SAB101 OLASILIK

BÖLÜM 2
RASGELE DEĞİŞKENLER
2.5 Ortak Dağılımlı Rasgele Değişkenler
2.6 Rasgele Değişkenin Fonksiyonları

Doç. Dr. Furkan BAŞER
Ankara Üniversitesi
Uygulamalı Bilimler Fakültesi

2.5 Ortak Dağılımlı Rasgele Değişkenler
2.5.1 Ortak Dağılımlı Rasgele Değişkenler (1/4)

- Ortak Olasılık Dağılımı
  - Kesikli

$$P(X = x_i, Y = y_j) = p_{ij} \geq 0$$

$$\text{satisfying} \quad \sum_i \sum_j p_{ij} = 1$$

  - Sürekli

$$f(x, y) \geq 0 \quad \text{satisfying} \quad \iint_{\text{state space}} f(x, y)\,dxdx = 1$$

2.5.1 Ortak Dağılımlı Rasgele Değişkenler (2/4)

- Ortak dağılım fonksiyonu
  - Kesikli

$$F(x,y) = P(X \le x_i, Y \le y_j)$$

  - Sürekli

$$F(x,y) = \sum_{i:x_i \le x} \sum_{j:y_j \le y} p_{ij}$$

$$F(x,y) = \int_{w=-\infty}^{x} \int_{z=-\infty}^{y} f(w,z)\,dz\,dw$$

2.5.1 Ortak Dağılımlı Rasgele Değişkenler (3/4)

- Örnek: Klima Bakımı
  - Klima üniteleri için rezidanslarda ve ofis bloklarında hizmet veren bir firma, teknisyenlerini en verimli şekilde nasıl programlayabilecekleriyle ilgileniyor
  - Servis süresi (X) rasgele değişkeni, 1,2,3 and 4, değerlerini almakta
  - Klimadaki ünite sayısı (Y) rasgele değişkeni de 1,2 and 3 değerlerini almaktadır.

2.5.1 Ortak Dağılımlı Rasgele Değişkenler (4/4)

Y= Ünite sayısı	X=servis zamanı			
	1	2	3	4
1	0.12	0.08	0.07	0.05
2	0.08	0.15	0.21	0.13
3	0.01	0.01	0.02	0.07

- Ortak olasılık fonksiyonu

$$\sum_i \sum_j p_{ij} = 0.12 + 0.18 + \cdots + 0.07 = 1.00$$

- Ortak dağılım fonksiyonu

$$\begin{aligned} F(2,2) &= p_{11} + p_{12} + p_{21} + p_{22} \\ &= 0.12 + 0.18 + 0.08 + 0.15 \\ &= 0.43 \end{aligned}$$

2.5.2 Marjinal Olasılık Dağılımı (1/2)

- Marjinal olasılık dağılımı
  - Diğer rasgele değişkenin değerleri üzerinden ortak olasılık dağılımının toplanması veya integrallenmesi ile elde edilir.
  - Kesikli

$$P(X = i\ ) = p_{i+} = \sum_{j} p_{ij}$$

  - Sürekli

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy$$

2.5.2 Marjinal Olasılık Dağılımı (2/2)

- Örnek
  - X in marjinal olasılık fonksiyonu

$$P(X=1)=\sum_{j=1}^{3} p_{1j}=0.12+0.08+0.01=0.21$$

  - Y nin marjinal olasılık fonksiyonu

$$P(Y=1)=\sum_{i=1}^{4} p_{i1}=0.12+0.08+0.07+0.05=0.32$$

- Örnek: (Sürekli durum)

- Ortak olasılık yoğunluk fonksiyonu: $f(x, y)$

- X ve Y nin marjinal olasılık yoğunluk fonksiyonu::

$$f_X(x) = \int f(x, y)dy$$

$$f_Y(y) = \int f(x, y)dx$$

2.5.3 Koşullu Olasılık Dağılımları (1/2)

- Koşullu Olasılık Dağılımları

  Y değerinin bilindiği koşulu altında X rasgele değişkeninin olasılığı

  - Kesikli

$$p_{i|j} = P(X = i \mid Y = j) = \frac{P(X = i, Y = j)}{P(Y = j)} = \frac{p_{ij}}{p_{+j}}$$

  - Sürekli

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

  - Koşullu olasılık dağılımı da bir olasılık dağılımıdır..

2.5.3 Koşullu Olasılık Dağılımları (2/2)

- Örnek
  - Y nin marjinal olasılık dağılımı

$$P(Y=3)=p_{+3}=0.01+0.01+0.02+0.07=0.11$$

  - X in koşullu dağılımı

$$p_{1|Y=3}=P(X=1\,|\,Y=3)=\frac{p_{13}}{p_{+3}}=\frac{0.01}{0.11}=0.091$$

2.5.4 Bağımsızlık ve Kovaryans (1/5)

- X ve Y rasgele değişkenleri bağımsızdır denilebilir eğer ;
  - Kesikli

$$p_{ij} = p_{i+} p_{+j} \quad \text{for all values } i \text{ of } X \text{ and } j \text{ of } Y$$

  - Sürekli

$$f(x, y) = f_X(x) f_Y(y) \quad \textit{for all } x \textit{ and } y$$

  - Bağımsızlığın, olaylar arasındaki bağımsızlıktan farkı nedir?

2.5.4 Bağımsızlık ve Kovaryans (2/5)

- Kovaryans

$$\begin{aligned}\mathrm{Cov}(X,Y) &= E((X-E(X))(Y-E(Y))) \\ &= E(XY)-E(X)E(Y)\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}(X,Y) &= E((X-E(X))(Y-E(Y))) \\ &= E(XY-XE(Y)-E(X)Y+E(X)E(Y)) \\ &= E(XY)-E(X)E(Y)-E(X)E(Y)+E(X)E(Y) \\ &= E(XY)-E(X)E(Y)\end{aligned}$$

- Herhangi bir pozitif veya negatif sayı alabilir.
- Bağımsız olayların kovaryansı 0 dır.
- Peki ya kovaryans 0 ise?

2.5.4 Bağımsızlık ve Kovaryans (3/5)

- Örnek: (Klima bakımı)

$$E(X) = 2.59, \qquad E(Y) = 1.79$$

$$E(XY) = \sum_{i=1}^{4}\sum_{j=1}^{3} ijp_{ij}$$

$$= (1\times1\times0.12) + (1\times2\times0.08)$$

$$+ \cdots + (4\times3\times0.07) = 4.86$$

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$$

$$= 4.86 - (2.59\times1.79) = 0.224$$

2.5.4 Bağımsızlık ve Kovaryans (4/5)

- Korelasyon:

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

  - −1 ve 1 arasındaki değerler alır.
  - Bağımsız rasgele değişkenler sıfır korelasyona sahiptir.

2.5.4 Bağımsızlık ve Kovaryans (5/5)

- Örnek: (Klima bakımı)

$$\mathrm{Var}(X) = 1.162, \quad \mathrm{Var}(Y) = 0.384$$

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

$$= \frac{0.224}{\sqrt{1.162 \times 0.384}} = 0.34$$

- X ve Y rasgele değişkenleri doğrusal bir ilişkiye sahipse ne olur?

$$Y = aX + b$$

$$a \neq 0$$

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

$$= E[X(aX + b)] - E[X]E[aX + b]$$

$$= aE[X^2] + bE[X] - aE^2[X] - bE[X]$$

$$= a(E[X^2] - E^2[X]) = aVar(X)$$

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{aVar(X)}{\sqrt{Var(X)a^2Var(X)}}$$

Yani, Corr(X,Y) = 1 eğer a>0; −1 eğer a<0.

2.6 Rasgele Değişkenlerin Fonksiyonları

2.6.1 Rasgele Değişkenlerin Doğrusal Fonksiyonları (1/4)

- Rasgele Değişkenlerin Doğrusal Fonksiyonları
  - X rasgele değişkeni ise ve $Y = aX + b$ $a, b \in \mathrm{R}$ için;

$$E(Y) = aE(X) + b$$

$$\mathrm{Var}(Y) = a^2 \mathrm{Var}(X)$$

- **Standartlaştırma**
  - X rasgele değişkeninin beklenen değeri μ varyansı σ^2 ise

$$Y = \frac{X-\mu}{\sigma} = \frac{1}{\sigma} X + \left(-\frac{\mu}{\sigma} \right)$$

Y r.d nin beklenen değeri 0, varyansı 1 olur.

2.6.1 Rasgele Değişkenlerin Doğrusal Fonksiyonları (2/4)

- Örnek: Test Sonuçlarının Standartlaştırılması
  - Belirli bir test prosedüründen elde edilen ham puanın (X) , beklenen değeri 10 ve varyans 7 ile –5 ve 20 arasında dağıtıldığını varsayalım. Puanları 0 ile 100 arasında bir değerde olacak şekilde standartlaştırmak için $Y = 4X + 20$ doğrusal dönüşümü uygulanır.

2.6.1 Rasgele Değişkenlerin Doğrusal Fonksiyonları (3/4)

- Örneğin, x=12 için standartlaştırılmış değer y=(4 x 12)+20=68

$$E(Y) = 4E(X) + 20 = (4 \times 10) + 20 = 60$$

$$\mathrm{Var}(Y) = 4^2\,\mathrm{Var}(X) = 4^2 \times 7 = 112$$

$$\sigma_Y = \sqrt{112} = 10.58$$

2.6.1 Rasgele Değişkenlerin Doğrusal Fonksiyonları (4/4)

- Rasgele Değişkenlerin Toplamı
  - X_1 ve X_2 rasgele değişkenleri ise

$$E(X_1 + X_2) = E(X_1) + E(X_2) \cdots\cdots\cdots\cdots (why?)$$

ve

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2\mathrm{Cov}(X_1, X_2)$$

  - Eğer X_1 ve X_2 r.d leri bağımsızsa $\mathrm{Cov}(X_1, X_2) = 0$
    Bu durumda;

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2)$$

- $Cov(X_1, X_2)$ ın Özellikleri:

$$Cov(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2]$$

$$\Rightarrow Cov(X_1, X_2) = Cov(X_2, X_1)$$

$$Cov(X_1, X_1) = Var(X_1) \quad Cov(X_2, X_2) = Var(X_2)$$

$$Cov(X_1 + X_2, X_1) = Cov(X_1, X_1) + Cov(X_2, X_1)$$

$$Cov(X_1 + X_2, X_1 + X_2) = Cov(X_1, X_1) + Cov(X_1, X_2) +$$

$$Cov(X_2, X_1) + Cov(X_2, X_2)$$

$$\Rightarrow Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2Cov(X_1, X_2)$$

2.6.2 Rasgele Değişkenlerin Doğrusal Bileşimleri(1/5)

- Rasgele Değişkenlerin Doğrusal Bileşimleri
  - Eğer $X_1, \ldots, X_n$ r.d leri ve $a_1, \ldots, a_n$ ve b sabitlerse

$$E(a_1 X_1 + \cdots + a_n X_n + b) = a_1 E(X_1) + \cdots + a_n E(X_n) + b$$

  - Rasgele değişkenler bağımsızsa:

$$\mathrm{Var}(a_1 X_1 + \cdots + a_n X_n + b) = a_1^{2} \mathrm{Var}(X_1) + \cdots + a_n^{2} \mathrm{Var}(X_n)$$

2.6.2 Rasgele Değişkenlerin Doğrusal Bileşimleri(2/5)

- Bağımsız rasgele değişkenlerin ortalaması
  - $X_1, \ldots, X_n$ ortalaması μ ve varyası σ^2 olan bağımsız rasgele değişkenlerin bir dizisi olsun.
  - $$\overline{X} = \frac{X_1 + \cdots + X_n}{n}$$
  - $$E(\overline{X}) = \mu$$

$$\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$$

2.6.2 Rasgele Değişkenlerin Doğrusal Bileşimleri(3/5)

$$E(\overline{X}) = E\left(\frac{1}{n}X_1 + \cdots \frac{1}{n}X_n\right) = \frac{1}{n}E(X_1) + \cdots + \frac{1}{n}E(X_n)$$

$$= \frac{1}{n}\mu + \cdots \frac{1}{n}\mu = \mu$$

$$\mathrm{Var}(\overline{X}) = \mathrm{Var}\left(\frac{1}{n}X_1 + \cdots + \frac{1}{n}X_n\right) = \left(\frac{1}{n}\right)^2 \mathrm{Var}(X_1) + \cdots \left(\frac{1}{n}\right)^2 \mathrm{Var}(X_n)$$

$$= \left(\frac{1}{n}\right)^2 \sigma^2 + \cdots + \left(\frac{1}{n}\right)^2 \sigma^2 = \frac{\sigma^2}{n}$$

2.6.2 Rasgele Değişkenlerin Doğrusal Bileşimleri(4/5)

- Örnek
  - İki testin standart puanı;

$$Y_1 = \frac{10}{3} X_1 \quad \text{and} \quad Y_2 = \frac{5}{3} X_2 + \frac{50}{3}$$

  - Final puanı;

$$Z = \frac{2}{3} Y_1 + \frac{1}{3} Y_2 = \frac{20}{9} X_1 + \frac{5}{9} X_2 + \frac{50}{9}$$

2.6.2 Rasgele Değişkenlerin Doğrusal Bileşimleri(5/5)

– Final puanının beklenen değeri

$$\begin{aligned} E(Z) &= \frac{20}{9}E(X_1)+\frac{5}{9}E(X_2)+\frac{50}{9} \\ &= \left(\frac{20}{9}\times 18\right)+\left(\frac{5}{9}\times 30\right)+\frac{50}{9} \\ &= 62.22 \end{aligned}$$

– Final puanının varyansı

$$\mathrm{Var}(Z)=\mathrm{Var}\left(\frac{20}{9}X_1+\frac{5}{9}X_2+\frac{50}{9}\right)$$

$$=\left(\frac{20}{9}\right)^2\times\mathrm{Var}(X_1)+\left(\frac{5}{9}\right)^2\times\mathrm{Var}(X_2)$$

$$=\left(\frac{20}{9}\right)^2\times 24+\left(\frac{5}{9}\right)^2\times 60=137.04$$

2.6.3 Rasgele Değişkenlerin Doğrusal Olmayan Fonksiyonları(1/3)

- Rasgele Değişkenlerin Doğrusal Olmayan Fonksiyonları
  - g, X rasgele değişkeninin doğrusal olmayan bir fonksiyonu, için bir başka rasgele değişken Y = g (X) ‘dir

$$Y = X^2, \; Y = \sqrt{X}, \; Y = e^X$$

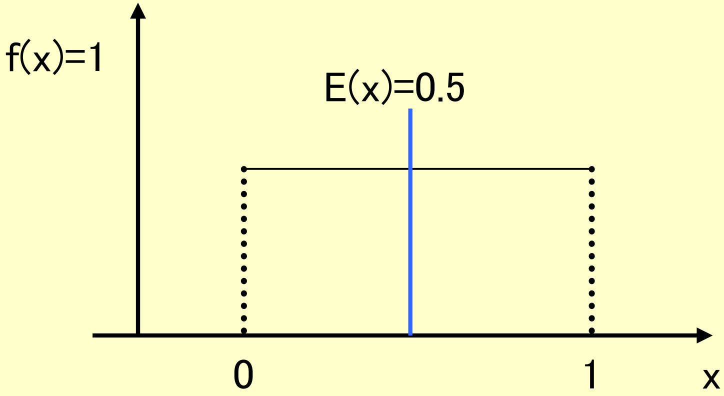
  - Y rasgele değişkeninin beklenen değeri ve varyansını, X rasgele değişkeninin beklenen değeri ve varyansı ile ilişkilendiren genel bir sonuç yoktur.

2.6.3 Rasgele Değişkenlerin Doğrusal Olmayan Fonksiyonları(2/3)

- Örneğin,

$$f_X(x) = 1 \quad \text{for} \quad 0 \le x \le 1$$
$$f(x) = 0 \quad \text{elsewhere}$$

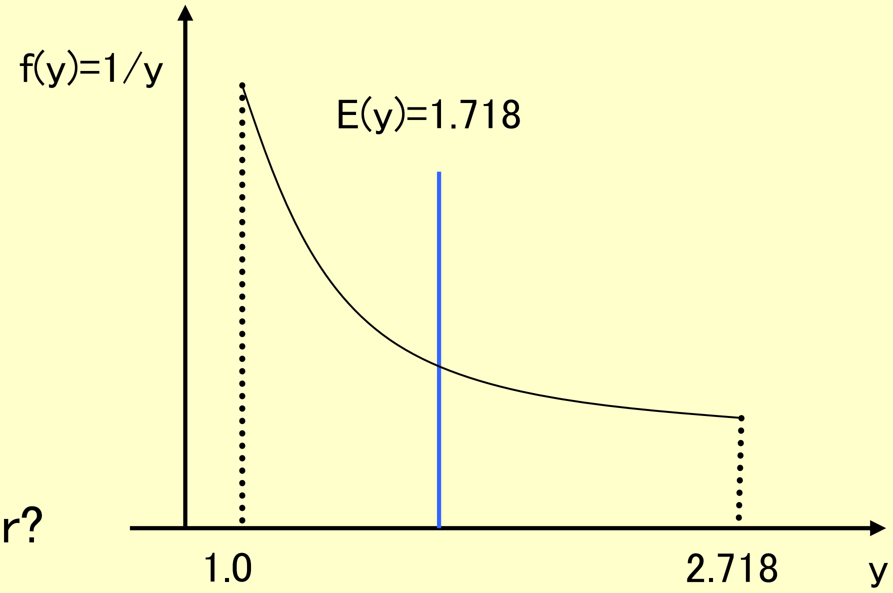
$$F_X(x) = x \quad \text{for} \quad 0 \le x \le 1$$



- Eğer

$$Y = e^X$$

$$\text{where} \quad 1 \le Y \le 2.718$$



İse Y nin dağılım fonksiyonu nedir?

2.6.3 Rasgele Değişkenlerin Doğrusal Olmayan Fonksiyonları(3/3)

- **Dağılım fonksiyonu metodu**

$$F_Y(y) = P(Y \le y) = P(e^X \le y) = P(X \le \ln(y)) = F_x(\ln(y)) = \ln(y)$$

  - Y nin dağılım fonksiyonunun türevlenebilir olması gerekir:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{y} \qquad \text{for } 1 \le y \le 2.718$$

- Dikkat

$$E(Y) = \int_{z=1}^{2.718} z f_Y(z)dz = \int_{z=1}^{2.718} 1dz = 2.718 - 1 = 1.718$$

$$E(Y) \ne e^{E(X)} = e^{0.5} = 1.649$$

- Rasgele değişkenin doğrusal olmayan dönüşümünün olasılık (yoğunluk) fonksiyonu

  $f_X(x)$ ve $Y = g(X)$ verildiğinde $f_Y(y)$ ne olur?

Eğer $x_1, x_2, \cdots, x_n$ hepsi gerçek kökse yani;

$$y = g(x_1) + g(x_2) + \cdots + g(x_n)$$

$$f_Y(y) = \sum_{i=1}^{n} \frac{f_X(x_i)}{|g'(x_i)|}$$

$$g'(x) = \frac{dg(x)}{dx}$$

- Örnek: $f_Y(y)$ belirleme

$$f_X(x) = 1 \quad \text{for} \quad 0 \le x \le 1 \qquad\qquad Y = e^X$$

$$f(x) = 0 \quad \text{elsewhere}$$

$$\text{where} \quad 1 \le Y \le 2.718$$

$$y = g(x) = e^x$$

-> Bir kök mümkün

$$\Rightarrow \ln y = x \qquad\qquad 0 \le x \le 1$$

$$\frac{dg}{dx} = e^x = y$$

$$f_Y(y) = \frac{1}{|y|} = \frac{1}{y}$$