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

RASTGELE DEĞİŞKENLER

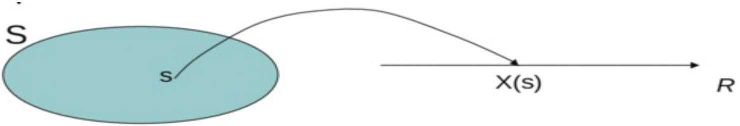
Fatih Kızılaslan

Marmara Üniversitesi

2019-2020 Bahar 9. Hafta

Rastgele Değişken

Rastgele (rassal veya rastlantı) değişken S örnek uzayındaki her rastgele olaya sayısal değerler atayan bir fonksiyondur.
Bu fonksiyon örnek uzayın her elamanını reel sayılar kümesi $\mathbb{R} = (-\infty, \infty)$'ye taşıyan bir fonksiyondur.



- Rastgele değişkenleri X, Y, Z gibi büyük harfler ile gösteririz.
- Rastgele değişkenin aldığı değerleri ise x, y, z gibi küçük harfler ile gösteririz.
- Böylece, $s \in S$ için bir X rastgele değişkeninin alacağı değeri $X(s) = x \in \mathbb{R}$ biçiminde ifade ederiz.

Örnek 1: Bir madeni para 3 kez atılsın. X rastgele değişkenini bu 3 atışta gelen turaların sayısı olarak tanımlansın. X'in aldığı değerleri ve olasılıklarını bulunuz.

Çözüm: Bu deney için örnek uzayımız

$$S = \{YYY, YYT, YTY, TYY, TTY, TYT, YTT, TTT\}$$

biçiminde 8 elemandan oluşur. X rastgele değişkeni bu 3 atıştaki tura sayısı olduğundan alabilecek olduğu değerler **Değer Kümesi:**$\{0, 1, 2, 3\}$ yani $x \in \{0, 1, 2, 3\}$ olur.

S	X r.d. (Tura sayısı)	Olasılık
YYY	0	$1/8$
YYT, YTY, TYY	1	$3/8$
TTY, TYT, YTT	2	$3/8$
TTT	3	$1/8$

Böylece, $P(X = 0) = 1/8$, $P(X = 1) = 3/8$, $P(X = 2) = 3/8$ ve $P(X = 3) = 1/8$ dir. Ayrıca, her $x \in \{0, 1, 2, 3\}$ için $P(X = x) > 0$ ve bu olasılıkların toplamları yani $\sum_{x=0}^{3} P(X = x) = 1$ dır.

Kesikli Rastgele Değişken

X rastgele değişkeninin tüm olası değerlerinin (değer kümesi) sayısı sonlu veya sayılabilir ise X'e **kesikli rastgele değişken** denir.

Örneğin,

- Bir paranın n defa atılmasında gelen yazıların sayısı,
- Bir hastaneye bir saat içinde gelen hastaların sayısı,
- Bir otoparka bir günde gelen araçların sayısı,
- Kasa sırasında bekleyen müşteri sayısı gibi.

Bu örneklerdeki gibi oluşturulan her bir X rastgele değişkeninin alabilecek olduğu değerler sonludur.

Sürekli Rastgele Değişken

X rastgele değişkeni bir aralıktaki ya da aralıklar kümesindeki tüm değerleri alabiliyor ise X'e **sürekli rastgele değişken** denir. (**Not:** Herhangi bir aralıkta sonsuz tane reel sayı vardır.)

Örneğin,

- Bir telefon görüşmesinin süresi,
- Sınıftaki öğrencilerin ağırlıkları, uzunlukları,
- Hastaların tansiyon değeri,
- Hasta muayene süresi gibi.

Bu örneklerdeki gibi oluşturulan her bir X rastgele değişkeninin alabilecek olduğu değerler reel sayılar kümesinin bir alt kümesindeki değerlerdir. Mesela, hasta muayene süresi 1dk ile 10 dk arasında değişmektedir. Bu durumda muayene süresi $[1, 10]$ aralığındaki herhangi bir reel sayıdır.

Olasılık Fonksiyonu

X kesikli rastgele değişken olmak üzere X'in her olası x değeri için

$$p(x) = P(X = x)$$

fonksiyonuna X'in **olasılık fonksiyonu** denir.

Bir $p(x)$ fonksiyonunun olasılık fonksiyonu olabilmesi için

1. Tüm olası değerleri $\{x_1, ..., x_N\}$ olmak üzere $p(x_i) = P(X = x_i) \geq 0$, $i = 1, ..., N$
2. $\sum_{i=1}^{N} p(x_i) = \sum_{i=1}^{N} P(X = x_i) = 1$,

olmalıdır.

Örnek 1'de tanımlanan X kesikli bir rastgele değişkendir ve olasılık fonksiyonu

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

biçiminde olur. $\sum_{x=0}^{3} p(x) = \sum_{x=0}^{3} P(X = x) = 1$ dir.

Örnek 2: Bir X rastgele değişkeninin olasılık fonksiyonu

x	5	10	15	20
$p(x)$	0.1	c	0.2	0.4

biçiminde verilsin.

Bu durumda $c = ?$, $P(X < 11) = ?$ ve $P(X \geq 14) = ?$

Çözüm: Verilen $p(x)$ bir olasılık fonksiyonu olduğundan tüm değerlerinin toplamı 1 olmalıdır. Böylece,

$$\sum_{x} p(x) = 1 \Leftrightarrow 0.1 + c + 0.2 + 0.4 = 1 \Rightarrow c = 0.3$$

$P(X < 11) = P(X = 5) + P(X = 10) = 0.1 + 0.3 = 0.4,$
$P(X \geq 14) = P(X = 15) + P(X = 20) = 0.6$ veya $P(X \geq 14) =$
$1 - P(X < 14) = 1 - \{P(X = 5) + P(X = 10)\} = 1 - 0.4 = 0.6$

Olasılık Yoğunluk Fonksiyonu

X sürekli rastgele değişken olmak üzere

- her x için $f(x) \geq 0$
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- $P(a \leq X \leq b) = \int_a^b f(x)dx$, $-\infty < a < b < \infty$ (Herhangi bir $[a, b]$ aralığındaki olasılıktır.)

koşullarını sağlayan $f(x)$ fonksiyonuna X sürekli rastgele değişkeninin **olasılık yoğunluk fonksiyonu** denir.

- Sürekli rastgele değişkenin bir noktadaki olasılığı 0 dır. Yani, $P(X = c) = 0$.
- Sürekli rastgele değişken için **integral**, kesikli rastgele değişken için **toplam sembolü** kullanılır.

Örnek 3: $f(x) = \begin{cases} ax & , 0 \leq x < 1 \\ a & , 1 \leq x \leq 2 \\ 0 & , \text{diğer durumlar} \end{cases}$ biçiminde verilen $f(x)$ fonksiyonunun X sürekli rastgele değişkeninin olasılık yoğunluk fonksiyonu olması için a ne olmalıdır ? Ayrıca, $P(1/2 < X < 3/4)$ olasılığını hesaplayınız.

Çözüm: $f(x)$'in olasılık yoğunluk fonksiyonu olması için $\int_{-\infty}^{\infty} f(x)dx = 1$ olmalıdır. Bu nedenle,

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{0} 0dx + \int_{0}^{1} ax\ dx + \int_{1}^{2} a\ dx + \int_{2}^{\infty} 0\ dx \\ &= 0 + a\ \left.\frac{x^2}{2}\right|_{x=0}^{1} + a\ \ x|_{x=0}^{1} + 0 \\ &= a\frac{1}{2} + a = \frac{3a}{2} \end{aligned}$$

ve $\frac{3a}{2} = 1 \Rightarrow a = \frac{2}{3}$ bulunur.

Böylece, $f(x) = \begin{cases} \frac{2}{3}x & , 0 \leq x < 1 \\ \frac{2}{3} & , 1 \leq x \leq 2 \\ 0 & , \text{diğer durumlar} \end{cases}$ olur.

$$P(1/2 < X < 3/4) = \int_{1/2}^{3/4} f(x)dx = \int_{1/2}^{3/4} \frac{2}{3}x \, dx$$

$$= \frac{2}{3} \frac{x^2}{2}\Big|_{x=1/2}^{3/4} = \frac{2}{3}\frac{1}{2}\left(\frac{9}{16} - \frac{1}{4}\right) = \frac{5}{48}.$$

Beklenen Değer

Beklenen değer bir rastgele değişkenin ortalama değeridir. Bir X rastgele değişkeni için beklenen değeri $E(X)$ olarak gösteririz.

Örneğin, bir hilesiz paranın 1000 kez atıldığını varsayalım. X rastgele değişkeni gelen yazıların sayısı olarak tanımlansın. Bu durumda bu 1000 atıştaki yazı sayısana ilişkin **matematiksel beklentimiz** $1000\ \frac{1}{2} = 500$ olacaktır.

X kesikli rastgele değişkeni için **beklenen değer** aldığı değerler ile onlara karşılık gelen olasılıkların çarpımlarının toplamıdır.
Böylece, X kesikli rastgele değişkeni için beklenen değer

$$E(X) = \sum_{x} x\ p(x)$$

biçiminde tanımlanır.

X sürekli rastgele değişkeni için beklenen değer $E(X) = \int_{-\infty}^{\infty} xf(x)dx$ biçiminde tanımlanır.

Örnek 4: Örnek 1'deki bir paranın 3 kez atılması deneyinde X rastgele değişkeni gelen turaların sayısı olmak üzere olasılık fonksiyonu

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

biçimindedir. Bu X rastgele değişkeni için beklenen değer $E(X)$ bulalım.

Çözüm: Kesikli rastgele değişken olduğundan

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{3} x\, p(x) = 0(1/8) + 1(3/8) + 2(3/8) + 3(1/8) \\
&= 3/8 + 6/8 + 3/8 = 12/8 = 3/2
\end{aligned}
$$

dir. Bunun anlamı: 3 kez para atılması deneyi çok defa tekrar edildiğinde her 3 atıştan 3/2 si yazı olacaktır.

Gelecek hafta: Varyans ve Kesikli Olasılık Dağılımları