Gazi Üniv. End. Sant. Eğt. Fak.Der.
Sayı : 3, Haziran 1995, Sayfa : 61 - 70

KESİKLİ DAĞILIM FONKSİYONLARI İÇİN BİLGİSAYAR ALGORİTMASI VE UYGULAMASI

Öğr. Gör. Nursal ARICI*

1- GİRİŞ

Bu çalışmada, klasik istatistiksel dağılım fonksiyonlarından kesikli dağılımlar için olasılık (yoğunluk) fonksiyonu ve dağılım fonksiyonu hesaplayan tek bir algoritma ve bu algoritmanın binom ve poisson dağılımları için Fortran dilinde kodlanmış programları tanıtılacaktır. Buna ek olarak kodlanmış programların binom ve poisson dağılımlarına uyan bir uygulama problemi ile nasıl kullanabileceği ve kullanıcıya neler sağlayacağı açıklanacaktır.

2. KESİKLİ DAĞILIMLAR İÇİN ALGORİTMA

Tablo-2'de verdiğimiz algoritma Tablo-1'de(1) olasılık yoğunluk fonksiyonu, parametreleri ve değişim aralıkları verilen kesikli dağılımların fonksiyonlarının tümü için kullanılabilecek tek bir algoritmadır.

Algoritmanın genel mantığı, kesikli değişkenlerin olasılık (yoğunluk) fonksiyonu ve dağılım fonksiyonu tanımlarında yatmaktadır.

Bilindiği üzere;

"Bir X (S) kümesinin ∀xi elemanını P(X=xi) şeklinde tanımlanan bir olasılık karşı gelecek şekilde bir olasılık uzayına eşlediğimizde $P(X=x_i) = f(X_i)$ şeklinde tanımlanan fonksiyona X'in olasılık (yoğunluk) fonksiyonu (2)

* Gazi Üniversitesi Endüstriyel Sanatlar Eğitimi Fakültesi, Bilgisayar Eğitimi Bölümü

TABLO 1 : BAZI KESİKLİ DAĞILIMLARIN ÖZELLİKLERİ			
DAĞILIM	f(x):olasılık fonksiyonu	Parametre(leri)	Değişim Aralığı
Binomiyal	$\binom{n}{x} p^{x} q^{n-x}$	$n \geq 1$ $0 \leq p \leq 1$	x=0,1,...,n
Poisson	$\frac{e^{-\alpha} \alpha^{x}}{x!}$	$\alpha > 0$	x=0,1,...
Negatif Binomiyal	$\binom{i-1}{r-1} p^{r} q^{i-r}$	$0 < p \leq 1$ r=1,2,...	x=0,1,2,...
Hipergeometrik	$\frac{\binom{M1}{x}\binom{M2}{n-x}}{\binom{M1+M2}{n}}$	$0 \leq p \leq 1$	x=0,1,...,n

X kesikli tesadüf değişkeninin x gibi bir değer eşit ya da ondan daha küçük herhangi bir değer olması olasılığını gösteren ve

$$P(X \leq x) = F(X) = \sum_{xi \leq k}^{x} f(xi) \text{ yada}$$

$$F(x) = \sum_{t=-\infty}^{x} f(t)$$

şeklinde tanımlanan fonksiyona dağılım fonksiyonu denir. (3)"

Tanımlardan da görüleceği gibi bir kesikli değişkene ait dağılım fonksiyonu, herhangi bir X değerine eşit yada ondan daha küçük olan olasılık (yoğunluk) fonksiyonlarının eklenmiş değerlerinden oluşur.

TABLO 2 : KESİKLİ DAĞILIM FONKSİYONLARI İÇİN ALGORİTMA	
Adım 1	: Dağılımın parametlerini veriniz, x değişim aralığını tespit ediniz.
Adım 2	: f(x) olasılık fonksiyonunu hesaplayınız (k=1, ..., x)
Adım 3	: F(x) = F(x) + f(x) (k=1, ...,x)
Adım 4	: x, f(x) ve F(x)'i yazınız

Tablo-2'de verilen algoritma tamamen yukarıda verdiğimiz tanımlardan yola çıkarak geliştirilmiş ve diğer kesikli dağılımlar için de uygulanabilecek bir algoritmadır.

Algoritmada, fonksiyon değeri hesaplanacak olan dağılımın parametleri ve değişim aralığı belirlendikten sonra, istenen değişim aralığında ve parametrelere bağlı olarak f(x) olasılık (yoğunluk) fonksiyon değerleri hesaplanmakta ve bu değerler F(x) de üst üste eklenerek dağılım fonksiyonları elde edilerek, sonuçlar yazdırılmaktadır. Tablo 2'den de anlaşılacağı üzere algoritmanın karmaşıklık fonksiyonu**

C.0(x) dir.

Yukarıda verdiğimiz algoritmanın uygulanabilirliğini gösterebilmek amacıyla, en çok uygulama alanı bulunan kesikli dağılımlardan binomiyal ve poisson dağılımları için Fortran dilinde kodlama yapılarak, Program 1 ve Program 2 adlarıyla listeleri verilen programlar bir uygulama Problemi üzerinde çalıştırılmıştır.

Aşağıda bu programlar ve uygulamalı çözümleri hakkında bilgi verilmektedir;

3. BİNOM VE POİSSON DAĞILIM FONKSİYONU HESAPLAYAN PROGRAMLAR

PROGRAM 1: BİNOM OLASILIK YOĞUNLUK VE DAĞILIM FONKSİYONU HESAPLAYAN PROGRAM

```
C VERİLEN N, P PARAMETRELERİ İÇİN
C BİNOM DAĞILIM FONKSİYONU VE OLASILIK FONKSİYONU HESAPLAR.
  OPEN (1, FILE='PRN', STATUS='NEW')
  WRITE (*,*)'N, P DEĞERLERİNİ GİRİNİZ'
  READ (*,1) N, TETA
  WRITE (1,55)
55 FORMAT (5X,'N=', I3,3X,'VE P=',1X,F4.2,2X,
  D'İÇİN BİNOMİYAL DAĞILIMA AİT')
  WRITE (1,56)
56 FORMAT (/,10X,'OLASILIK FONKSİYONU',10X,'DAĞILIM FONKSİYONU')
  WRITE (1,57)
57 FORMAT (10X,'------------------------------------------------------------')
```

(**) Algoritma karmaşıklığı kavramı çeşitli şekillerde tamamlanabilir. Bu tanımlardan birisi "Bir sorunu çözecek algoritmanın giriş verilerinin o işletmede en çok işlevi gerektirecek (en olumsuz) dağılımı gösterdiği varsayılarak, işlemler için üst sınır fonksiyonu bulmaktır" şeklinde ifade edilmektedir (4).

```
1  FORMAT (I3,F4.2)
   N1=N+1
   T=0
   DO 2 J=1,N1
   J1=J-1
   I1=N-J1
   PR=FAKT(N)*TETA**J1*(1.-TETA)**I1/(FAKT(J1)*FAKT(I1)
   T=T=PR
2  WRITE(1,4) J1,PR,T
4  FORMAT (5X,I3,5X,F11.5,10X,F11.5)
   STOP
   END
   FUNCTION FAKT(N)
   FAKT=1
   IF(N.LT.1) GOTO 10
   DO 1 I=1,N
   AI=I
1  FAKT=FAKT*AI
10 RETURN
   END
```

PROGRAM 2: POISSON OLASILIK (YOĞUNLUK) VE DAĞILIM FONKSİYONU HESAPLAYAN PROGRAM

```
C BU PROGRAM VERİLEN ALFA PARAMETRESİNE BAĞLI OLARAK
C VERİLEN DEĞİŞİM ARALIĞINDA POISSON DAĞILIM FONKSİYONU VE
C OLASILIK (YOĞUNLUK) FONKSİYONU HESAPLAR.
   OPEN (1, FILE='PRN', STATUS='NEW')
   WRITE(*,*)'ALFA PARAMETRENİZİ VERİNİZ'
```



```
   READ (*,100) ALFA
   WRITE (*,*)'DEĞİŞİM ARALIĞINI VERİNİZ'
   READ (*,103) NX
103 FORMAT (I3)
100 FORMAT (F4.2)
 7 alf=-ALFA
   TOP=0
   WRITE (1,101) ALFA
101 FORMAT (2X,'ALFA=', 2X, F4.2,2X,'İÇİN POISSON DAĞILIMINA AİT',/,8X,
   D 'OLASILIK FONKSIYONU', 10X,'DAĞILIM FONKSİYONU')
   WRITE (1,102)
102 FORMAT(8X,50('-'))
   DO 10 K=0,NX
     TERİM = EXP(ALF)*(-ALF)**K/KFAKT(K)
     TOP=TOP+TERIM
     WRITE(1,8) K, TERIM, TOP
10 CONTINUE
 8 FORMAT (2X,I3,5X,F5.3,25X,F6.3)
   STOP
   END


   FUNCTION KFAKT(N)
   KFAKT=1
   DO 5 I=2,N
5  KFAKT=KFAKT*1
   RETURN
   END
```

UYGULAMA : Bilgisayar çipleri üreten bir firmada üretilen IC çiplerinin % 10'unun kusurluğu olduğu tespit edilmiştir. Firmanın tesadüfi olarak seçtiği 20 örneklem içinde 5'den daha fazla kusurlu çip bulunması olasılğı nedir? (5)

AÇIKLAMA : Bu problemde istenen olayı kusurlu çip bulunması olarak kabul edecek olursak,

örneklem genişliği n=20,

kusurlu çip bulunması olasılığı p=10/100 = 0.10

kusurlu çip bulunmaması olasılığı q = 1-0.10 = 0.90

istenen olayın gerçekleşme olasılığı : p(x>5)=?

Binom dağılımına uyan bu problem için yukarıda listesi verilen Program1'i çalıştıracak olursak dağılımın parametreleri n için 20, p için 0.10 değerlerini terminalden girersek x=0,1,......,n değişim aralığında elde edilen çıktı şekil-1'deki gibi olur.

N=20 ve P= .10 İÇİN BİNOMİYAL DAĞILIMA AİT

X	OLASILIK FONKSİYONU	DAĞILIM FONKSİYONU
0	.12158	.12158
1	.27017	.39175
2	.28518	.67693
3	.19012	.86705
4	.08978	.95683
5	.03192	.98875
6	.00887	.99761
7	.00197	.99958
8	.00036	.99994
9	.00005	.99999
10	.00001	1.00000
11	.00000	1.00000
12	.00000	1.00000
13	.00000	1.00000
14	.00000	1.00000
15	.00000	1.00000
16	.00000	1.00000
17	.00000	1.00000
18	.00000	1.00000
19	.00000	1.00000
20	.00000	1.00000

ŞEKİL - 1

Şekil-1'den de görüldüğü gibi 0-20 arasındaki x değişik değerleri için birinci sütunda

$$p(X=x) = \binom{20}{x} (0,1)^x (0,9)^{20-x}$$

olasılık (yoğunluk) fonksiyonları ikinci sütunda,

$$p(X\leq x) = \sum_{k=0}^{x} \binom{n}{x} p^x q^{n-x}\ , \ x=0,1,...,n$$

dağılım fonksiyonları elde edilmektedir.

Tablodaki x=5'e karşı gelen 0.03192 tek bir kusurlu çip üretme olasılığı 0.10 olan üretimden alınan 20 birimlik örneklem içinde, 5 kusurlu çip bulunması olasılığıdır.

Aynı örneklem içinde 5 yada daha az sayıda kusurlu çip bulunma olasılığı ise 0.98875 yada diğer deyişle 20 birim içinde 5'den daha çok sayıda kusurlu çip olması olasılığı 1-0.98875=0.01 şeklinde de ifade edilebilir. Şekil-1'de görüleceği gibi sadece x=5 için değil de 0-20 arasındaki bütün x değerleri için bu tablodan benzer şekilde yararlanılabileceği aşikardır.

Bilindiği üzere Binom dağılımını belirleyen parametrelerden birisi olan n çok büyük, p de çok küçük olduğunda ($n\to\infty$, $p\to 0$) olasılıkların hesaplanması oldukça güçtür. Böyle durumlarda binom dağılımı yerine "gerçekleşme olasılığı çok küçük olan olayların tekrarlı denemeleri için uygun bir dağılım olan" poisson dağılımı kullanılır.

Şimdi yukarıda verilen problemi n=20'yi büyük kabul ederek poisson dağılımına uygulayalım. Bu durumda α=n.p = 20*(0.10)=2 olacaktır. Program 2'yi çalıştırarak α parametresi sorulduğunda terminalden 2 değerini girdiğimizde elde edilen çıktı şekil-2'deki gibi olur.

ALFA = 2.00 İÇİN POISSON DAĞILIMINA AİT

X	OLASILIK FONKSİYONU	DAĞILIM FONKSİYONU
0	.135	.135
1	.271	.406
2	.271	.677
3	.180	.857
4	.090	.947
5	.036	.983
6	.012	.995
7	.003	.999
8	.001	1.000
9	.000	1.000
10	.000	1.000
11	.000	1.000
12	.000	1.000
13	.000	1.000
14	.000	1.000
15	.000	1.000
16	.000	1.000
17	.000	1.000
18	.000	1.000
19	.000	1.000
20	.000	1.000

ŞEKİL - 2

Şekil-2'de 0-20 arasındaki x değişim aralığında birinci sütunda

$$p(X=x)=f(x)=\frac{e^{-\alpha}\alpha^{x}}{x!}$$

olasılık (yoğunluk) fonksiyonları ve ikinci sütunda

$$p(X\leq x)=\sum_{k=0}^{x}\frac{e^{-\alpha}\alpha^{x}}{x!}$$

dağılım fonksiyonları elde edilmektedir.

Tabloda x=5'e karşı gelen 0.036 tek bir kusurlu çip üretme olasılığı 0.10 olan üretimden alınan 20 birimlik örneklem içinde 5 kusurlu çip bulunması olasılığıdır. Aynı örneklem içinde 5 yada daha az sayıda kusurlu çip bulunma olasılığı ise 0.983, diğer deyişle 5'den daha fazla sayıda kusurlu çip bulunma olasılığı ise 1-0.983=0.017'dir. Benzer şekilde 0-20 aralığındaki diğer x değerleri için de aynı tablodan yararlanılabilir.

Buradan da görüleceği gibi n değeri büyüdükçe binom ve poisson dağılımlarına ilişkin olasılık fonksiyon ve dağılım fonksiyon değerleri birbirine çok yakın değerler çıkmaktadır(6).

Demek ki n'nin büyük (n≥20) değerleri için binom dağılım yerine poisson dağılımın uygulanması çok büyük sakınca yaratmamaktadır(6).

SONUÇ

Bu çalışmada, kesikli tesadüf değişkenlere ilişkin olasılık yoğunluk ve dağılım fonksiyonu hesaplayan bir algoritma geliştirilmiş, bu algoritmanın binomiyal ve poisson dağılımları için kodlanması yapılarak uygulanabilirliği gösterilmiştir. Ayrıca binomiyal ve poisson dağılımlarına uyan bir problem, hazırlanan programlar aracılığı ile çözüme ulaştırılmıştır.

Gerek sürekli dağılımlar için gerekse kesikli dağılımlar için üzerinde çalıştığımız probleme uygun olasılık yoğunluk fonksiyonunu ve dağılım fonksiyonunu elle hesap etmek oldukça zor ve zaman alıcı bir iştir. Bu amaçla geliştirilmiş istatistiksel dağılım tabloları olsa da, özellikle kesikli dağılımlar için tablolara ulaşmak güçtür. Tablolara ulaşılabilse dahi bazı durumlarda üzerinde çalıştığımız probleme uygun parametreler ve değişim aralığında hesaplanmış olasılık yoğnuluk ve dağılım fonksiyonu değerlerini tablolardan elde edemeyebiliriz.

Bütün kesikli dağılımlar için uygulamaya koyabileceğiniz ve karmaşıklık fonksiyonu C.0 (x) olan bu algoritma sayesinde, kullanıcı tarafından belirlenen parametre ve değişim aralığı ile olasılık (yoğunluk) ve dağılım fonksiyonu kolayca hesaplanabilmektedir.

Binomial ve Poisson dağılımları için sınırlayarak uygulamaya koyduğumuz algoritma, hipergeometrik, geometrik, multinomial vb. diğer kesikli dağılımlar için de aynı mantıkla yeniden kodlanabilir.

Böylece üzerinde çalıştığımız problemin hangi kesikli dağılıma uygun bir model olduğuna karar verdikten sonra uygun program seçilerek, elle hesap yapmaya gerek kalmadan ve istatistiksel dağılım tablolarına ihtiyaç duymadan olasılık ve dağılım fonksiyon değerlerini elde edebiliriz.

KAYNAKLAR

1- Myers L.B. ve Enrick N.; Statistical Functions, USA 1970 S. 46

2- Ersoy N. ve Oral S.; Olasılık ve İstatistiğe Giriş, Ankara-1991 S. 135

3- Saracoğlu B. ve Çevik F.; Matematiksel İstatistik, Ankara-1985 S.86

4- Karakoç Ü.; "Veri Yapıları ve Algoritmalar", Ankara-1987- S.26

5- Kishor S.T.; "Probability and Statistics with Reliability, Queuing, and Computer Science Aplications", USA, 1982, S. 64-65

6- Kavuncu O.; "Teorik Dağılımlar Ders Notları", Ankara, S. 21-25