Olasılık ve İstatistik

Fırat İsmailoğlu, PhD

Ayrık Rastgele Değişkenler- 2

Ortak Olasılık Dağılımı (Joint Probability Distribution)

Bazen bir deneyde tek bir rastgele değişkenin alabileceği değerle ilgilenmeyiz. Birden fazla rastgele değişkenin birbirleriyle olan ilişkisi ile ilgileniriz.

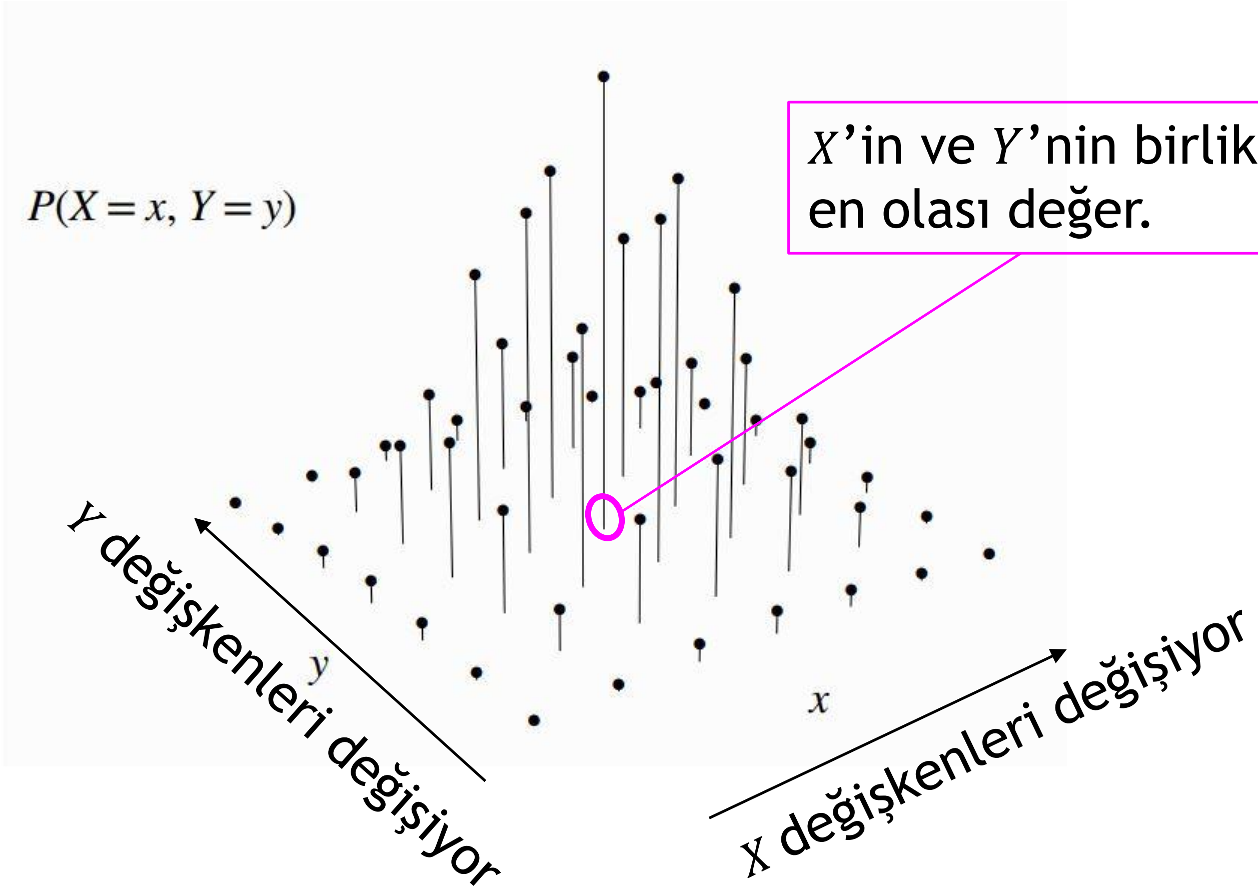
Örneğin kişinin kanser teşhisi konulduğu yaşı ile kanserin ciddiyeti arasında bir ilişki arayabiliriz, yada üniveritedeki yıl ve alınan ders sayısı arasında bir ilişki.

X ve Y iki rastgele değişken olsun. $X: x_1,\ x_2, \dots, x_n$ değişkenlerini, $Y: y_1,\ y_2, \dots, y_m$ değişkenlerini alsın. Bu durumda, X'in x_i değişkenini **ve** Y'nin y_j değişkenini alma olasılığı:

$$\boxed{P(X = x_i, Y = y_j)}$$

$P(X = x_i, Y = y_j)$ olasılıklarına X'in ve Y'nin ortak olasılık dağılımı diyeceğiz.

X'in ve Y'nin birlikte alabileceği en olası değer.



$$\boxed{P(X = x_i, Y = y_j) = P(X = x_i) \times P(Y = y_j | X = x_i)}$$

(çarpım kuralından)

Ortak Olasılık Dağılımı (Joint Probability Distribution)

ör. Bir zar atılma deneyinde hem zarın tek gelmesiyle hemde zarın asal gelmesi ile ilgilenelim. X rastgele değişkeni zar tekse 1, değilse 0 değerini alsın; Y rastgele değişkeni zar asalsa 1 değilse 0 değerini alsın.

$$X = \begin{cases} 1, zar\ tekse \\ 0, zar\ tek\ değilse \end{cases} \qquad Y = \begin{cases} 1\ , zar\ asalsa \\ 0, zar\ asal\ değilse \end{cases}$$

Bu iki rastgele değişken birleşerek aşağıdaki gibi bir ortak olasılık dağılımı oluştururlar:

	$Y = 1$	$Y = 0$
$X = 1$	$P(X = 1, Y = 1)$	$P(X = 1, Y = 0)$
$X = 0$	$P(X = 0, Y = 1)$	$P(X = 0, Y = 0)$

X'in 1 ve Y'nin 0 almasının olasılığı. Kendine sor: ne zaman hem X,1 değerini alır, hem de Y, 0 değerini alır. Bu zarın 1 gelmesiyle mümkündür.

O halde olasılık $\frac{1}{6}$'dır.

Ikinci bir yöntem olarak çarpım kuralı ile bulabiliriz:

$$P(X = 1, Y = 0) = P(X = 1)P(Y = 0|X = 1) = \frac{3}{6} \cdot \frac{1}{3} = \frac{1}{6}$$

Ortak Olasılık Dağılımı (Joint Probability Distribution)

$$X = \begin{cases} 1, zar\ tekse \\ 0, zar\ tek\ değilse \end{cases} \qquad Y = \begin{cases} 1, zar\ asalsa \\ 0, zar\ asal\ değilse \end{cases}$$

	$Y = 1$	$Y = 0$	Toplam
$X = 1$	$P(X = 1, Y = 1) = \frac{2}{6}$	$P(X = 1, Y = 0) = \frac{1}{6}$	$P(X = 1) = \frac{3}{6}$
$X = 0$	$P(X = 0, Y = 1) = \frac{1}{6}$	$P(X = 0, Y = 0) = \frac{2}{6}$	$P(X = 0) = \frac{3}{6}$
Toplam	$P(Y = 1) = \frac{3}{6}$	$P(Y = 0) = \frac{3}{6}$	1

Burası tek başına (X'e bağımlı olmaksızın) Y'nin 1 olma olasılığı ör. Bir fakültedeki öğrencilerin %75'i yerel, %25'i başka şehirlerden gelsin. Yerel öğrencilerin %60'ı bilgisayar mühendisliği, %30'u makine mühendisliği ve %10'u inşaat mühendisliği bölümünde olsun. Başka şehirlerden gelen öğrencilerin %30'u bilgisayar mühendisliği, %40'ı makine ve %30'u inşaat okusun. X rastgele değişkeni öğrencinin yerel mi dışarıdan mı olduğunu; Y rastgele değişkeni öğrencinin bölümünü göstersin. Buna göre X ve Y rastgele değişkenlerinin ortak olasılık dağılımı nasıl olur?

Çözüm.

Y'nin alabileceği değerler

X'in alabileceği değerler

	Bilgisayar	Makine	İnşaat
Yerel	$P(X=Yer, Y=Bil)=$ $P(X=Yer)P(Y=Bil \mid X=Yer)$ $=0.75\times 0.6=0.45$	$P(X=Yer, Y=Mak)=$ $P(X=Yer)P(Y=Mak \mid X=Yer)$ $=0.75\times 0.3=0.225$	$P(X=Yer, Y=\dot{I}nş)=$ $P(X=Yer)P(Y=\dot{I}ns \mid X=Yer$ $=0.75\times 0.1=0.075$
Dış.	$P(X=Dış, Y=Bil)=$ $P(X=Dış)P(Y=Bil \mid X=Dış)$ $=0.25\times 0.3=0.075$	$P(X=Dış, Y=Mak)=$ $P(X=Dış)P(Y=Mak \mid X=Dış)$ $=0.25\times 0.4=0.1$	$P(X=Dış, Y=\dot{I}nş)=$ $P(X=Dış)P(Y=\dot{I}ns \mid X=Dış)$ $=0.25\times 0.3=0.075$

Rastgele seçilen birinin yerel olması ve bilgisayar mühendisliği bölümünde olması olasılığı

	Bilgisayar	Makine	İnşaat	Toplam
Yerel	0.45	0.225	0.075	0.75
Dışarıdan	0.075	0.1	0.075	0.25
Toplam	0.525	0.325	0.15	1

Rastgele seçilen birinin makine mühendisliği bölümünde olması olasılığı. Dikkat edilirse bu, soruda açıkça verilmemiş, ortak olasılıklardan kendimiz hesapladık.

Marjinal Olasılık

Marjinal olasılık, ortak olasılıklar kullanılarak bulunan tek bir rastgele değişkenin belirli bir değeri alma olasılığıdır.

$$P(X = x_i) = \sum_{j=1}^{m} P(X = x_i, Y = y_j) \qquad P(Y = y_j) = \sum_{i=1}^{n} P(X = x_i, Y = y_j)$$

Marjinal Olasılık

$P(X = x_i) = \sum_{j=1}^{m} P\left(X = x_i, Y = y_j\right)$ marjinal olasılığı hesaplanırken X rastgele değişkeninin x_i değerini alması **sabit** iken, Y'nin diğer değerleri alma olasılıklarının toplamını buluyoruz.

Bir önceki örnekten Y'nin makine mühendiliği olması marjinal olasılığını hesaplayalım. Bir öğrenci yerel iken yada dışarıdan gelmiş iken makine mühendisi olabilir. Yani burada değişen şey öğrencinin nereden geldiğidir, makine mühendisliği sabittir.

$P(Y = makine) = P(X = yerel, Y = makine) + P(X = dışarıdan, Y = makine) =$

$= 0.225 + 0.1 = 0.325$

	Bilgisayar	Makine	İnşaat	Toplam
Yerel	0.45	0.225	0.075	0.75
Dışarıdan	0.075	0.1	0.075	0.25
Toplam	0.525	0.325	0.15	1

X'in marjinal dağılımı

Y'nin marjinal dağılımı

ör.

	Sigara İçmeyen	Sigara İçen	Toplam
Kanser Değil	40	10	50
Kanser	7	3	10
Toplam	47	13	60

60 kişilik bir grubun kanser olma ve sigara içme durumlarına göre dağılımı.

Ortak olasılık dağılımı	Sigara İçmeyen	Sigara İçen	Toplam
Kanser Değil	40/60	10/60	50/60
Kanser	7/60	3/60	10/60
Toplam	47/60	13/60	1

$$P(kanser\ degil\ ,sigara\ icen)$$
$$= P(kanser\ degil)P(sigara\ icen|kanser\ degil)$$
$$\frac{50}{60}\times\frac{40}{50}=\frac{40}{60}$$

$$P(kanser) = P(kanser, sigara\ icmeyen) + P(kanser, sigara\ icen)$$
$$= \frac{7}{60}+\frac{3}{60}=\frac{10}{60}$$

Kanserin marjinal olasılığı

Bazı Özel Ayrık Rastgele Değişkenler

1) Bernoulli Rastgele Değişkeni

Bir deneyde yalnızca 2 sonuç varsa (başarı, başarısızlık) bu bir Bernoulli deneyidir.

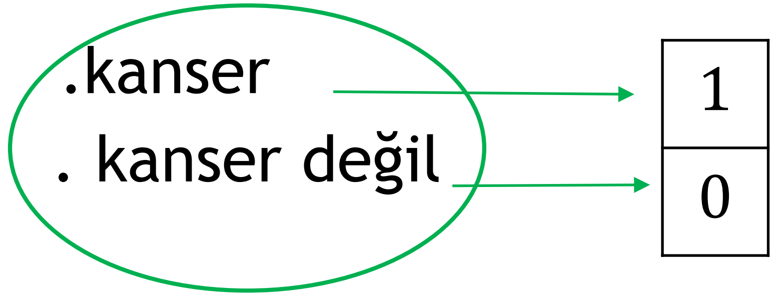
Bernoulli deneyleri örnekleri:

- Seçilen kişi başarılı, başarısız.
- Seçilen kişi kanser, kanser değil.
- Atılan bozuk para yazı, tura.
- Hava yağmurlu, kuru.
- Bir ilacın yan etkisi var, yan etkisi yok …

Bernoulli deneyinde 2 sonuç olduğu için, bu sonuçların olasılıkları p ve $1-p$ dir.

Bernoulli deneyinde ‘*başarı*’ sonucunu 1’e; ‘*başarısızlık*’ sonucunu 0’a götüren rastgele değişkene özel olarak Bernoulli Rastgele Değişkeni diyeceğiz.

Bir kişiyi
seçme deneyi



S: örnek uzay

X, Bernoulli rastgele değişkeni 1 yada 0 değerini alır.

Bu değişkenin $X = 1$ değerini alması (yani başarı olma) olasılığı: $P(X = 1) = p$

$X = 0$ değerini alması (yani başarısızlık olma) olasılığı: $P(X = 0) = 1 - p$

Bernoulli rastgele değişkeninin dağılımına yani $P(X = 1)$ ve $P(X = 0)$ değerlerine Bernoulli olasılık dağılımı denir.

ör. Bir toplumda sigara içme oranı yüzde 20’dir. Bu durumda Bernoulli rastgele değişkeninin olasılık dağılımı ne olur?

Seçilen kişi sigara içer ise $X = 1,\ \ P(X = 1) = 0.2$

Seçilen kişi sigara içer ise $X = 0,\ \ P(X = 0) = 0.8$

Bernoulli Rastgele Değişkeninin Beklenen Değeri

X, bir Bernoulli rastgele değişkeni olsun. Şu halde X'in alabileceği değerler 1 ve 0 dır. Ayrıca $P(X=1)=p$ ve $P(X=0)=1-p$ dir.

$$E[X]=1\times p+0\times(1-p)=p$$
$$E[X]=p$$

Bir Bernoulli rastgele değişkenin beklenen değeri olma olasılığnın ta kendisidir.

2) Binomal Rastgele Değişkeni

n tane birbirinden *bağımsız* Bernoulli deneyinde (yani iki sonuçlu deneyde) toplam başarı sayısı binomial rastgele değişkeni verir.

Burada X rastgele değişkeni $0,1,\ldots,n$ değerlerini alabilir. Bu değerlerden birini alması olasılığı $(k\ \in\ \{0,1,\ldots,n\}\ iken)$

$$\boxed{P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}}$$

k defa başarı isterken, $(n-k)$ defa başarısızlık istediğimizi unutmayalım. n tane Bernoulli deneyinden $\binom{n}{k}$ tane farklı şekilde k başarı seçilebilir.

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

k defa başarı olasılığı

k başarının kaç farklı şekilde seçilebileceği

k defa başarı

$n-k$ defa başarısızlık

ör. Bir A takımının, bir B takımını yenme olasılığı 0.3 olsun. Bu iki takım arasında yapılan 5 maçın 3'ünü A'nın yenmesi olasılığı nedir?

Çözüm.

X rastgele değişkeni 5 maçtan, A'nın toplam başarı sayısını göstersin. $P(X=3)$ ile ilgileniyoruz.

$$P(X=3)=\binom{5}{3}0.3^3 0.7^2=0.132$$

ör. 3 tane birbirinden bağımsız her birinde başarı oranı p olan Bernoulli deneylerinde 2 tane başarı olma olasılığı:

İlgilendiğimiz sonuçlar:

			Olasılık:
başarı	başarı	başarısızlık	$p \times p \times (1-p) = p^2(1-p)$
başarı	başarısızlık	başarı	$p \times (1-p) \times p = p^2(1-p)$
başarısızlık	başarı	başarı	$(1-p) \times p \times p = p^2(1-p)$

$$+$$

$$3 \times p^2(1-p)$$

$$\binom{3}{2}$$

(kaç farklı şekilde başarı olacağı)

ör. 20 kişilik bir toplulukta 8 kişi araba kullanmayı biliyor olsun. Rastgele seçilen 5 kişiden ikisinin araba kullanma olasılığı nedir?

Çözüm.

X seçilen 5 kişiden araba kullanmayı bilen toplam kişi sayısı olsun.

p: seçilen bir kişinin araba kullanma olasılığı: $\frac{8}{20} = 0.4$

$$P(X = 2) = \binom{5}{2} 0.4^2 (1 - 0.4)^3 = 0.3456$$

Binomial Rastgele Değişkenin Beklenenen Değeri

$$E[X] = \sum_{i=0}^{n} i \times P(X = i)$$

$$E[X] = \sum_{i=0}^{n} i \times \binom{n}{i} p^i (1-p)^{n-i}$$

$$E[X] = np$$

Burada $E[X], n$ deneyde beklenen toplam başarı sayısı olarak yorumlanır.

ör. 100 defa atılan düzgün bir bozuk paranın toplam kaç defa tura gelmesi beklenir?

Çözüm.

X rastgele değişkeni 100 atıştaki toplam tura sayısını versin. Her bir atış bir Bernoulli deneyidir, iki sonuç vardır, ve deneyler birbirinden bağımsızdır. Her bir deneydeki başarı olasılığı $p = 0.5'$tir.

X'in beklenen değeri 100 atıştaki beklenen toplam başarı sayısını verir.

$$E[X] = 100 \times 0.5 = 50$$

100 atışta 50 defa başarı olması (tura gelmesi) beklenir.

ör. Yeni geliştirilen bir korona virüs aşısının 0.95 olasılıkla fayda sağladığı biliniyor.

8 kişiye bu aşı yapılıyor. Bu kişilerden en az ikisinin bu aşıdan fayda görme olasılığı nedir?

Çözüm.

X rastgele değişkeni ilaçtan fayda gören kişi sayısını göstersin. Bu halde X, 0-8 arası bir tam sayıdır. Biz X'in 2 veya 8 arasında (2 ve 8 dahil) bir sayı olmasıyla ilgileniyoruz.

Aradığımız olasılık

$$P(X=2)+P(X=3)+\cdots+P(X=8)$$

‘dir. Burada toplamdaki her bir olasılığı binomial olasılıkla bulabilriz. Burdan ilerlersek 7 defa olasılık hesaplamamız gerekir.

Fakat öte yandan

$$P(X=0)+P(X=1)+P(X=2)+\cdots+P(X=8)=1$$

olduğunu biliyoruz. Aradığımız olasılık aynı zamanda

$$1-P(X=0)+P(X=1)$$

olasılığına eşittir.

$$P(X=0)=\binom{8}{0}0.95^{0}(1-0.95)^{\ 8}$$

$$P(X=1)=\binom{8}{1}0.95^{1}(1-0.95)^{\ 7}$$

ör.Bir önceki soru için en fazla 2 kişinin aşıdan fayda görmeme olasılığı nedir?

Çözüm.

Fayda görmeme olasılılığı $1-0.95=0.05$

X rastgele değişkeni fayda görmeyen kişi sayısını göstersin. Biz bunun 0 veya 1 veya 2 olması olasılığı ile ilgileniyoruz:

$$P(X=0)+P(X=1)+P(X=2)$$

$$\binom{8}{0}0.05^0(1-0.05)^{8}+\binom{8}{1}0.05^1(1-0.05)^{7}+\binom{8}{2}0.05^2(1-0.05)^{6}$$