8-5 توليد البيانات للتوزيعات المتقطعة

يحتوي المتغير العشوائي المتقطع على دالة رياضية تسمى دالة الكتلة الاحتمالية الهدف منها حساب احتمال كل قيمة من قيم المتغير العشوائي المستمر X وكذلك يمكن حساب دالة التوزيع او دالة التراكم الاحتمالي لكل توزيع احتمالي وسوف نقدم في هذا الفصل توليد البيانات العشوائية لبعض التوزيعات الاحتمالية المتقطعة الاكثر شيوعاً وهذه التوزيعات هي (التوزيع برنولي ، التوزيع ثنائي الحدين ، توزيع بواسون ، التوزيع الهندسي) وطريقة توليد البيانات العشوائية لهذه التوزيعات باستعمال طريقة الرفض والقبول وطريقة التحويل المعكوس .

التوزيع الهندسي Geometric Distribution :

افرض ان X يتبع التوزيع الهندسي بدالة الكتلة الاحتمالية الاتية:

$$p(x) = p(1-p)^{x-1}, \quad x = 1,2,\ldots$$

اذ ان $0 < p < 1$

وان الدالة التوزيعية التجميعية هي:

$$F(x) = 1-(1-p)^{x} \quad ,x = 1,2,\ldots$$

وان التوقع والتباين للتوزيع هي الاتي:

- توليد الاعداد العشوائية باستعمال طريقة التحويل المعكوس

سوف نستعمل خوارزمية التحويل المعكوس لتوليد بيانات تتبع التوزيع الهندسي وعلى النحو الاتي:

١. توليد الأعداد العشوائية التي تتبع التوزيع المنتظم (u_i) على المدة $(0,1)$
٢. يتم مساواة قيم U_i بدالة التوزيع وكما في المعادلة الاتية :

$$u_i = F(x_i)$$

٣. تحويل العدد العشوائي المنتظم بطريقة المعكوس وكما مبين في المعادلة الآتية :

$$x_i = F^{-1}(u_i)$$

للحصول على متغير عشوائي يصف الأنموذج تحت التجربة .
ولتحويل الاعداد العشوائية الى بيانات تتبع توزيع الهندسي بأسلوب رياضي وعلى النحو الاتي:

$$\mathrm{u}_i = F(x_i)$$

$$\mathrm{u}_i = 1 - q^{x+1}$$

$$q^{x+1} = 1 - \mathrm{u}$$

وبأخذ اللوغاريتم الطبيعي للطرفين نحصل على الاتي:

$$(x+1)\log(q) = \log(1-u)$$

$$\mathrm{x}\log(q) + \log(q) = \log(1-u)$$

$$x log(q) = \log(1-u) - log(q)$$

ان الدالة المولدة للارقام العشوائية التي تتبع التوزيع الهندسي، هي:

$$X = 1 + \left[\frac{\ln(1-R)}{\ln(1-p)} - 1\right]$$

وبعد اشتقاق معادلة توليد الاعداد العشوائية للتوزيع الهندسي نأتي الى كتابة البرنامج الخاص بهذه الصيغة باستعمال لغة البرمجة MATLAB وعلى النحو الاتي:

```
clc
clear all
n=input('n=');
p=input('p=');
q=1-p;
for i=1:n
    u(i)=rand;
    x(i)=ceil((log(1-u(i))/log(q))-1);
end
x'
hist(x)
```

نلاحظ تم استعمال الايعاز ceil لتقريب العدد لأقرب عدد صحيح باتجاه الموجب لان المتغير هو عشوائي متقطع يأخذ الاعداد ...,0,1,2 لذلك لابد من تقريب الاعداد.

وبعد تنفيذ البرنامج بوضع n=10 و p=0.4 حصلنا على النتائج الاتية :

```
Command Window
n=10
p=0.4

ans =

     0
     2
     1
     0
     0
     1
     0
     2
     0
     4
```

مثال: كون عينة عشوائية من عشرة ارقام تتبع التوزيع الهندسي بمعدل قدره 2، بالاستناد الى الاتي من الارقام العشوائية التي تتبع التوزيع المنتظم القياسي. ومن ثم قارن بين الوسط الحسابي والانحراف المعياري للعينة مع الوسط الحسابي والانحراف المعياري النظري.

الحل: بما ان

$$E(X) = \frac{1}{p}$$

$$2 = \frac{1}{p};\ p = \frac{1}{2}$$

	R	$\frac{\ln(1-R)}{\ln(1-p)} - 1$	$X = 1 + \left[\frac{\ln(1-R)}{\ln(1-p)} - 1\right]$
1	0.310599	-0.46342	1
2	0.441619	-0.15932	1
3	0.646532	0.50035	2
4	0.053014	-0.92142	1
5	0.599849	0.32139	2
6	0.534656	0.10363	2
7	0.989918	5.63208	7
8	0.386234	-0.29576	1
9	0.793572	1.27629	3
10	0.643247	0.48700	2

$$\bar{X} = \frac{\sum x_i}{n} = 2.2$$

$$SD = 1.81353$$

$$E(X) = 2$$

$$Var(X) = \frac{1-p}{p^2} = \frac{1-0.5}{0.5^2} = 2$$

$$\sigma = 1.4142$$

توزيع برنولي Bernoulli Distribution

ليكن المتغير العشوائي X يتبع توزيع برنولي بالمعلمة p، فان دالة الكتلة الاحتمالية للتوزيع هي:

$$p(X) = p^x(1-\mathrm{p})^{1-x} \quad , x = 0,1$$

$$p(X=1) = 1 - p(X=0) = p$$

وان الوسط الحسابي والتباين للتوزيع هو

$$E(X) = p$$

$$Var(X) = pq$$

سوف يتم استعمال طريقة الرفض – القبول لتوليد الاعداد العشوائية التي تتبع توزيع برنولي وحسب الخوارزمية الاتية :

الخطوة 1: ولد ارقام عشوائية تتبع التوزيع المنتظم القياسي R.

الخطوة 2: اذا كان $R \leq p$، عندئذ $X = 1$، وبخلافه فان $X = 0$.

مثال:

كون عينة عشوائية من خمسة ارقام تتبع توزيع برنولي بالاستناد الى الاتي من الارقام العشوائية التي تتبع التوزيع المنتظم القياسي، علما ان $p = 0.3$. قارن بين الوسط الحسابي والانحراف المعياري للعينة مع الوسط الحسابي والانحراف المعياري النظري.

R		X
0.595808	$R \nleq p$	0
0.930725	$R \nleq p$	0
0.838439	$R \nleq p$	0
0.124105	$R \leq p$	1
0.834745	$R \nleq p$	0

$$\bar{X} = \frac{1}{5} = 0.2$$

$$SD = 0.447214$$

$$E(X) = 0.3$$

$$Var(X) = 0.3 * 0.7 = 0.21$$

$$\sigma = 0.4582$$

توزيع ثنائي الحدين .Binomial Dist:

ليكن المتغير العشوائي X يتبع توزيع ثنائي الحدين بالمعلمتين n و p، فان دالة الكتلة الاحتمالية للتوزيع هي:

$$p(X) = \binom{n}{x} p^x (1-\mathrm{p})^{n-x} \quad , X = 0,1,\dots,n$$

وان الوسط الحسابي والتباين للتوزيع هو

$$E(X) = np$$

$$Var(X) = npq$$

سوف يتم استعمال طريقة الرفض – القبول لتوليد الاعداد العشوائية التي تتبع توزيع برنولي وحسب الخوارزمية الاتية :

1. ولد ارقام عشوائية تتبع توزيع برنولي بالمعلمة p للمتغيرات العشوائية $Y_1, Y_2, \dots, Y_n$.
2. ضع $X = Y_1 + Y_2 + \cdots + Y_n$.

مثال : كون عينة عشوائية تتبع توزيع ثنائي الحدين بالمعلمتين $p = 0.3, n = 3$ بالاستناد الى الاتي من الارقام العشوائية التي تتبع التوزيع المنتظم القياسي (R). ثم قارن بين الوسط الحسابي والانحراف المعياري للعينة مع الوسط الحسابي والانحراف المعياري النظري.

الحل:

R	Bernoulli	Binomial (3, 0.3)
0.653413	**0**	0
0.774356	**0**	
0.945399	**0**	
0.330428	0	2
0.032944	1	
0.246109	1	
0.867240	**0**	2
0.022977	**1**	
0.054329	**1**	
0.924427	0	1
0.251232	1	
0.740529	0	
0.701664	**0**	1
0.430707	**0**	
0.122860	**1**	
0.785539	0	0
0.354465	0	
0.338171	0	
0.349359	**0**	1
0.387035	**0**	
0.242699	**1**	
0.991664	0	1
0.718505	0	
0.270150	1	
0.957290	**0**	2
0.221598	**1**	
0.123605	**1**	
0.721382	0	2
0.249868	1	
0.060961	1	
0.515272	**0**	0
0.768593	**0**	
0.394115	**0**	

$$E(X) = np = 3 * 0.3 = 0.9$$

$$Var(X) = \mathrm{npq} = 3 * 0.3 * 0.7 = 0.63$$

$$\sigma = 0.7937$$

$$\bar{X} = 1.09091$$

$$SD = 0.831209$$

وبعد كتابة الخوارزمية الخاصة بتوليد الاعداد العشوائية التي تتبع توزيع برنولي نأتي الان الى كتابة البرنامج الخاصة بها باستعمال MATLAB وعلى النحو الاتي:

```
clc
clear all
n=input('n=');
p=input('p=');
for i=1:n
  u(i)=rand;
  if u(i)<=p
      x1(i)=1;
  else
      x1(i)=0;
  end
        x(i)=sum(x1);
end
x'
hist(x)
```

وبعد كتابة البرنامج يتم تنفيذ البرنامج بوضع n=10,p=0.4 نحصل على النتائج الاتية :

```
Command Window
n=10
p=0.4

ans =

     0
     0
     0
     0
     0
     1
     1
     1
     2
     2
```

مثال : كون عينة عشوائية تتبع توزيع ثنائي الحدين بالمعلمتين $p = 0.3, n = 3$ بالاستناد الى الاتي من الارقام العشوائية التي تتبع التوزيع المنتظم القياسي (R). ثم قارن بين الوسط الحسابي والانحراف المعياري للعينة مع الوسط الحسابي والانحراف المعياري النظري.

الحل بالاستناد الى الدالة التجميعية:

الخطوة 1: ولد ارقام عشوائية تتبع التوزيع المنتظم القياسي (R).

الخطوة 2: اذا كانت $F(x_i) < R_i \leq F(x_i + 1)$ ، عندئذ ضع $X = x_i + 1$.

اذ ان:

$$F(X) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-\mathrm{p})^{n-k}$$

$$\begin{aligned}
p(X=0) &= 0.343;\ p(X \leq 0) = 0.343 \\
p(X=1) &= 0.441;\ p(X \leq 1) = 0.784 \\
p(X=2) &= 0.189;\ p(X \leq 2) = 0.973 \\
p(X=3) &= 0.027;\ p(X \leq 3) = 1.000
\end{aligned}$$

$$X = \begin{cases} 0 & 0 < R \leq 0.343 \\ 1 & 0.343 < R \leq 0.784 \\ 2 & 0.784 < R \leq 0.973 \\ 3 & 0.973 < R \leq 1 \end{cases}$$

R	Binomial (3, 0.3)
0.653413	**1**
0.774356	**1**
0.945399	**2**
0.330428	0
0.032944	0
0.246109	0
0.867240	**2**
0.022977	**0**
0.054329	**0**
0.924427	2
0.251232	0
0.740529	1
0.701664	**1**
0.430707	**1**
0.122860	**0**
0.785539	2
0.354465	1
0.338171	0

0.349359	**1**
0.387035	**1**
0.242699	**0**
0.991664	3
0.718505	1
0.270150	0
0.957290	**2**
0.221598	**0**
0.123605	**0**
0.721382	1
0.249868	0
0.060961	0
0.515272	**1**
0.768593	**1**
0.394115	**1**

$$E(X) = np = 3 * 0.3 = 0.9$$

$$Var(X) = \mathrm{npq} = 3 * 0.3 * 0.7 = 0.63$$

$$\sigma = 0.7937$$

$$\bar{X} = 0.787879$$

$$\mathrm{SD} = 0.819969$$

توزيع ذي الحدين السالب .Negative Binomial Dist:

ليكن المتغير العشوائي X يتبع توزيع ذي الحدين السالب بالمعلمتين r و p، فان دالة الكتلة الاحتمالية للتوزيع هي:

$$p(X) = \binom{r+x-1}{x} p^r (1-\mathrm{p})^x \quad , x = 0,1, \ldots$$

وان الوسط الحسابي والتباين للتوزيع هو

$$E(X) = \frac{r}{p}$$

$$Var(X) = \frac{rq}{p^2}$$

خوارزمية توليد ارقام عشوائية تتبع توزيع ثنائي الحدين السالب

1. ولد ارقام عشوائية تتبع التوزيع الهندسي بالمعلمة p للمتغيرات العشوائية $Y_1, Y_2, \ldots, Y_n$.
2. ضع $X = Y_1 + Y_2 + \cdots + Y_n$.

مثال : من خلال الارقام العشوائية المولدة الاتية التي تتبع التوزيع الهندسي المطلوب توليد بيانات تتبع توزيع ذي الحدين السالب بالمعلمتين n=3 , p=0.3 ثم قارن بين الوسط الحسابي والانحراف المعياري للعينة مع النظري

Geometric	NB
1	6
4	
1	
5	7
1	
1	
5	11
5	
1	
1	5
3	
1	
1	9
6	
2	
2	4
1	
1	
10	18
5	
3	
1	7
4	
2	
8	14
4	
2	
5	9
1	
3	

$$E(X) = \frac{r}{p} = \frac{3}{0.3} = 10$$

$$Var(X) = \frac{rq}{p^2} = \frac{3 * 0.7}{0.3^2} = 23.33$$

$$\sigma = 4.830$$

$$\bar{X} = 9$$

$$SD = 4.32049$$