuçları pozitif çıktı ve ben de tam ameliyat olmasını tavsiye edecekken Ahmet diyabet hastası olduğunu söyledi. Bu bilgi sorunu karmaşık hale getirdi, çünkü (hala Ahmet'in %60 ihtimalle o hastalığa sahip olduğunu düşünsem de) bu durum A testinin sonuçlarının okunmasını değiştirecektir. Şöyle ki A testi hasta sağlıklı iken hiçbir zaman pozitif sonuç üretmese de diyabet olan ama bu hastalığa sahip olmayan insanlar için %30 ihtimalle pozitif sonuç üretmektedir. Bu durumda ne yapmalıyım? Daha fazla test mi yoksa ameliyat mı?"
- H: Ahmet gerçekten o hastalığa sahip
- T: Test sonucu pozitif
- $P(H|T)=?$

Problem 2

- “Eğer hastamın belirli bir hastalığa %80 sahip olduğuna inanırsam, o zaman her zaman ameliyat öneririm. Eğer o kadar emin değilsem, yeni testler isterim. Fakat bu testler pahalı ve acı vericidir. Başlangıçta Ahmet’in hasta olduğundan %60 emindim, bu yüzden A serisi testlerini yapmasını istedim. Bu testler hasta ise her zaman pozitif sonuç verir ve eğer değilse hiçbir zaman pozitif sonuç vermez. Test sonuçları pozitif çıktı ve ben de tam ameliyat olmasını tavsiye edecekken Ahmet diyabet hastası olduğunu söyledi. Bu bilgi sorunu karmaşık hale getirdi, çünkü (hala Ahmet’in %60 ihtimalle o hastalığa sahip olduğunu düşünsem de) bu durum A testinin sonuçlarının okunmasını değiştirecektir. Şöyle ki A testi hasta sağlıklı iken hiçbir zaman pozitif sonuç üretmese de diyabet olan ama bu hastalığa sahip olmayan insanlar için %30 ihtimalle pozitif sonuç üretmektedir. Bu durumda ne yapmalıyım? Daha fazla test mi yoksa ameliyat mı?”
- $P(H|T) = P(TH)/P(T)$

$$= P(T|H)P(H) \,/\, [P(TH)+P(TH')]$$

$$= P(T|H)P(H) \,/\, [P(T|H)P(H) + P(T|H')P(H')]$$

$$= (1)(0{,}6) \,/\, [(1)(0{,}6) + (0{,}3)(0{,}4)]$$

$$= 0{,}833$$

Yani Ameliyat!