Практическое занятие № 9:
Генерация случайных чисел с заданным законом распределения

Цели занятия - закрепление теоретических знаний по следующим методам генерации случайных чисел:

- формирование выборки случайных чисел с равномерным распределением в заданном интервале;
- метод взятия обратной функции. Формирование выборки случайных чисел с экспоненциальным распределением;
- генерация нормальной случайной величины на основе центральной предельной теоремы;
- генерация случайной величины, имеющей биномиальное распределение;
- генерация случайной величины, имеющей распределение Пуассона;
- генерация случайных величин по графику функции плотности распределения.

Задание:

- на основе лекционного материала произвести в среде MS Excel генерацию случайных чисел с заданными законами распределения;
- вычислить для каждой полученной последовательности случайных чисел выборочное среднее и выборочную дисперсию; сравнить полученные значения с теоретическими значениями математического ожидания и дисперсии исследуемого распределения;
- варианты индивидуальных заданий.

Формирование выборки случайных чисел с равномерным распределением в заданном интервале

Непрерывная случайная величина X имеет равномерный закон распределения на отрезке [a, b], если её плотность вероятности φ(x) постоянна на этом отрезке и равна нулю вне его, т.е.

$$\varphi(x)=\begin{cases}\dfrac{1}{b-a} & \text{при} \quad a\le x\le b,\\ 0 \quad \text{при} & x<a,\, x>b.\end{cases}$$

M(X) = (a+b)/2
D(X) = (b-a)²/12

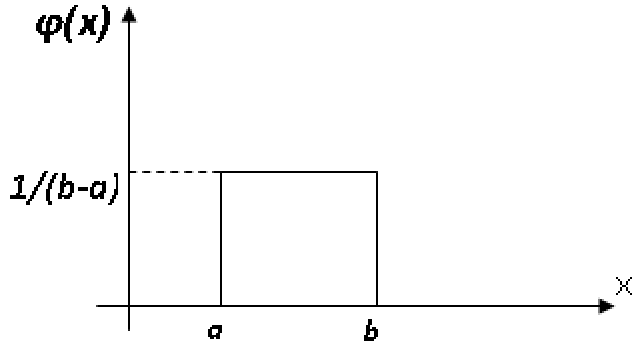


Рис. 1

Функция распределения случайной величины X, распределённой по нормальному закону:

$$F(x)=\begin{cases}0 \quad \text{при} & x\le a,\\ \dfrac{(x-a)}{(b-a)} & \text{при} \quad a<x\le b,\\ 1 \quad \text{при} & x>b.\end{cases}$$

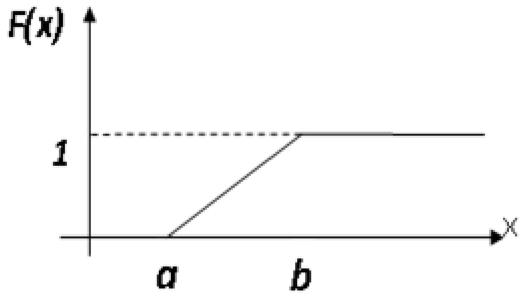


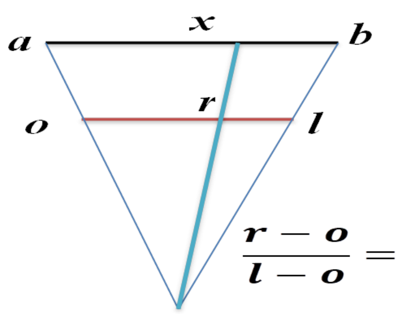
Рис. 2

Математическое ожидание $\boldsymbol{m_r}$ и дисперсия $\boldsymbol{D_r}$ такой последовательности, состоящей из $\boldsymbol{n}$ случайных чисел $\boldsymbol{r_i}$, должны быть следующими:

$$m_r=\frac{\sum_{i=1}^{n} r_i}{n}=0.5$$

$$D_r=\frac{\sum_{i=1}^{n}(r_i-m_r)^2}{n}=\frac{1}{12}$$

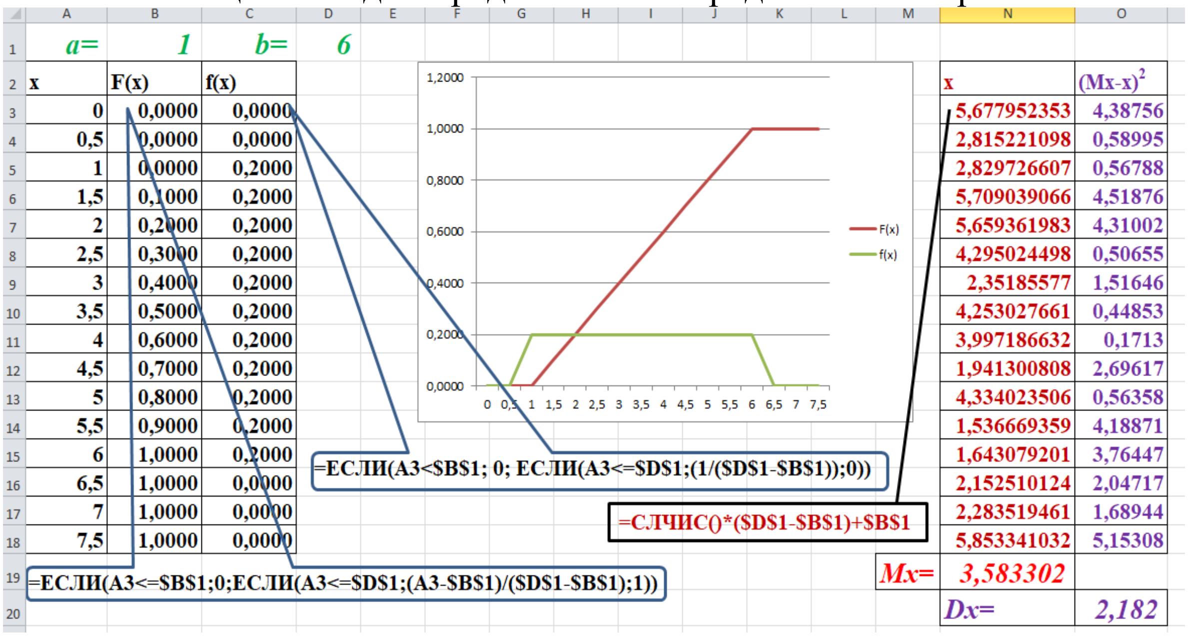
Если пользователю потребуется, чтобы случайное число x находилось в интервале (a; b), отличном от (0; 1), нужно воспользоваться формулой (рис. 3) $\boldsymbol{x = a + (b - a)\, r}$,
где r — случайное число из интервала (0; 1).



$$\frac{r-o}{l-o}=\frac{x-a}{b-a} \quad => x=a+(b-a)r$$

Рис. 3

Реализация метода в среде MS Excel представлена на рис 4.



a=	1	b=	6
x	F(x)	f(x)	
0	0,0000	0,0000	
0,5	0,0000	0,0000	
1	0,0000	0,2000	
1,5	0,1000	0,2000	
2	0,2000	0,2000	
2,5	0,3000	0,2000	
3	0,4000	0,2000	
3,5	0,5000	0,2000	
4	0,6000	0,2000	
4,5	0,7000	0,2000	
5	0,8000	0,2000	
5,5	0,9000	0,2000	
6	1,0000	0,2000	
6,5	1,0000	0,0000	
7	1,0000	0,0000	
7,5	1,0000	0,0000	

x	$(Mx-x)^2$
5,677952353	4,38756
2,815221098	0,58995
2,829726607	0,56788
5,709039066	4,51876
5,659361983	4,31002
4,295024498	0,50655
2,35185577	1,51646
4,253027661	0,44853
3,997186632	0,1713
1,941300808	2,69617
4,334023506	0,56358
1,536669359	4,18871
1,643079201	3,76447
2,152510124	2,04717
2,283519461	1,68944
5,853341032	5,15308
Mx= 3,583302	
Dx=	2,182

Рис. 4

Метод взятия обратной функции. Формирование выборки случайных чисел с экспоненциальным распределением

Непрерывная случайная величина X имеет показательный (экспоненциальный) закон распределения с параметром λ>0, если её плотность вероятности φ(x) имеет вид:

$$\varphi(x)=\begin{cases}\lambda e^{-\lambda x} & при\ x\ge 0,\\ 0 & при\ x<0.\end{cases}$$

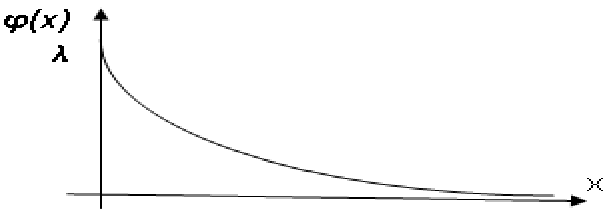


Рис. 5

$$M(X)=\frac{1}{\lambda},$$

$$D(X)=\frac{1}{\lambda^2}.$$

$$F(x)=\begin{cases}0 & при\ x<0,\\ 1-e^{-\lambda x} & при\ x\ge 0.\end{cases}$$

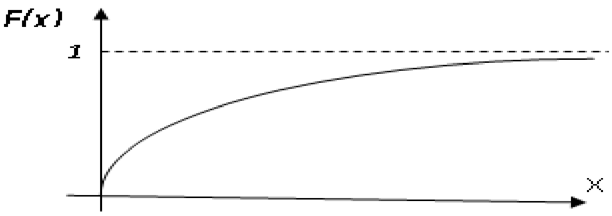


Рис. 6

Так как ***r*** и ***F*** в данном методе предполагаются аналогичными и расположены в одном интервале, то, заменяя F на случайное число r, имеем**:**

$$r = 1 - e^{-\lambda x}.$$

Тогда $\quad$ $\boldsymbol{x = -1/\lambda\ \ln(1 - r)}$.

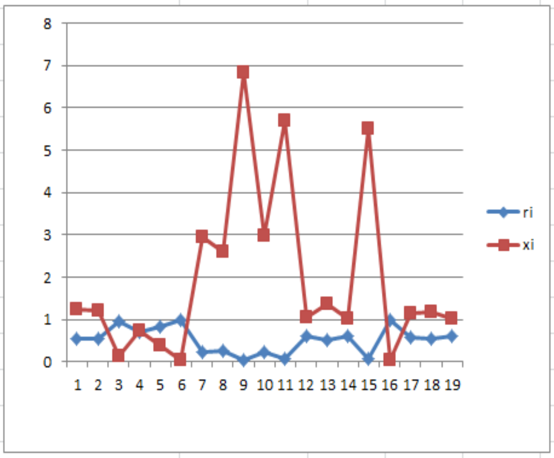
Так как в статическом смысле $\boldsymbol{(1 - r)}$ и $\boldsymbol{r}$ — это одно и тоже, то

$$\boldsymbol{x = -1/\lambda\ \ln(r)}.$$

Реализация метода в среде MS Excel при M(x)=2 представлена на рис 7.

C21 $\quad$ fx =СРЗНАЧ(C2:C20)

	A	B	C	D	E	F
1		ri	xi	$(Mx\text{-}xi)^2$	M(X)=1/ λ =	2
2		0,535870907	1,247724	0,4998004	Отсюда λ =	0,5000
3		0,542398814	1,2235075	0,5346273		
4		0,942329618	0,1188003	3,3704895		
5		0,696334201	0,7238511	1,5149633		
6		0,825792568	0,3828233	2,4707635		
7		0,973136047	0,0544628	3,610862		
8		0,227847585	2,9581567	1,0069463		
9		0,273735087	2,591189	0,4051314		
10		0,032880035	6,8297793	23,7665		
11		0,224578612	2,9870589	1,0657865		
12		0,058204846	5,6875733	13,934421		
13		0,595842025	1,0355594	0,8448003		
14		0,501062754	1,3820479	0,3279186		
15		0,6002721	1,0207445	0,8722535		
16		0,063492874	5,5136552	12,666236		
17		0,984135742	0,0319829	3,6968011		
18		0,567309303	1,1337012	0,6740219		
19		0,553814091	1,1818524	0,5972772		
20		0,596118243	1,0346325	0,8465051		
21		*Mx=*	*1,95469*			
22			*Dx=*	*4,01146*		



$$\varphi(x) = \begin{cases} \lambda e^{-\lambda x} & при \quad x \ge 0, \\ 0 & при \quad x < 0. \end{cases}$$

$$F(x) = \begin{cases} 0 & при \quad x < 0, \\ 1 - e^{-\lambda x} & при \quad x \ge 0. \end{cases}$$

= - 1/$F$2*LN(B2)

Рис. 7

Генерация нормальной случайной величины на основе центральной предельной теоремы

Непрерывная случайная величина X имеет нормальный закон распределения (закон Гаусса) с параметрами a и σ2 , если её плотность вероятности φ(x) имеет вид:

$$\varphi_N(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

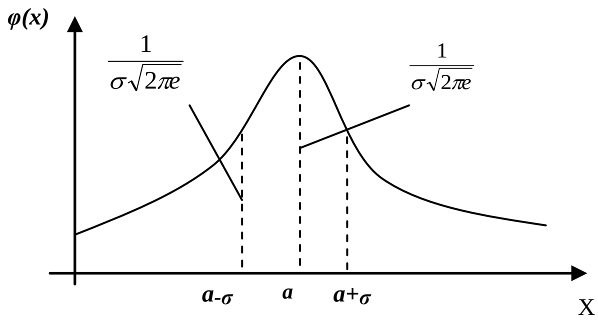


Рис. 8

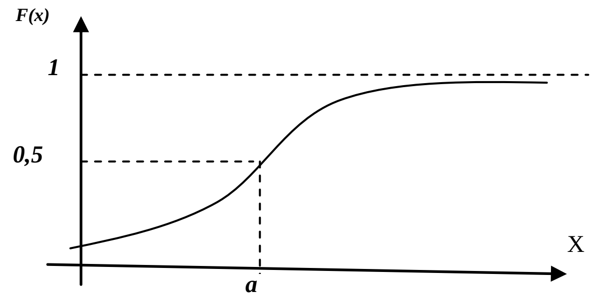


Рис. 9

Математическое ожидание случайной величины X, распределённой по нормальному закону, равно параметру a этого закона, а её дисперсия – параметру σ2, т.е. $\boldsymbol{M(X) = a;}$ $\boldsymbol{D(X) = \sigma^2}$.

Центральная предельная теорема: сумма ***n*** одинаково распределенных независимых случайных величин x со средним ***Mx*** и дисперсией ***Dx*** стремится к нормально распределенной величине с параметрами $\boldsymbol{nM_x}$ и $\boldsymbol{nD_x}$ при бесконечном увеличении ***n***.

Идея алгоритма: Определим новую случайную величину s в виде суммы базовых чисел ***Ri***, (i=1,2,3, …,n): $\boldsymbol{s = R_1 + R_2 + \ldots + R_n}$.

Согласно утверждению центральной предельной теоремы, случайная величина *s является асимптотически нормальной* величиной с математическим ожиданием и дисперсией равными соответственно:

$$\boldsymbol{M_s = n/2,\ D_s = n/12.}$$

Введем вспомогательную случайную величину z равную

$$z = \frac{(s - n/2)}{\sqrt{n/12}}.$$

z - случайная величина, распределенная по нормальному закону с нулевым средним и единичной дисперсией.

Тогда для ***любого нормального распределения со средним M_x и дисперсией σ^2 случайное отклонение y, соответствующее указанным выше n случайным числам, получается из формулы***

$$(y - M_x)/\sigma = z$$

Следовательно $y = M_x + \sigma * z =$

$$= M_x + \left(\frac{\delta}{\sqrt{\frac{n}{12}}}\right)\left(\sum_{i=1}^{n} R_i - n/2\right)$$

Согласно той же предельной теореме, нормальность достигается быстро даже при сравнительно небольших значениях n.

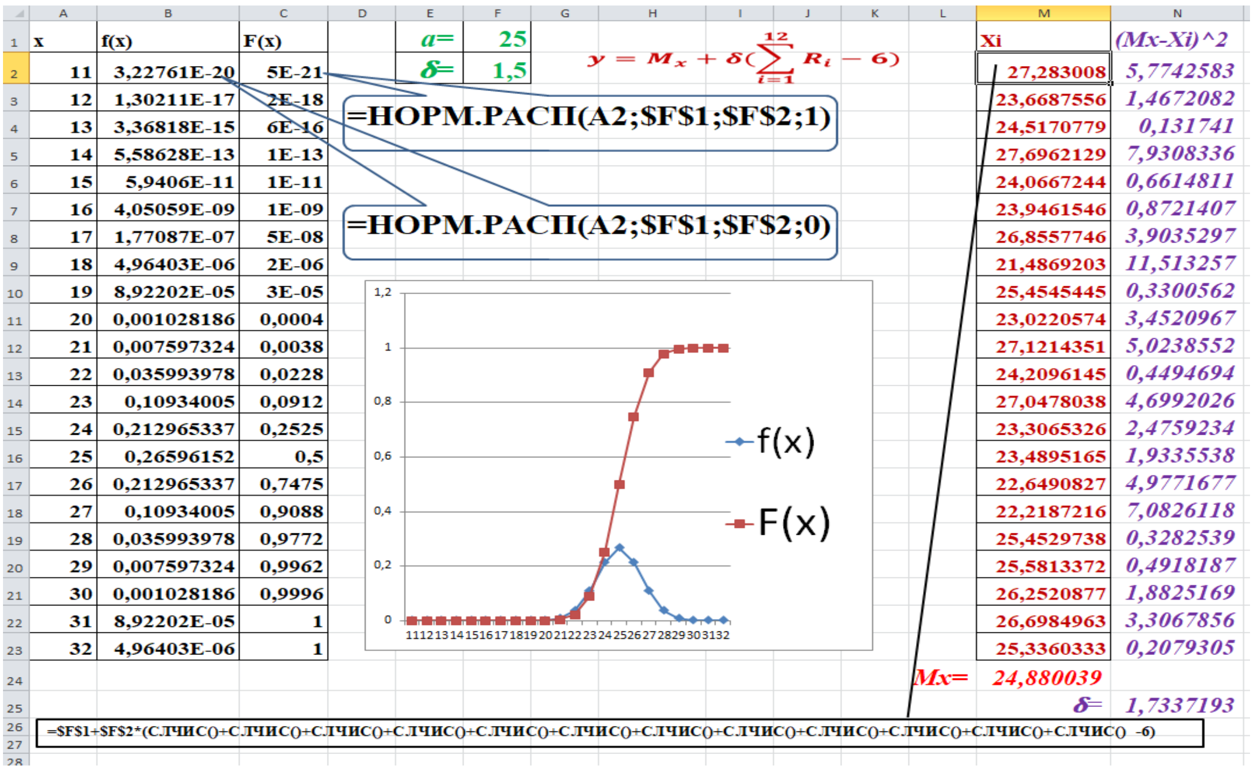
В практических задачах обычно принимается n =12.

При этом последняя формула упрощается и принимает вид

$$y = M_x + \delta\left(\sum_{i=1}^{12} R_i - 6\right)$$

и ***дает алгоритм моделирования нормальных случайных чисел с требуемыми параметрами M_x и σ^2.***

Реализация метода в среде MS Excel при M(x)=25 и =1,5 представлена на рис 10.



x	f(x)	F(x)	Xi	(Mx-Xi)^2
11	3,22761E-20	5E-21	27,283008	5,7742583
12	1,30211E-17	2E-18	23,6687556	1,4672082
13	3,36818E-15	6E-16	24,5170779	0,131741
14	5,58628E-13	1E-13	27,6962129	7,9308336
15	5,9406E-11	1E-11	24,0667244	0,6614811
16	4,05059E-09	1E-09	23,9461546	0,8721407
17	1,77087E-07	5E-08	26,8557746	3,9035297
18	4,96403E-06	2E-06	21,4869203	11,513257
19	8,92202E-05	3E-05	25,4545445	0,3300562
20	0,001028186	0,0004	23,0220574	3,4520967
21	0,007597324	0,0038	27,1214351	5,0238552
22	0,035993978	0,0228	24,2096145	0,4494694
23	0,10934005	0,0912	27,0478038	4,6992026
24	0,212965337	0,2525	23,3065326	2,4759234
25	0,26596152	0,5	23,4895165	1,9335538
26	0,212965337	0,7475	22,6490827	4,9771677
27	0,10934005	0,9088	22,2187216	7,0826118
28	0,035993978	0,9772	25,4529738	0,3282539
29	0,007597324	0,9962	25,5813372	0,4918187
30	0,001028186	0,9996	26,2520877	1,8825169
31	8,92202E-05	1	26,6984963	3,3067856
32	4,96403E-06	1	25,3360333	0,2079305

Рис. 10

Генерация случайной величины, имеющей биномиальное распределение

Дискретная случайная величина X имеет биномиальный закон распределения с параметрами n и , p если она принимает значения $0, 1, 2, \dots, m, \dots , n$ с вероятностями

$$P(X=m) = C_n^m p^m q^{n-m} ,$$

где $0<p<1, q = 1 – p$.

Биномиальный закон распределения – закон распределения числа $X=m$ наступлений события A в n независимых испытаниях, в каждом из которых оно может произойти с одной и той же вероятностью p.

Математическое ожидание случайной величины X, распределённой по биномиальному закону $M(X) = np$, а её дисперсия $D(X) = npq$.

Моделирующий алгоритм основан на представлении случайной величины X, подчиненной биноминальному закону распределения, в виде суммы n независимых случайных величин Xi, каждая из которых имеет распределение

Xi	0	1
Pi	$q=1-p$	p

Процедура получения значений случайной величины X с биномиальным распределением:

а) реализуется случайная величина Ri;

б) для каждого члена последовательности $r_1, r_2, \dots, r_n$ проверяется выполнение неравенства $r_i<p (i=1, 2, \dots, n)$.

Если неравенство выполняется, принимается $Xi=1$, в противном случае $Xi=0$.

в) вычисляется сумма $X_1 + X_2 + \dots + X_n$ значений n случайных величин Xi, которая принимается за значение случайной величины **X=S**.

При повторении этой процедуры k раз получаем последовательность значений $s_1, s_2, \dots, s_k$, которые являются реализацией биноминально распределенной случайной величины.

Реализация метода в среде MS Excel при n=20 и p=0,5 представлена на рис 11.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Число испытаний n =	20	Вероятность успеха P=	0,5															
2	i	Ri	Xi		i	Ri	Xi		i	Ri	Xi		i	Ri	Xi		i	Ri	Xi
3	0	0,085501215	1		0	0,085	1		0	0,151	1		0	0,112	1		0	0,808	0
4	1	0,923876978	0		1	0,075	1		1	0,788	0		1	0,009	1		1	0,599	0
5	2	0,445248301	1		2	0,826	0		2	0,523	0		2	0,879	0		2	0,645	0
6	3	0,358545159	1		3	0,8	0		3	0,536	0		3	0,211	1		3	0,816	0
7	4	0,15477303	1		4	0,919	0		4	0,057	1		4	0,326	1		4	0,61	0
8	5	0,576880062	0		5	0,652	0		5	0,463	1		5	0,059	1		5	0,497	1
9	6	0,130593064	1		6	0,48	1		6	0,034	1		6	0,211	1		6	0,512	0
10	7	0,33691868	1		7	0,429	1		7	0,581	0		7	0,895	0		7	0,156	1
11	8	0,352281123	1		8	0,94	0		8	0,421	1		8	0,605	0		8	0,118	1
12	9	0,702322688	0		9	0,302	1		9	0,312	1		9	0,307	1		9	0,882	0
13	10	0,524705104	0		10	0,873	0		10	0,545	0		10	0,203	1		10	0,776	0
14	11	0,601610876	0		11	0,758	0		11	0,312	1		11	0,397	1		11	0,028	1
15	12	0,769452263	0		12	0,106	1		12	0,372	1		12	0,442	1		12	0,312	1
16	13	0,763005319	0		13	0,79	0		13	0,141	1		13	0,771	0		13	0,942	0
17	14	0,762419642	0		14	0,848	0		14	0,193	1		14	0,338	1		14	0,43	1
18	15	0,247903332	1		15	0,334	1		15	0,577	0		15	0,687	0		15	0,493	1
19	16	0,742135127	0		16	0,604	0		16	0,136	1		16	0,45	1		16	0,327	1
20	17	0,989523609	0		17	0,282	1		17	0,409	1		17	0,623	0		17	0,788	0
21	18	0,87659718	0		18	0,306	1		18	0,428	1		18	0,679	0		18	0,503	0
22	19	0,616950186	0		19	0,533	0		19	0,915	0		19	0,667	0		19	0,36	1
23	20	0,338254825	1		20	0,7	0		20	0,312	1		20	0,457	1		20	0,929	0
24		ΣXi=	9			ΣXi=	9			ΣXi=	14			ΣXi=	13			ΣXi=	9
25																			
26	Реализации биноминально распределенной случ. величины:								9	9	14	13	9						
27	M(X) = np =	10	=ЕСЛИ(B3<$D$1;1;0)					M(X) =		10,8			(Mx-Xi)^2=		3,24	3,24	10,2	4,84	3,24
28	D(X) = npq=	5						D(X) =		4,96									

Рис. 11

Генерация случайной величины, имеющей распределение Пуассона

Дискретная случайная величина X имеет распределение Пуассона с параметром ***λ>0***, если она принимает значения ***0, 1, 2, …, m,*** …(бесконечное, но счётное множество значений) с вероятностями

$$P(X=m)=\frac{\lambda^m e^{-\lambda}}{m!}=P_m(\lambda).$$

Если вероятность события ***A*** в каждом испытании стремится к нулю при неограниченном увеличении числа испытаний, причём произведение ***np*** стремится к постоянному числу ***λ*** (***np*** → ***λ***), то вероятность ***Pm,n*** того, что событие ***A*** появится ***m*** раз в ***n*** независимых испытаниях стремится к закону Пуассона.

Математическое ожидание и дисперсия случайной величины, распределённой по закону Пуассона совпадают и равны параметру ***λ*** этого закона, т.е. ***M(X) = D(X) = λ.***

<u>**Моделирующий алгоритм основывается на следующем утверждении**</u>: если случайные величины X1, X2, … независимы и все имеют экспоненциальное распределение с математическим ожиданием, равным 1, то неотрицательное целое число n, для которого выполняется неравенство

$$\sum_{i=1}^{n} x_i \le \lambda < \sum_{i=1}^{n+1} x_i ,$$

имеет распределение Пуассона с параметром ***λ***.

В связи с тем, что $X_i = -\ln r_i$,

где r_i - случайная величина R с равномерным распределением на [0;1], условие можно записать в виде

$$\prod_{i=1}^{n+1} r_i \le e^{-\lambda} < \prod_{i=1}^{n} r_i ,$$

где правая часть неравенства - произведение всех r_i, таких, что значение i – целое и выполняется соотношение $i=1 \ldots n$.

<u>Алгоритм получения случайной величины, распределенной по закону Пуассона с параметром λ:</u>

а) реализуются последовательности $r_1, r_2, \ldots, r_n$ независимых случайных величин, равномерно распределенных на [0; 1);

б) вычисляются произведения $r_1, r_1r_2, r_1r_2r_3, \ldots$ до тех пор, пока не выполнится указанное условие.

В качестве значения случайной величины X принимается число n.

Если неравенству удовлетворяет первое из равномерно распределенных чисел r_1, то $X=0$.

Реализация метода в среде MS Excel при λ=10 представлена на рис 12.

J28

Среднее λ= 10 | $e^{-\lambda}$ = 4,53999E-05 | $P(X=m)=\frac{\lambda^m e^{-\lambda}}{m!}=P_m(\lambda).$ | $\prod_{i=1}^{n+1} r_i \le e^{-\lambda} < \prod_{i=1}^{n} r_i ,$

i	Ri	П Ri		i	Ri	П Ri		i	Ri	П Ri		i	Ri	П Ri	
0	0,428955238	0,428955238	0	0	0,154	0,1537	0	0	0,635	0,635	0	0	0,802	0,80245	0
1	0,224860633	0,096455146	0	1	0,214	0,0329	0	1	0,555	0,352	0	1	0,794	0,63747	0
2	0,236432949	0,022805175	0	2	0,566	0,0186	0	2	0,793	0,279	0	2	0,225	0,14356	0
3	0,947736721	0,021613301	0	3	0,362	0,0067	0	3	0,522	0,146	0	3	0,932	0,13376	0
4	0,980890574	0,021200284	0	4	0,472	0,0032	0	4	0,774	0,113	0	4	0,52	0,0695	0
5	0,505548301	0,010717767	0	5	0,761	0,0024	0	5	0,731	0,082	0	5	0,524	0,03642	0
6	0,962360194	0,010314353	0	6	0,228	0,0006	0	6	0,353	0,029	0	6	0,406	0,01477	0
7	0,010857178	0,000111985	0	7	0,783	0,0004	0	7	0,187	0,005	0	7	0,875	0,01293	0
8	0,5155186	5,77302E-05	0	8	0,519	0,0002	8	8	0,261	0,001	0	8	0,173	0,00224	0
9	0,965684814	5,57492E-05	9	9	0,131	3E-05	0	9	0,641	9E-04	0	9	0,034	7,7E-05	9
10	0,356116927	1,98532E-05	0	10	0,399	1E-05	0	10	0,923	8E-04	0	10	0,146	1,1E-05	0
11	0,241926646	4,80303E-06	0	11	0,198	2E-06	0	11	0,984	8E-04	0	11	0,587	6,6E-06	0
12	0,632906556	3,03987E-06	0	12	0,219	5E-07	0	12	0,807	7E-04	0	12	0,299	2E-06	0
13	0,01795156	5,45704E-08	0	13	0,208	1E-07	0	13	0,964	6E-04	0	13	0,8	1,6E-06	0
14	0,531947386	2,90286E-08	0	14	0,545	6E-08	0	14	0,096	6E-05	14	14	0,819	1,3E-06	0
15	0,219496153	6,37166E-09	0	15	0,228	1E-08	0	15	0,258	2E-05	0	15	0,358	4,6E-07	0
16	0,798327043	5,08667E-09	0	16	0,971	1E-08	0	16	0,698	1E-05	0	16	0,997	4,6E-07	0
17	0,871026898	4,43062E-09	0	17	0,809	1E-08	0	17	0,214	2E-06	0	17	0,738	3,4E-07	0
18	0,96165316	4,26072E-09	0	18	0,393	4E-09	0	18	0,233	6E-07	0	18	0,72	2,5E-07	0
19	0,205266208	8,74582E-10	0	19	0,158	6E-10	0	19	0,052	3E-08	0	19	0,827	2E-07	0
20	0,024504298	2,1431E-11	0	20	0,798	5E-10	0	20	0,826	2E-08	0	20	0,313	6,4E-08	0

=B4 | =C4*B5 | X= 9 | X= 8 | X= 14 | X= 9

=ЕСЛИ(C4>$E$1;ЕСЛИ(C5<=$E$1;A4;0);0)

M(X) = D(X) = λ   M(X) = D(X) = λ= 10

Рис. 12

Генерация случайных величин по графику функции плотности распределения

Пусть функция плотности распределения случайной величины f(x) задана графически (рис. 13).

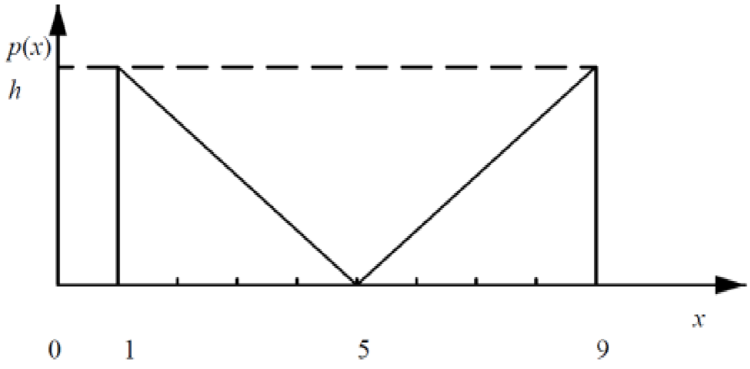


Рис. 13

На первом этапе необходимо получить аналитическое описание заданной функции.

Площадь S области, ограниченной функцией распределения плотности вероятности и осью x, равна 1. Это позволяет найти высоту h.
Площадь двух треугольников равна

S=1=1/2*h*4+1/2*h*4=4h => h=1/4

Найдем аналитическое описание функции на отрезке [1, 5].

Возьмем уравнение прямой, проходящей через две заданные точки в форме ***y=ax+b***; и найдем значения коэффициентов ***a*** и ***b***.

Для этого подставим в это уравнения координаты двух базовых точек (***x = 1, y = ¼***) (***x = 5, y = 0***).

Получим два алгебраических уравнения

¼= a*1+b;   0=a*5+b. =>   a = – 1/16, b = 5/16.

Поэтому функция плотности вероятности имеет вид

на отрезке [1, 5]   *f(x) = -1/16 x + 5/16*

на отрезке [5, 9]   *f(x) = 1/16 x - 5/16*

Отсюда

$$F(x) = \begin{cases} \frac{1}{32}(-x^2 + 10x - 9) & \text{при } x \in [1, 5] \\ \frac{1}{32}(x^2 - 10x + 41) & \text{при } x \in [5, 9] \end{cases}$$

Решая эти уравнения как квадратные относительно x, получим выражения для обратных функций

$$x_{1,2} = 5 \pm 4\sqrt{1-2F}; F \subset [0,0.5]$$
$$x_{1,2} = 5 \pm 4\sqrt{2F-1}; F \subset [0.5,1] \qquad (6)$$

Знаки в этих формулах выбираются из граничных условий. Подставляя в первую формулу $F = 0$, получим, что необходим знак минус. Подставляя во вторую формулу F =0.5, получим, что необходим знак плюс. Окончательно получим

$$x = 5 - 4\sqrt{1-2F}; F \subset [0,0.5]$$
$$x = 5 + 4\sqrt{2F-1}; F \subset [0.5,1]$$

Варианты индивидуальных заданий

В соответствии с вариантом (табл. 2) составить аналитическое выражение для функции распределения плотности вероятности, заданной в табл. 1.

Таблица 1 - Варианты функции распределения плотности вероятности.

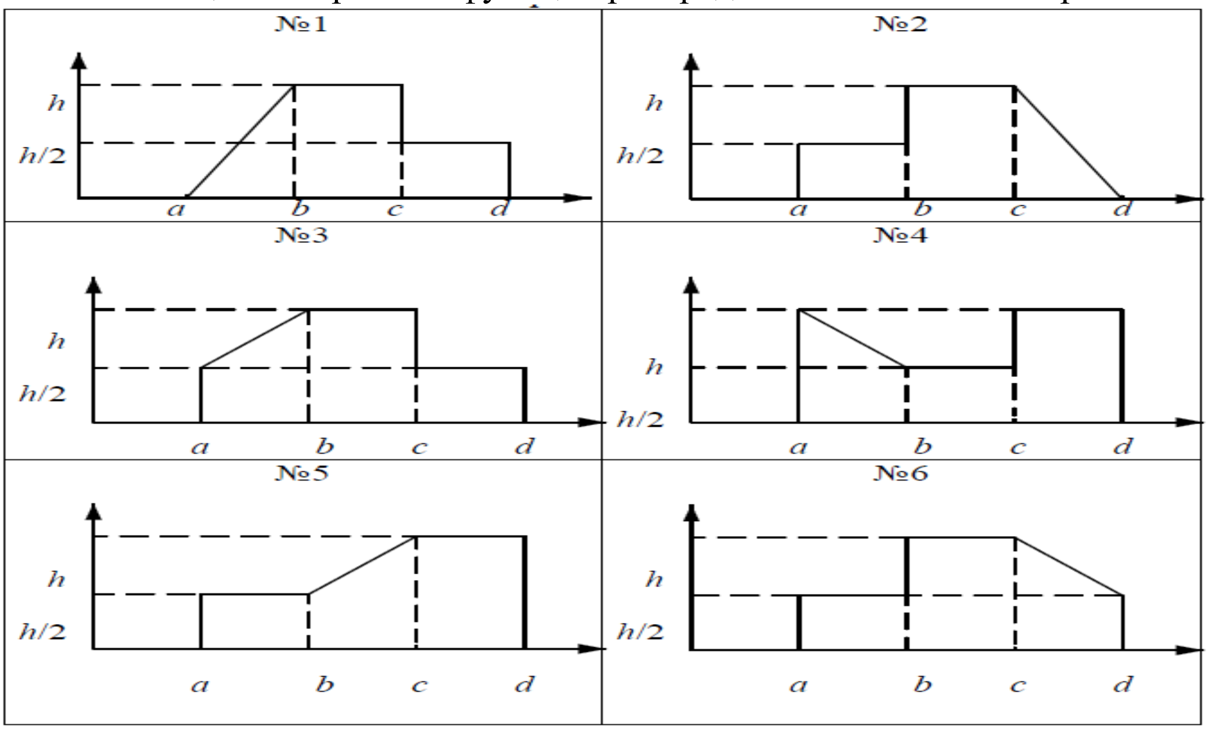


Таблица 2 - Варианты исходных данных

Вариант	*a*	*b*	*c*	*d*
1	0	1	4	6
2	0	2	4	6
3	0	1	3	5
4	0	2	3	6
5	0	1	4	5
6	0	2	3	5
7	0	1	4	6
8	0	2	4	6
9	0	1	3	5
10	0	2	3	6