İki Değişkenli Karma (Mixed) Dağılımlar

Bu bölümde buraya kadar kesikli ya da sürekli iki değişkenli dağılımlar üzerinde duruldu. Bazen, rastgele değişkenlerden birinin kesikli diğerinin ise sürekli dağılıma sahip olduğu iki değişkenli karışık dağılımlarla karşılaşılabilir. Böyle bir ortak olasılık fonksiyonunu $f_{X,Y}(x,y)$ ile gösterilir. Bu durumda (X,Y) çiftinin xy −düzleminde belirli bir bölgeye ait olma olasılığını, değişkenlerden birisi için $f_{X,Y}(x,y)$ değerleri toplanarak ve diğer değişken için de $f_{X,Y}(x,y)$'nin integrali alınarak bulunur. Her ortak olasılık fonksiyonu iki koşulu sağlamalıdır. Eğer X mümkün değerleri $x_1, x_2, \ldots$ olan kesikli bir rastgele değişken ve Y bir sürekli rastgele değişken ise, bu durumda $\forall\ x, y$ için

$$f_{X,Y}(x,y) \geq 0 \quad \text{ve} \quad \int_{-\infty}^{\infty} \sum_{i=1}^{\infty} f_{X,Y}(x_i,y)dy = 1$$

olmalıdır. f fonksiyonu negatif olmadığından, işlem kolaylığı sağlayacaksa integral ve toplama işlemlerinin sırası değiştirilebilir.

Örnek: (X,Y) iki boyutlu rastgele değişkenin ortak olasılık fonksiyonu aşağıdaki gibi verilmiş olsun,

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{xy^{x-1}}{3} & , \quad x = 1,2,3 \ \text{ ve } \ 0 < y < 1 \\ 0 & , \quad \text{d. d.} \end{cases}$$

Bu fonksiyon eşitliğini sağlamalıdır, ilk önce y değerleri üzerinden integral almak kolay olacağından,

$$\sum_{x=1}^{3} \int_{0}^{1} \frac{xy^{x-1}}{3} dy = \sum_{x=1}^{3} \frac{1}{3} = 1$$

gösterilir. $(X \geq 2,\ Y \geq 1/2)$ olayının olasılığı

$$P(X \geq 2,\ \ Y \geq 1/2) = \sum_{x=2}^{3} \int_{1/2}^{1} \frac{xy^{x-1}}{3} dy = \sum_{x=2}^{3} \frac{1 - \left(\frac{1}{2}\right)^x}{3} = 0{,}5417$$

olarak bulunur.

Koşullu Dağılımlar

Koşullu Olasılık Fonksiyonu

Tanım. (X,Y) iki boyutlu rastgele değişkeninin ortak olasılık fonksiyonu $P_{X,Y}(x,y)$ ve $X = x$ verildiğinde Y rastgele değişkeninin koşullu olasılık fonksiyonu

$$P_{Y|X}(y) = \frac{P_{X,Y}(x,y)}{P_X(x)}, \quad P_X(x) > 0$$

şeklindedir. $Y = y$ verildiğinde X rastgele değişkeninin koşullu olasılık fonksiyonu

$$P_{X|Y}(x) = \frac{P_{X,Y}(x,y)}{P_Y(y)}, \qquad P_Y(y) > 0$$

şeklindedir.

$P_{Y|X}(y)$ koşullu olasılık fonksiyonunun Özellikleri

1) $0 \leq P_{Y|X}(y) \leq 1, \quad \sum_{D_Y} P_{Y|X}(y) = 1$

2) Eğer X ve Y rastgele değişkenleri bağımsız ise,

$$P_{X|Y}(x) = P_X(x), \qquad P_{Y|X}(y) = P_Y(y)$$

Örnek: (X, Y) iki boyutlu kesikli rastgele değişkenin ortak olasılık fonksiyonu aşağıdaki gibidir.

$$P_{X,Y}(x,y) = \begin{cases} \left(\frac{4}{5}\right)^{x+y} \left(\frac{1}{5}\right)^{2-x-y} & , \quad x = 0,1; \; y = 0,1 \\ 0 & , \quad \text{d.d.} \end{cases}$$

a) $P_{X|Y}(x)$ koşullu olasılık fonksiyonunu bulunuz.

b) X' in marjinal olasılık fonksiyonunu bulunuz.

Çözüm.

a) Y rastgele değişkeninin marjinal olasılık yoğunluk fonksiyonu,

$$P_Y(y) = \sum_{D_X} P_{X,Y}(x,y)$$

eşitliğinden,

$$P_Y(y) = \left(\frac{4}{5}\right)^{y} \left(\frac{1}{5}\right)^{2-y} + \left(\frac{4}{5}\right)^{y+1} \left(\frac{1}{5}\right)^{1-y}, y = 0,1 \text{ için}$$

dir.

$$P_{X|Y}(x) = \frac{\left(\frac{4}{5}\right)^{x+y} \left(\frac{1}{5}\right)^{2-x-y}}{\left(\frac{4}{5}\right)^{y} \left(\frac{1}{5}\right)^{1-y}} = \left(\frac{4}{5}\right)^{x} \left(\frac{1}{5}\right)^{1-x}$$

$$P_{X|Y}(x) = \begin{cases} \left(\frac{4}{5}\right)^{x} \left(\frac{1}{5}\right)^{1-x} & , \quad x = 0,1 \\ 0 & , \quad \text{d.d.} \end{cases}$$

olarak bulunur.

b)

$$P_X(x) = \sum_{D_Y} P_{X,Y}(x,y)$$

$$P_X(x) = \left(\frac{4}{5}\right)^x \left(\frac{1}{5}\right)^{1-x}, x = 0,1 \text{ için}$$

$P_{X|Y}(x) = P_X(x)$ olduğundan X ve Y rastgele değişkenleri bağımsızdır.

Koşullu Olasılık Yoğunluk Fonksiyonu

Tanım. (X, Y) iki boyutlu rastgele değişkeninin ortak olasılık yoğunluk fonksiyonu $f_{X,Y}(x,y)$ ve $X = x$ verildiğinde Y rastgele değişkeninin koşullu olasılık yoğunluk fonksiyonu

$$f_{Y|X}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}, \qquad f_X(x) > 0$$

şeklindedir. $Y = y$ verildiğinde X rastgele değişkeninin koşullu olasılık yoğunluk fonksiyonu

$$f_{X|Y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \qquad f_Y(y) > 0$$

dir.

Koşullu olasılık yoğunluk fonksiyonu $f_{Y|X}(y)$'nin Özellikleri

a) $f_{Y|X}(y) \geq 0, \qquad \int_{-\infty}^{\infty} f_{Y|X}(y)\, dy = 1$

b) Eğer X ve Y rastgele değişkenleri bağımsız ise,

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

Benzer olarak, yukarıdaki özellikler $f_{X|Y}(x)$ içinde geçerlidir.

Kovaryans ve Korelasyon Katsayısı

Tanım. (Ω, U, P) olasılık uzayı, X ve Y bu uzayda tanımlı iki rastgele değişken olsun. (X, Y) iki boyutlu rastgele değişkeninin m.inci ve n.inci momenti hem kesikli hem sürekli rastgele değişkenler için

$$\mu_{mn} = E(X^m Y^n) = \begin{cases} \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} x^m y^n p_{X,Y}(x,y), \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^m y^n f_{X,Y}(x,y)\, dx\, dy, \end{cases} \qquad m,n \in \mathbb{N}$$

olarak tanımlanır. Bu tanımla ilgili aşağıdaki sonuçlar elde edilir:

1) Eğer $n = 0$ alınırsa X rastgele değişkeninin m.inci momenti elde edilir.

2) Eğer $m = 0$ alınırsa Y rastgele değişkeninin n.inci momenti elde edilir.

3) Eğer $n = 0,\ m = 1$ alınırsa $\mu_{10} = E(X) = \mu_X$

4) Eğer $n = 1,\ m = 0$ alınırsa $\mu_{01} = E(Y) = \mu_Y$

5) Eğer $n = 0,\ m = 2$ alınırsa $\mu_{20} = E(X^2)$

6) Eğer $n = 2,\ m = 0$ alınırsa $\mu_{02} = E(Y^2)$

7) Eğer $n = 1,\ m = 1$ alınırsa $\mu_{11} = E(XY)$

elde edilir. X ve Y rastgele değişkenlerinin varyansları momentler cinsinden sırasıyla şöyledir:

$$\sigma_X{}^2 = Var(X) = \mu_{20} - (\mu_{10})^2$$

$$\sigma_Y{}^2 = Var(Y) = \mu_{02} - (\mu_{01})^2$$

Tanım. X ve Y iki rastgele değişken olmak üzere bu rastgele değişkenlerin birlikte değişim ölçüsüne kovaryans denir ve $Cov(X,Y)$ veya σ_{XY} ile gösterilir.

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

Eğer X ve Y rastgele değişkenleri bağımsız ise $Cov(X,Y) = 0$ olur. $Cov(X,Y) = 0$ ise X ve Y rastgele değişkenleri ilişkisizdir.

Tanım. X ve Y iki rastgele değişken olmak üzere

$$\rho(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

eşitliğine korelasyon katsayıları denir. Momentler cinsinden

$$\rho_{XY} = \frac{\mu_{11} - \mu_{10}\mu_{01}}{\sqrt{(\mu_{20} - \mu_{10}{}^2)(\mu_{02} - \mu_{01}{}^2)}}$$

olarak da yazılabilir ve tanım aralığı,

$$-1 \leq \rho_{XY} \leq 1$$

dır. Korelasyon katsayısı X ve Y arasındaki doğrusal bağımlılığın bir ölçüsü olarak da düşünülebilir.

Örnek: (X,Y) rastgele değişkeninin ortak olasılık fonksiyonu

$Y \backslash X$	0	1	2	$P(Y=y)$
0	0	0	1/4	1/4
1	0	2/4	0	2/4
2	1/4	0	0	1/4
$P(X=x)$	1/4	2/4	1/4	1

olarak veriliyor.

a) $P(X = Y)$ ve $P(|X| \leq 3|[\![Y]\!] \geq 1)$ olasılıklarını bulunuz

b) $E(3X + 5Y)$ ve $Var(2X - 3Y)$ değerlerini bulunuz.

c) Korelasyon katsayısı ρ_{XY}'i bularak yorumlayınız.

Çözüm.

a) $P(X = Y) = p_{X,Y}(0,0) + p_{X,Y}(1,1) + p_{X,Y}(2,2) = 2/4$

$$P(|X| \leq 3|[\![Y]\!] \geq 1) = \frac{P[(|X| \leq 3) \cap ([\![Y]\!] \geq 1)]}{P([\![Y]\!] \geq 1)}$$

$$= \frac{P[(X \leq 2) \cap (Y \geq 1)]}{P(Y \geq 1)}$$

$$P[(X \leq 2) \cap (Y \geq 1)]$$

$$= p_{X,Y}(0,1) + p_{X,Y}(0,2) + p_{X,Y}(1,1) + p_{X,Y}(1,2) + p_{X,Y}(2,1) + p_{X,Y}(2,2) = 3/4$$

olduğundan,

$$P(|X| \leq 3|[\![Y]\!] \geq 1) = \frac{3/4}{3/4} = 1$$

bulunur.

a) $E(3X + 5Y) = 3E(X) + 5E(Y) = 8$

$$E(X) = E(Y) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{2}{4} + 2 \cdot \frac{1}{4} = 1$$

$$Var(2X - 3Y) = 4Var(X) - 12Cov(X, Y) + 9Var(Y)$$

$$E(X^2) = E(Y^2) = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{2}{4} + 2^2 \cdot \frac{1}{4} = \frac{6}{4}$$

$$Var(X) = Var(Y) = \frac{1}{2}$$

$$\sigma_X = \sigma_Y = 1/\sqrt{2}$$

$$E(XY) = \mu_{11} = \sum_{x=0}^{2} \sum_{y=0}^{2} xy\, P_{X,Y}(x, y) = 1.1.\frac{2}{4} = \frac{2}{4}$$

bulunur. Böylece

$$Cov(X, Y) = -0{,}5 \text{ ve } Var(2X - 3Y) = 7$$

bulunur.

c)

$$\rho_{XY} = -\frac{1/2}{(1/\sqrt{2})(1/\sqrt{2})} = -1$$

bulunur. Yani, X ve Y rastgele değişkenleri arasında zıt yönde tam bir ilişki vardır.

Kaynak ders kitabı: Sağlam, V., Sağır, M. ve Yücesoy, E. (2018). Olasılığa Giriş, Güncellenmiş 4. Baskı, Seçkin Yayınevi.