RD'ler İçin Bağımsızlık

- Bir RD'nin bir olaydan bağımsız olması için aşağıdaki şartın sağlanması gerekir.

$$P(X = x \; ve \; A) = P(X = x)P(A) = p_X(x)P(A)$$

- İki RD'nin bağımsız olması için ise kesikli ve sürekli için şu şartlar sağlanmalıdır.

$$p_{X,Y}(x, y) = p_X(x)p_Y(y)$$

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

- İki RD bağımsız ise çarpımların beklenen değeri, beklenen değerlerin çarpımına eşittir, yani

$$\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$$

Bağımsız RD'ler için Beklenen Değer

- Önceki bağıntı şu şekilde ispat edilebilir.

$$\mathsf{E}[XY] = \sum_x \sum_y xyp_{X,Y}(x,y) = \sum_x \sum_y xyp_X(x)p_Y(y)$$

$$= \sum_x xp_X(x) \sum_y yp_Y(y) = \mathsf{E}[X]\mathsf{E}[Y]$$

- Genellersek, iki bağımsız X ve Y gibi iki RD için

$$\mathsf{E}[g(X)h(Y)] = \mathsf{E}[g(X)]\mathsf{E}[h(Y)]$$

- Sürekli RD'ler için de benzer bir ispat yazılabilir.

Bağımsız RD'ler için Varyans/Sapma ve İkiden Fazla RD için Bağımsızlık

- İki bağımsız X ve Y gibi RD'nin toplamının varyansına bakalım

$$\sigma^2_{X+Y} = \mathrm{E}[(X+Y)^2] - \mathrm{E}[X+Y]^2$$

$$= \mathrm{E}[X^2] + \mathrm{E}[Y^2] + \mathrm{E}[2XY] - \mathrm{E}[X]^2 - \mathrm{E}[Y]^2 - 2\mathrm{E}[X]\mathrm{E}[Y]$$

$$= \mathrm{E}[X^2] + \mathrm{E}[Y^2] + 2\mathrm{E}[X]\mathrm{E}[Y] - \mathrm{E}[X]^2 - \mathrm{E}[Y]^2 - 2\mathrm{E}[X]\mathrm{E}[Y] = \sigma_x^2 + \sigma_y^2$$

- Sürekli RD'ler için de benzer bir ispat yapılabilir.

- Birden fazla RD için bağımsızlık benzer şekilde tanımlanır

$$p_{x_1,x_2,...,x_n}(x_1, x_2, ..., x_n) = p_{x_1}(x_1)p_{x_2}(x_2)...p_{x_n}(x_n)$$

$$f_{x_1,x_2,...,x_n}(x_1, x_2, ..., x_n) = f_{x_1}(x_1)f_{x_2}(x_2)...f_{x_n}(x_n)$$

İkiden Fazla RD için Bağımsızlık (Devam)

- Beklenen değer ve varyans kuralları da genellenebilir.

$$\mathrm{E}[X_1 X_2 ... X_n] = \mathrm{E}[X_1]\mathrm{E}[X_2]...\mathrm{E}[X_n]$$

$$\sigma^2_{X_1+X_2+...+X_n} = \sigma^2_{X_1} + \sigma^2_{X_2} + ... + \sigma^2_{X_n}$$

Sürekli Bayes Kuralı ve Çıkarım (Inference)

- Kesikli olana benzer şekilde sürekli Bayes Kuralı şu şekilde verilir:

$$f_{X/Y}(x/y) = \frac{f_X(x) f_{Y/X}(y/x)}{f_Y(y)}$$

- Bazen aşağıdaki $f_Y(y)$ için toplam olasılık kullanılır ve şu ifade elde edilir

$$f_{X/Y}(x/y) = \int \frac{f_X(x) f_{Y/X}(y/x)}{f_{Y/X}(y/x') f_X(x')\, dx'}$$

- Sürekli Bayes Kuralı sıkça bir gözlemden direkt olarak gözlemlenemeyen değişkenlerin tespitinde kullanılır.

Inference (Çıkarım)

- Böyle bir problem için, x bulmak istediğimiz gizli değişken, y ise bizim gözlemlediğimiz x ile ilişkili bir değişkendir.

- Biz herhangi bir bilinmeyen x'e bağlı olarak y'yi gözlemlediğimizden elimizde $f_{Y/X}(y/x)$ vardır. Biz ise ölçtüğümüz bir Y için X'i yani $f_{X/Y}(x/y)$'yi bulmak istiyoruzdur.

- Biz gözlemden gizli bir değişkenin bu şekilde bulunmasına çıkarım/inference denir.

- Kesikli Hali ise şu şekilde verilir

$$p_{N/Y}(n/y) = \frac{p_N(n) f_{Y/N}(y/n)}{\sum_i f_{Y/N}(y/i) p_N(i)}$$

Örnek

Tekrar tekrar bir yazılım programı yazmaktasın, ve her defasında önceki denemelerden bağımsız olarak p olasılık ile program düzgün çalışmaktadır. Program düzgün çalışana kadar yapılan denem sayısına X dersek, X'in beklenen değeri ve varyansı ne olur?

X'i geometrik rastgele değişken olarak tanımaktayız, PMF'i:

$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1, 2, ...$$

X'in beklenen değeri ve varyansı:

$$\mathbf{E}[X] = \sum_{k=1}^{\infty} k(1-p)^{k-1}p, \qquad var(X) = \sum_{k=1}^{\infty} (k - \mathbf{E}[X])^2(1-p)^{k-1}p,$$

fakat bu sonsuz toplamlar bir miktar can sıkıcıdır. Alternatif olarak, toplam beklenti teoremini uygulayacağız; $A_1 = \{X = 1\} =$ {ilk deneme başarılı} $A_2 = \{X > 1\} =$ {ilk deneme başarısız} ve daha basit işlemler ile sonuçlanır.

Eğer ilk deneme başarılı ise $X = 1$, ve

$$\mathbf{E}[X|X = 1] = 1.$$

Eğer ilk deneme başarısız ise $X > 1$, ilk denemeyi boşa harcamışız demek ve başladığımız yere geri dönmüşüzdür. Yani kalan denemelerin beklenen değeri $\mathbf{E}[X]$'dir ve

$$\mathbf{E}[X|X > 1] = 1 + \mathbf{E}[X].$$

Böylece,

$$\begin{aligned}\mathbf{E}[X] &= \mathbf{P}(X = 1)\mathbf{E}[X|X = 1] + \mathbf{P}(X > 1)\mathbf{E}[X|X > 1] = 1 \\ &= p + (1 - p)(1 + \mathbf{E}[X]),\end{aligned}$$

buradan:

$$\mathbf{E}[X] = \frac{1}{p}.$$

Aynı yöntem ile,

$$\mathbf{E}[X^2|X = 1] = 1, \qquad \mathbf{E}[X^2|X > 1] = \mathbf{E}[(1+X)^2] = 1+2\mathbf{E}[X]+\mathbf{E}[X^2]$$

buradan:

$$\mathbf{E}[X^2] = p * 1 + (1-p)(1 + 2\mathbf{E}[X] + \mathbf{E}[X^2]),$$

ve

$$\mathbf{E}[X^2] = \frac{1 + 2(1-p)\mathbf{E}[X]}{p},$$

ve $\mathbf{E}[X] = 1/p$ eşitliğini kullarak şunu elde ederiz:

$$\mathbf{E}[X^2] = \frac{2}{p^2} - \frac{1}{p}.$$

Son olarak:

$$var(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Örnek

Bir madeni para n defa atılmaktadır ve p olasılıkla yazı gelmektedir. Yazı gelen para sayısının varyansı nedir?
Her i değeri için X_i yazı geldiğinde 1 tura geldiğinde 0 değerini alan Bernoulli rastgele değişkenidir. Öyleyse $X = X_1 + X_2 + ... + X_n$ binominal rastgele değişkendir. Örnek 2.10'da bulunduğu gibi $\mathbf{E}[X] = np$'dir. Para atışlarının bağımsızlığına dayanarak, $X_1, X_2, ..., X_n$ rastgele değişkenleride bağımsızdır ve

$$var(X) = \sum_{i=1}^{n} var(X_i) = np(1-p).$$

Bölüm 2.2'de tartıştığımız gibi, λ parametreli bir Poisson rastgele değişkeni (Y) binominalın "limiti", $n \to \infty$, $p \to 0$, olarak görülebilir, $np = \lambda$ iken. Böylece, binominalin beklenen değerinin ve varyansının limitini alınca, Poisson'un beklenen değerini ve varyansını elde etmiş oluruz, $\mathbf{E}[Y] = var(Y) = \lambda$. Örnek 2.7'de $\mathbf{E}[Y] = \lambda$ formülünü doğrulamıştık.

$var(Y) = \lambda$ formülü ispatlamak için:

$$\begin{aligned}
\mathbf{E}[Y^2] &= \sum_{k=1}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \lambda \sum_{k=1}^{\infty} k \frac{e^{-\lambda} \lambda^{k-1}}{(k-1)!} \\
&= \lambda \sum_{m=0}^{\infty} (m+1) \frac{e^{-\lambda} \lambda^m}{m!} \\
&= \lambda(\mathbf{E}[Y] + 1) \\
&= \lambda(\lambda + 1),
\end{aligned}$$

ve,

$$var(Y) = \mathbf{E}[Y^2] - (\mathbf{E}[Y])^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda.$$

Örnek

Bir mesaj -1 veya +1 değerlerini alan s sinyali olarak başka bir noktaya gönderiliyor. Haberleşme kanalında normal dağılıma sahip bir gürültü (beklenen değer $\mu = 0$ ve varyans σ^2) ile bozulmaktadır. Alıcı tarafına gelen değer eğer < 0 (veya > 0) ise mesajı -1 (veya +1) değeri olarak bulur. Hata olasılığı nedir?

Hata, -1 gönderildiğinde ve gürültü (N) en az 1 olduğunda oluşur, $s + N = -1 + N \geq 0$. Ya da +1 gönderildiğinde ve gürültü (N) -1'den daha düşük olduğunda oluşur, $s + N = 1 + N < 0$. İlk durumda hata olasılığı:

$$\begin{aligned}\mathbf{P}(N \geq 1) &= 1 - \mathbf{P}(N < 1) = 1 - \mathbf{P}(\frac{N-\mu}{\sigma} < \frac{1-\mu}{\sigma}) \\ &= 1 - \Phi(\frac{1-\mu}{\sigma}) = 1 - \Phi(\frac{1}{\sigma}).\end{aligned}$$

Diğer durumda, hata olasılığı simetriden dolayı eşittir. $\Phi(1/\sigma)$ normal tablosundan elde edilebilir. $\sigma = 1$ olunca $\Phi(1/\sigma) = \Phi(1) = 0.8431$, ve hata olasılığı 0.1587 olur.

Örnek

Polis radarından geçen bir aracın hızı beklenen değeri 50 mil/saat olan üssel dağılımlı rastgele değişken, X, olarak modellenmiştir. Polis radarındaki araç hızının ölçümü, Y, beklenen değeri 0 ve standart sapması aracın hızının onda biri olan bir normal rastgele değişkendir. X ve Y'nin ortak PDF'i nedir?

$f_X(x) = (1/50)e^{-x/50}, x \geq 0$ için, ve $X = x$'e koşullanmış Y ölçümünün beklenen değeri x ve varyansı $x^2/100$ olan normal PDF'i vardır:

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}(x/10)} e^{-(y-x)^2/(2x^2/100)}.$$

Böylece bütün $x \geq 0$ ve bütün y değerleri için,

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x) = \frac{1}{50} e^{-x/50} \frac{10}{\sqrt{2\pi}x} e^{-50(y-x)^2/x^2}.$$

Örnek

İkilik bir S sinyali iletilecektir ve $\mathbf{P}(S=1)=p$ ve $\mathbf{P}(S=-1)=1-p$'dir. Alınan sinyal $Y=N+S$'dir, N beklenen değeri 0 ve birim varyansı olan bir normal gürültüdür, S'den bağımsızdır. $S=1$ olma olasılığını, elde ettiğimiz $Y=y$ sinyalinin bir fonksiyonu olarak belirtiniz.

$S=s$'ye koşullu Y rastgele değişkeni s beklenen değerli ve birim varyanslı normal dağılımdır. Yukarıda verilen formülleri uygularsak,

$$\mathbf{P}(S=1|Y=y)=\frac{p_s(1)f_{Y|S}(y|1)}{f_Y(y)}=\frac{\frac{p}{\sqrt{2\pi}}e^{-(y-1)^2/2}}{\frac{p}{\sqrt{2\pi}}e^{-(y-1)^2/2}+\frac{1-p}{\sqrt{2\pi}}e^{-(y+1)^2/2}},$$

sadeleştirilirse:

$$\mathbf{P}(S=1|Y=y)=\frac{pe^y}{pe^y+(1-p)e^{-y}}.$$

y değeri $-\infty$'a yaklaştırılırsa $\mathbf{P}(S=1|Y=y)$ olasılığının 0'a gittiği görülmektedir. y değeri $+\infty$'a yaklaştırılırsa $\mathbf{P}(S=1|Y=y)$ olasılığı bu sefer 1'e gider ve iki değer arasında monoton olarak artmaktadır. Bu sezgilerimizle uyumludur.