IST2084/ IST104.1/ IST104.2 Biyoistatistik (Biyoloji Bölümü)

OLASILIK

Fatih Kızılaslan

Marmara Üniversitesi

2019-2020 Bahar 8. Hafta

Toplam Olasılık Formülü

Birleşimleri örnek uzayı veren $A_1, A_2, ..., A_n$ biçiminde n tane ayrık olayımız olsun. Bu durumda bu olaylar

- $i \neq j$ olmak üzere $A_i \cap A_j = \varnothing$
- $S = \cup_{i=1}^{n} A_i$

koşullarını sağlar. B bu örnek uzaydaki herhangi bir olay olsun($B \subset S$). Bu durumda,

$$B = (B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3) \cup ... \cup (B \cap A_n) = \bigcup_{i=1}^{n} (B \cap A_i)$$

biçiminde yazabiliriz. Burada $(B \cap A_i), i = 1, ..., n$ ayrık olaylardır. Böylece,

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + ... + P(B \cap A_n) = \sum_{i=1}^{n} P(B \cap A_i)$$

olur.

Örnek 1: Bir fabrikada kullanılan 3 farklı makine ile eşit miktarda ürün üretilmektedir. 1. makine ile üretilen ürünlerin 0.02'si bozuk, 2. makine ile üretilenlerin 0.05'si bozuk ve 3. makine ile üretilenlerin 0.07'si bozuktur. Üretilen ürünlerden rastgele seçilen bir ürünün bozuk olma olasılığı nedir?

Çözüm: İlk olarak gerekli olan bazı kümeleri tanımlayalım. $B = \{$Seçilen ürünün bozuk olması$\}$, $A_1 = \{$Ürünün 1. makinede üretilmesi$\}$, $A_2 = \{$Ürünün 2. makinede üretilmesi$\}$ ve $A_3 = \{$Ürünün 3. makinede üretilmesi$\}$ olarak tanımlayalım. Bu tanımlara göre $A_1 \cap A_2 = \varnothing$, $A_1 \cap A_3 = \varnothing$ ve $A_2 \cap A_3 = \varnothing$ olur, yani bu kümeler ayrıktır. Rastgele seçilen bir ürün bu üç makinenin biri tarafından üretilmiştir. Dolayısıyla,

$$B = (B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3)$$

biçiminde yazabiliriz. $A_i, i = 1, 2, 3$ kümeleri ayrık olduğundan $B \cap A_i, i = 1, 2, 3$ kümeleri de ayrık olur.

Böylece, istenilen rastgele seçilen ürünün bozuk olma olasılığı

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3)$$

biçiminde heasplarız.
Ayrıca, bu üç makine eşit miktada ürün ürettiğinden üretilen herhangi bir ürünün A_1 veya A_2 veya A_3'de üretilmesi $P(A_1) = P(A_2) = P(A_3) = 1/3$ olur.
Her makine için bozuk ürün oranlarını bildiğimizden koşullu olasılıkları bulabiliriz. 1. makinede üretildiği bilinen ürünün bozuk olma olasılğı $P(B/A_1)$:

$$P(B/A_1) = 0.02 = \frac{2}{100} \text{ olur.}$$

Benzer biçimde, $P(B/A_2) = \frac{5}{100}$ ve $P(B/A_3) = \frac{7}{100}$ olur. Böylece, $P(B \cap A_1) = P(A_1)P(B/A_1) = \frac{1}{3}\frac{2}{100}$ bulunur ve

$$P(B) = \frac{1}{3}\frac{2}{100} + \frac{1}{3}\frac{5}{100} + \frac{1}{3}\frac{7}{100} = \frac{14}{300} = 0.047$$

yani binde 47 olarak bulunur.

Bayes Formülü

A_i, $i = 1, ..., n$ ayrık olaylar olmak üzere $S = \bigcup_{i=1}^{n} A_i$ olsun. S örnek uzayındaki herhangi bir B olayı için $P(B) \neq 0$ olmak üzere $P(A_i/B)$, $i = 1, ..., n$ koşullu olasılığı

$$P(A_i/B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B/A_i)}{\sum_{i=1}^{n} P(A_i)P(B \cap A_i)}, i = 1, ..., n$$

biçiminde hesaplanır.

Bayes formülünde yukarıda verdiğimiz toplam olasılık formülü kullanılır.

Örnek 2: 2012'deki başkanlık seçiminde Ohio eyaletinin sandık çıkış anketleri aşağıdaki sonuçları vermiştir.

	Obama	**Romney**
Üniversite mezunu olmayan (%60)	%52	%47
Üniversite mezunu (%40)	%47	%50

Rastgele seçilmiş bir katılımcı Obama'ya oy vermişse üniversite mezunu olma olasılığı nedir?

Çözüm: Gerekli kümeler $O = \{$Obama'ya oy verenler$\}$, $\ddot{U} = \{$Üniversite mezunu olanlar$\}$ ve $\ddot{U}^C = \{$Üniversite mezunu olmayanlar$\}$. İstenilen olasılık

$$P(\ddot{U}/O) = \frac{P(\ddot{U} \cap O)}{P(O)} = \ ?$$

Obama'ya oy verenleri üniversite mezunu ve mezun olmayan olmak üzere iki gruba ayırabiliriz (ayrık kümelerdir).

Dolayısıyla, toplam olasılık formülü ile

$$\begin{aligned} P(O) &= P(\ddot{U}\cap O) + P(\ddot{U}^C \cap O) \\ &= P(O/\ddot{U})P(\ddot{U}) + P(O/\ddot{U}^C)P(\ddot{U}^C) \\ &= \frac{47}{100}\frac{40}{100} + \frac{52}{100}\frac{60}{100} = \frac{50}{100} = \frac{1}{2}. \end{aligned}$$

$P(O)$: Ankete katılanlardan rastgele seçilen birinin Obama'ya oy verme olasılığıdır.
Böylece,

$$\begin{aligned} P(\ddot{U}/O) &= \frac{P(\ddot{U}\cap O)}{P(O)} = \frac{P(O/\ddot{U})P(\ddot{U})}{P(O)} \\ &= \frac{\frac{47}{100}\frac{40}{100}}{\frac{50}{100}} \\ &= 0.376 \end{aligned}$$

olarak yani %37.6 bulunur.

Örnek 3: Bir hastalığın teşhisinde kullanılan test ile iligili şu bilgilere sahibiz: Testin doğru bir biçimde hasta olan bir kişiyi pozitif (yani hasta) olarak tanımlama olasılığı %99'dur. Testin doğru bir biçimde hasta olmayan bir kişiyi negatif (yani hasta değil) olarak tanımlama olasılığı %95'tir. Genel popülasyonda bu hastalığın görülme oranı 0.0001'dır. Bu durumda testinin sonucu pozitif olan bir kişinin gerçekten hasta olma olasılığı nedir?

Çözüm: İstenilen olasılık

$$P(\text{Hasta olma } / \text{ Test pozitif}) = \frac{P(\text{Hasta olma } \cap \text{ Test pozitif})}{P(\text{Test pozitif})} = \ ?$$

Kümelerimiz $H = \{\text{Hasta olma}\}$, $H^C = \{\text{Hasta olmayanlar}\}$, $T = \{\text{Testi pozitif}\}$ ve $T^C = \{\text{Testi negatif}\}$ olsun. Bu durumda,

$$P(H/T) = \frac{P(H \cap T)}{P(T)} = \frac{P(T/H)P(H)}{P(T/H)P(H) + P(T/H^C)P(H^C)}$$

biçiminde hesaplarız.

Test ile ilgili bilgilere göre:

- $P(T/H) = 0.99$, $P(T/H^C) = 1 - 0.95 = 0.05$ çünkü $P(T^C/H^C) = 0.95$.
- $P(H) = 0.0001$ ve $P(H^C) = 1 - 0.0001$ dır.

Böylece,

$$
\begin{aligned}
P(H/T) &= \frac{P(T/H)P(H)}{P(T/H)P(H) + P(T/H^C)P(H^C)} \\
&= \frac{0.99\ 0.0001}{(0.99\ 0.0001) + (0.05\ (1 - 0.0001))} \\
&= 0.00197 \approx 0.002
\end{aligned}
$$

bulunur.

Bu sonuca göre test pozitif sonuç verdiğinde kişinin hasta olma olasılığı binde 2'dir. (**Not:** Popülasyonda hastalığın görülme oranı on binde 1'dir.)

Örnek 4: Bir fabrikadaki A, B ve C makinelerinin üretimdeki payları sırasıyla %50, %40 ve %10'dur. Bu makinelerin kusurlu üretim oranları ise sırasıyle %5, %7 ve %1'dir. Bu makinelerden üretilmiş bir ürün rastgele seçiliyor.
a) Seçilen ürünün kusurlu olma olasılığını bulunuz.
b) Kusurlu ürünün A veya C makinelerinde üretilmiş olması olasılığını bulunuz.

Çözüm: Kümelerimiz $K = \{$Seçilen ürünün kusurlu olması$\}$, $A = \{$A makinesinde üretilen ürünler$\}$, $B = \{$B makinesinde üretilen ürünler$\}$ ve $C = \{$C makinesinde üretilen ürünler$\}$ olsun.

a) Toplam olasılık formülü ile

$$\begin{aligned} P(K) &= P(K \cap A) + P(K \cap B) + P(K \cap C) \\ &= P(K/A)P(A) + P(K/B)P(B) + P(K/C)P(C) \\ &= \frac{5}{100}\frac{50}{100} + \frac{7}{100}\frac{40}{100} + \frac{1}{100}\frac{10}{100} \\ &= 0.054 \end{aligned}$$

b) Kusurlu ürünün A makinesinde üretilme olasılığı:

$$\begin{aligned} P(A/K) &= \frac{P(A \cap K)}{P(K)} = \frac{P(K/A)P(A)}{P(K)} \\ &= \frac{\frac{5}{100}\frac{50}{100}}{0.054} = 0.463. \end{aligned}$$

Kusurlu ürünün C makinesinde üretilme olasılığı:

$$P(C/K) = \frac{P(K/C)P(C)}{P(K)} = \frac{10/10000}{0.054} = 0.0185.$$

Kusurlu ürünün A veya C makinesinde üretilme olasılığı:
$P(A/K) + P(C/K) = 0.4815$ bulunur.

Gelecek hafta: Rastgele değişkenler