Olasılık Dağılımlarından Sayı Üretme

1. Ters Dönüşüm Yöntemi

F dağılım fonksiyonuna sahip bir X rasgele değişkenin dağılımından sayı üretmek için en çok kullanılan yöntemlerden biri, F dağılım fonksiyonunun genelleştirilmiş tersi denen

$$F^{-} : (0,1) \to R$$

$$u \to F^{-}(u) = \inf\{x: F(x) \geq u\}$$

fonksiyonuna dayalı $X = F^{-}(U)$ dönüşümünü kullanmaktır. Burada U rasgele değişkeni (0,1) aralığı üzerindeki düzgün dağılıma, yani $U(0,1)$ dağılımına sahiptir. $X = F^{-}(U)$ rasgele değişkenin dağılım fonksiyonu,

$$P(X \leq x) = P[F^{-}(U) \leq x] = P\left[F\left(F^{-}(U)\right) \leq F(x)\right]$$

$$= P[U \leq F(x)] = F(x)$$

dır.

$X = F^{-}(U)$ dönüşümü integral dönüşümü olarak bilinmektedir. $U(0,1)$ düzgün dağılımdan üretilen sayılar integral dönüşümü sonucunda X rasgele değişkenin dağılımından üretilmiş sayılar olacaktır.

Algoritma

1. $U(0,1)$ dağılımından U üretilir
2. $X = F^{-}(U)$ hesaplanır

Örnek1: X rasgele değişkeninin olasılık yoğunluk fonksiyonu ve dağılım fonksiyonu,

$$f(x) = \begin{cases} 2x, & 0 \leq \mathrm{x} \leq 1 \\ 0, & \text{d.y.} \end{cases}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

Olup, $X = F^{-}(U) = U^{1/2} = \sqrt{U}$ dönüşümü ile X rasgele değişkenin dağılımından sayı üretilebilir.

MATLAB KODU

```
u=rand(50,1);
x=sqrt(u);
hist(x);
```

Ödev1: En az 50 sayı üretiniz ve üretilen sayıların bu dağılımdan gelip gelmediğini irdeleyiniz.

Örnek2: θ parametreli Üstel dağılımın olasılık yoğunluk fonksiyonu ve dağılım fonksiyonu aşağıda verilmiştir.

$$f(x) = \begin{cases} \dfrac{1}{\theta} e^{\frac{-x}{\theta}}, & x \geq \theta \\ 0, & \text{d.y.} \end{cases} \qquad F(x) = \begin{cases} 0 & , x < 0 \\ 1 - e^{-\frac{x}{\theta}}, & x \geq 0 \end{cases}$$

$X = F^{-}(U) = -\theta \ln(1-U)$ dönüşümü ile üretilen X rasgele sayıları üstel dağılımdan üretilmiş sayılardır.

MATLAB KODU

```
u=rand(100,1);
x= -5*log(1-rand(100,1));
x=-5*log(u);
hist(x)
```

expinv komutu ile sayı üretme

```
u=rand(100,1);
x=expinv(u,5); % üstel dağılıma dönüştürme
hist(x)
```

Ödev2: θ parametresine bir değer verip en az 50 tane sayı üretiniz ve üretilen sayıların bu dağılımdan gelip gelmediğini irdeleyiniz.

Örnek3: X rasgele değişkeninin olasılık fonksiyonu

X=x	1	2	3
f(x)	0.2	0.5	0.3

olsun. Kesikli bir rasgele değişken olan X ’in dağılım fonksiyonu

$$F(x)=\begin{cases}0 & , \quad x<1 \\ 0.2 & , \quad 1\le x\le 2 \\ 0.7 & , \quad 2\le x<3 \\ 1 & , \quad x\ge 3\end{cases}$$

dır. F^{-} fonksiyonu,

$$F^{-}(u)=\begin{cases}1, & 0<u\le 0.2 \\ 2, & 0.2<u\le 0.7 \\ 3, & 0.7<u<1\end{cases}$$

olmak üzere,

$$X=F^{-}(U)=\begin{cases}1, & 0<U\le 0.2 \\ 2, & 0.2<U\le 0.7 \\ 3, & 0.7<U<1\end{cases}$$

dönüşümü ile X rasgele değişkenin dağılımından sayı üretilebilir.

MATLAB KODU

```
u=rand
if (u>0 && u<=0.2)
x=1
else if (u>0.2 && u<=0.7)
x=2
else
x=3
end
end
```

2. Kabul-Red Yöntemi

Reddetme tekniği, sürekli ve sınırlı olan herhangi bir $f(x)$ olasılık yoğunluk fonksiyonundan rasgele değişken üretmek için kullanılan genel bir metottur. Diğer teknikler başarısız veya etkin olmadığında kullanılır.

X rasgele değişkenin olasılık yoğunluk fonksiyonu f ve aldığı değerler kümesi D olsun. V rasgele değişkenin olasılık yoğunluk g ve aldığı değerlerin kümesi D olmak üzere V sayıları kolayca üretilebilsin. $a > 0$ sabiti ve $\forall x \in D$ için

$$f(x) \leq a \cdot g(x)$$

koşulu sağlansın. Bu durumda aşağıdaki algoritma ile olasılık yoğunluk fonksiyonu f olan dağılımından sayı üretilebilir.

Algoritma

1) V sayısı olasılık yoğunluk fonksiyonu g olan dağılımından üretilsin

2) V 'den bağımsız olarak $U \sim U(0,1)$ üretilsin.

$U \cdot a \cdot g(V) \leq f(V)$ ise $X = V$ kabul edilsin yani bir X sayısı üretilmiş olsun, aksi durumda reddedilsin yani 1. adıma geçilsin (başka bir ifade ile $Y \sim U(0, a \cdot g(V))$ üretilsin. $Y < f(V)$ ise X kabul edilsin, aksi durumda reddedilsin.)

Olasılık yoğunluk fonksiyonu f olan dağılımdan X sayılarını üretmek için yukarıdaki algoritmayı aşağıdaki gibi de yazabiliriz.

1) Birbirinden bağımsız olarak $V \sim g$ ve $U \sim U(0,1)$ üretilsin.

2) Eğer $U \leq \dfrac{f(V)}{ag(V)}$ ise $X = V$ olsun ve X sayısı kabul edilsin.

Örnek: X rasgele değişkenin olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} \frac{2}{\pi}\sqrt{x(2-x)} & , \ 0 < x < 2 \\ 0 & , \ \text{d.y.} \end{cases}$$

olsun, uygun bir $g(x)$ fonksiyonu da

$$g(x) = \begin{cases} \frac{1}{2} & , 0 < x < 2 \\ 0 & , \text{d.y.} \end{cases}$$

şeklinde seçilsin. Bu durumda $a = 2$ için $f(x) \le 2g(x)$ ($x \in (0,2)$) dır. Buna göre algoritma aşağıdaki gibi olacaktır.

1) $U \sim U(0,1)$ dağılımından bir U_1 sayısı üretilip $V = 2U_1$ alınır.

U(0,1) dağılımından bir U_2 sayısı üretilir.

2) Eğer $U_2 \le \dfrac{\frac{2}{\pi}\sqrt{V(2-V)}}{2.\frac{1}{2}}$ yani $U_2 \le \frac{2}{\pi}\sqrt{V(2-V)}$ ise

$X = V$ alınır ve X sayısı kabul edilir.

Bilgisayar programında V sayısını X ile göstermek (X adresine yazdırmak) ikinci adımda kolaylık sağlamaktadır. Karışıklığa yol açmadığı takdirde algoritmanın birinci adımında g olasılık yoğunluk fonksiyonuna sahip dağılımdan üretilen V sayısı yerine X yazılabilir. Buna göre algoritma aşağıdaki şekli alır.

1) Birbirinden bağımsız olarak $X \sim g$ ve $U \sim U(0,1)$ üretilir.

2) Eğer $U \le \dfrac{f(X)}{ag(X)}$ ise X kabul edilir aksi halde red edilir.

MATLAB KODU

```
u1=rand
x=2*u1
u2=rand
k=0;
S=(2/3.14)*sqrt(x*(2-x))
if u2 < S
   k=k+1;
end
k
```

Kolayca sayı üretilebilen yardımcı dağılımın olasılık yoğunluk fonksiyonu olan g fonksiyonu f fonksiyonuna ne kadar çok benziyorsa ve grafikleri birbirine yakınsa simülasyon zamanı kısadır, red olunmalar o kadar az olur. g fonksiyonu seçildikten sonra a sabiti $\forall x \in D$ için $a \geq \frac{f(x)}{g(x)}$ olacak şekilde ve $\frac{f(x)}{ag(x)}$ değerleri bire yakın olacak şekilde seçilmelidir. Bu şartlar altında g seçildikten sonra a sabiti

$$a = \sup_{x \in D} \frac{f(x)}{g(x)}$$

olarak seçilebilir (mevcut olduğu taktirde $a = \max_{x \in D} \frac{f(x)}{g(x)}$ olarak seçilebilir).

Yukarıdaki örnekte sayı üretilmek istenen dağılımın olasılık yoğunluk fonksiyonu,

$$f(x) = \begin{cases} \frac{2}{\pi}\sqrt{x(2-x)} & , \ 0 < x < 2 \\ 0 & , \ \text{d.y.} \end{cases}$$

ve yardımcı dağılımın olasılık yoğunluk fonksiyonu

$$g(x) = \begin{cases} \frac{1}{2} & , \ 0 < x < 2 \\ 0 & , \ \text{d.y.} \end{cases}$$

dır.

$$\frac{f(x)}{g(x)} = 2\frac{2}{\pi}\sqrt{x(2-x)}$$

olmak üzere, maksimum değerini $x = 1$’de almaktadır. $a = \frac{f(1)}{g(1)} = \frac{4}{\pi}$ olarak seçilirse algoritmadaki $Uag(X) \leq f(X)$ eşitsizliği $U \leq \sqrt{X(2-X)}$ biçiminde olur. Buna göre algoritma:

1) Birbirinden bağımsız olarak $X \sim U(0,2)$ ve $U \sim U(0,1)$ üretilir.
2) Eğer $U \leq \sqrt{X(2-X)}$ ise X kabul edilir aksi halde red edilir.

MATLAB kodu:

```
x=2*rand
S= sqrt(x*(2-x))
k=0; r=rand
if r < S
k=k+1;
```

```
end
k
```

Ödev1: Bu algoritma (program) ile bir öncekini karşılaştırmak amacıyla her ikisi ile üretilen 100'er tane X sayısı için döngü sayılarını gözleyiniz.

Ödev2: $B(\alpha,\beta)$ dağılımına sahip X rasgele değişkenin olasılık yoğunluk fonksiyonu

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, 0 \le x \le 1$$

ve g, $U(0,1)$ düzgün dağılımın olasılık yoğunluk fonksiyonu olmak üzere a sabiti 135/64 olsun. B($\alpha = 2, \beta = 4$) olan Beta dağılımından sayı üreten algoritmayı ve MATLAB kodunu yazınız.