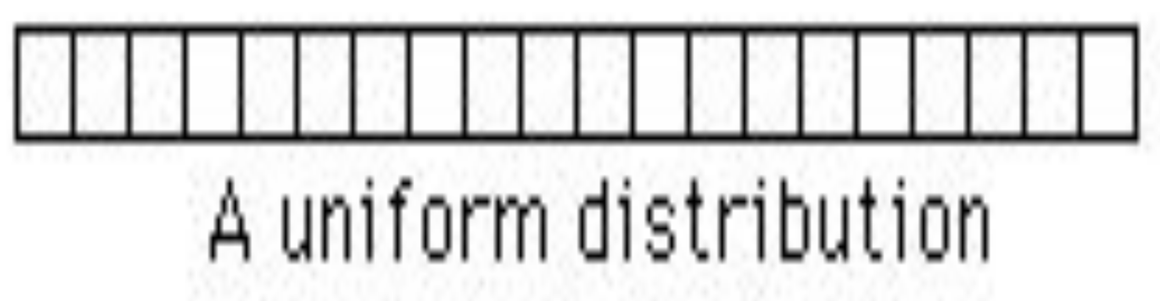
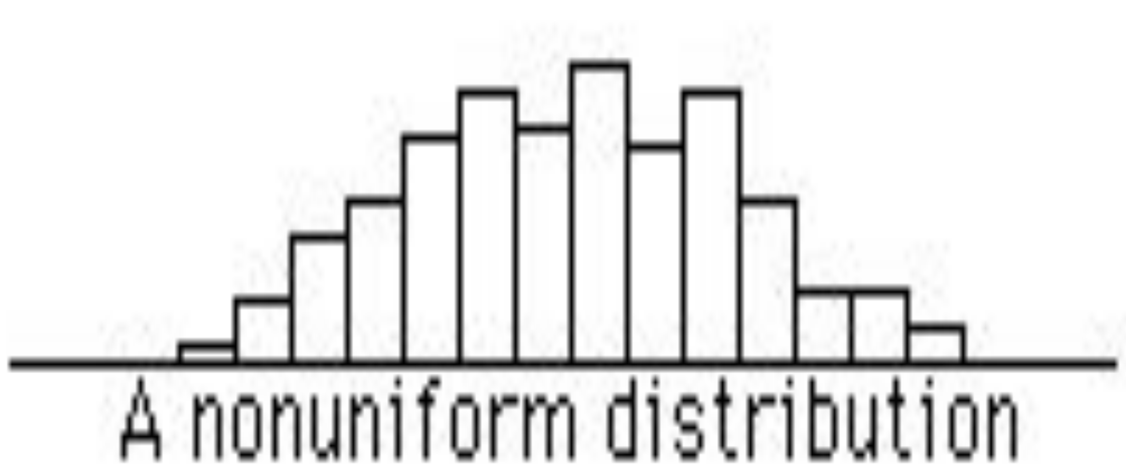
Secuencias en otras distribuciones

750098M Simulación computacional

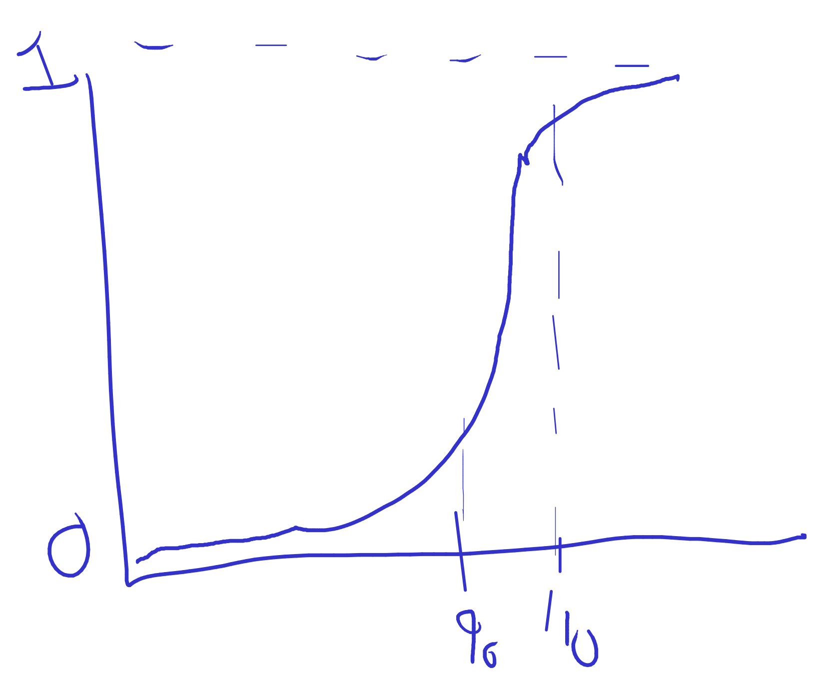
Contenido

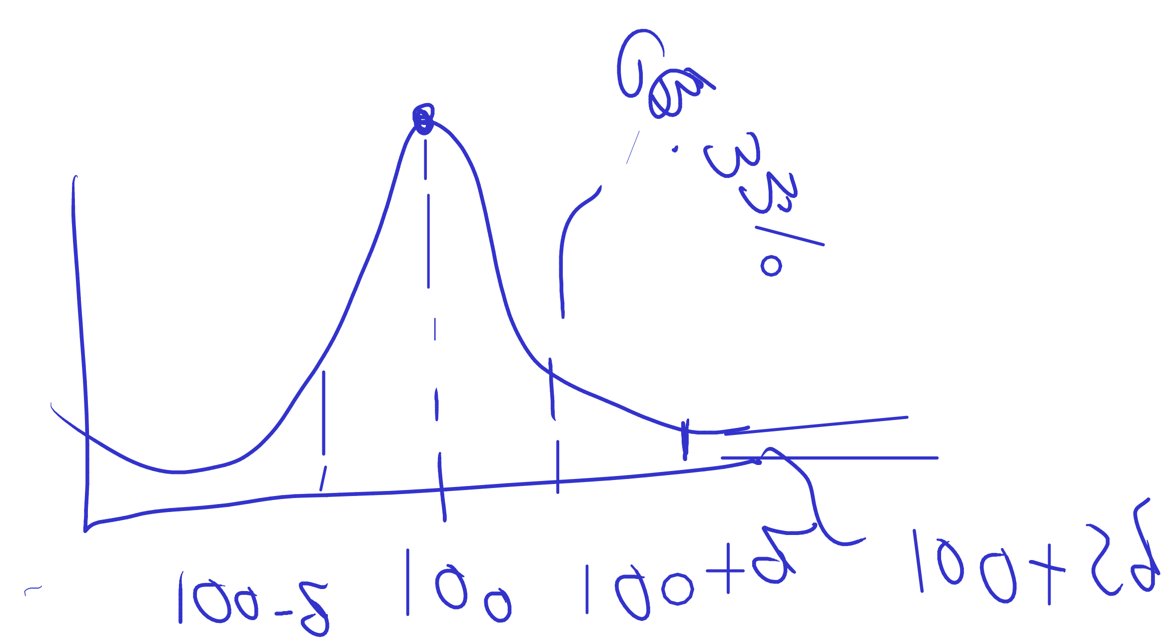




Se generan números con distribución uniforme U(0,1), se puede convertir a otra:

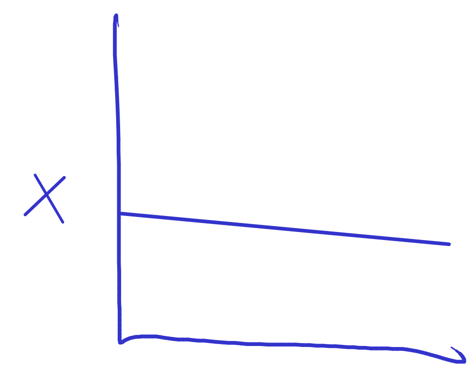
1. Método de la transformada inversa
2. Método de la convolución





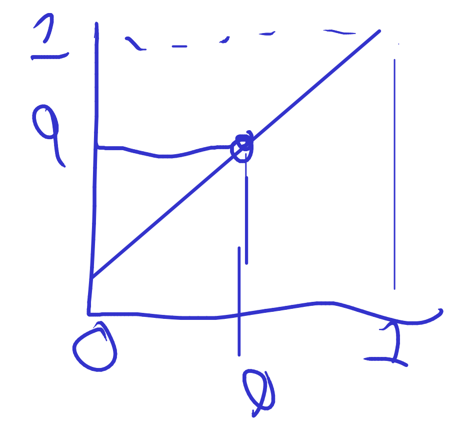
Método transformada inversa

Sea F la función de distribución acumulada de la distribución que se quiere generar.



Sea F(a) = r.

→ 0,58



Sea $p_u([0,r])$ la probabilidad del intervalo [0,r], suponiendo distribución la uniforme.

y pF ((−∞, a]) la probabilidad del intervalo (−∞, a], suponiendo distribución deseada:

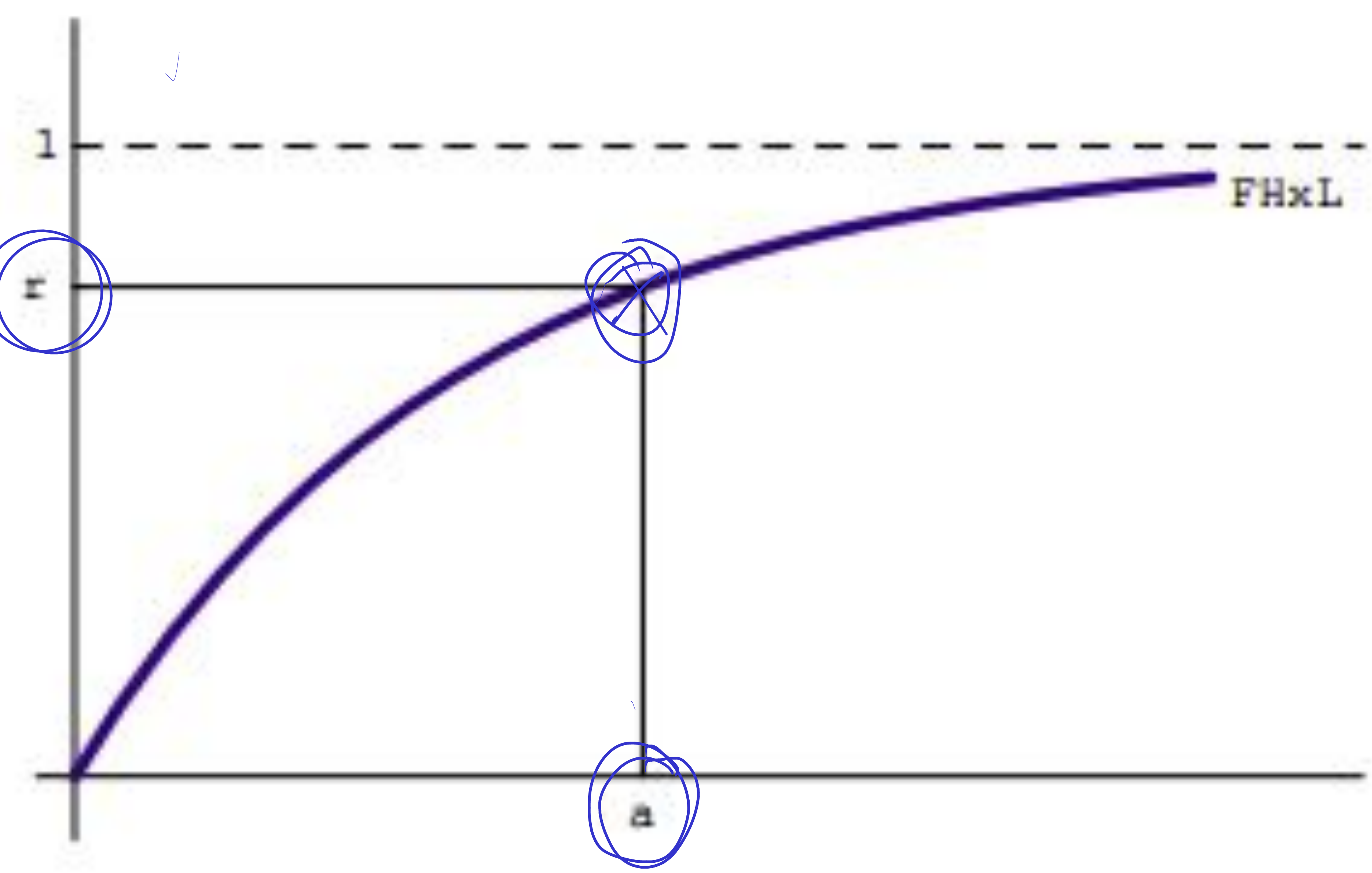
$$p_U([0,r]) = pF((-\infty, a]) = r$$

Método transformada inversa

Es decir: si ri es una secuencia de aleatorios ~ U[0, 1], entonces

$$a_i = F^{-1}(r_i)$$

es una secuencia de números que siguen la distribución deseada F.



F(0,5)
= 0,38
r ∈ U(0,1)

F(100)
= 5835%
0,5835
r = 0,5835

Distribución exponencial

La función de distribución acumulada de la distribución exponencial con parámetro λ está dada por:

$$\boxed{F(x) = 1 - e^{-\lambda x}}$$

$F(x) =$ $\lambda = \frac{1}{m}$

$$F(a) = r \Rightarrow \boxed{1 - e^{-\lambda a} = r} \Rightarrow a = F^{-1}(r) = -\frac{1}{\lambda}\ln(1-r)$$

Datos con distribución exponencial si r es uniforme en (0,1]

$$a = -\frac{1}{\lambda}\ln(r)$$

$1 - [0, 1)$

$[0, 1)$

Distribución exponencial

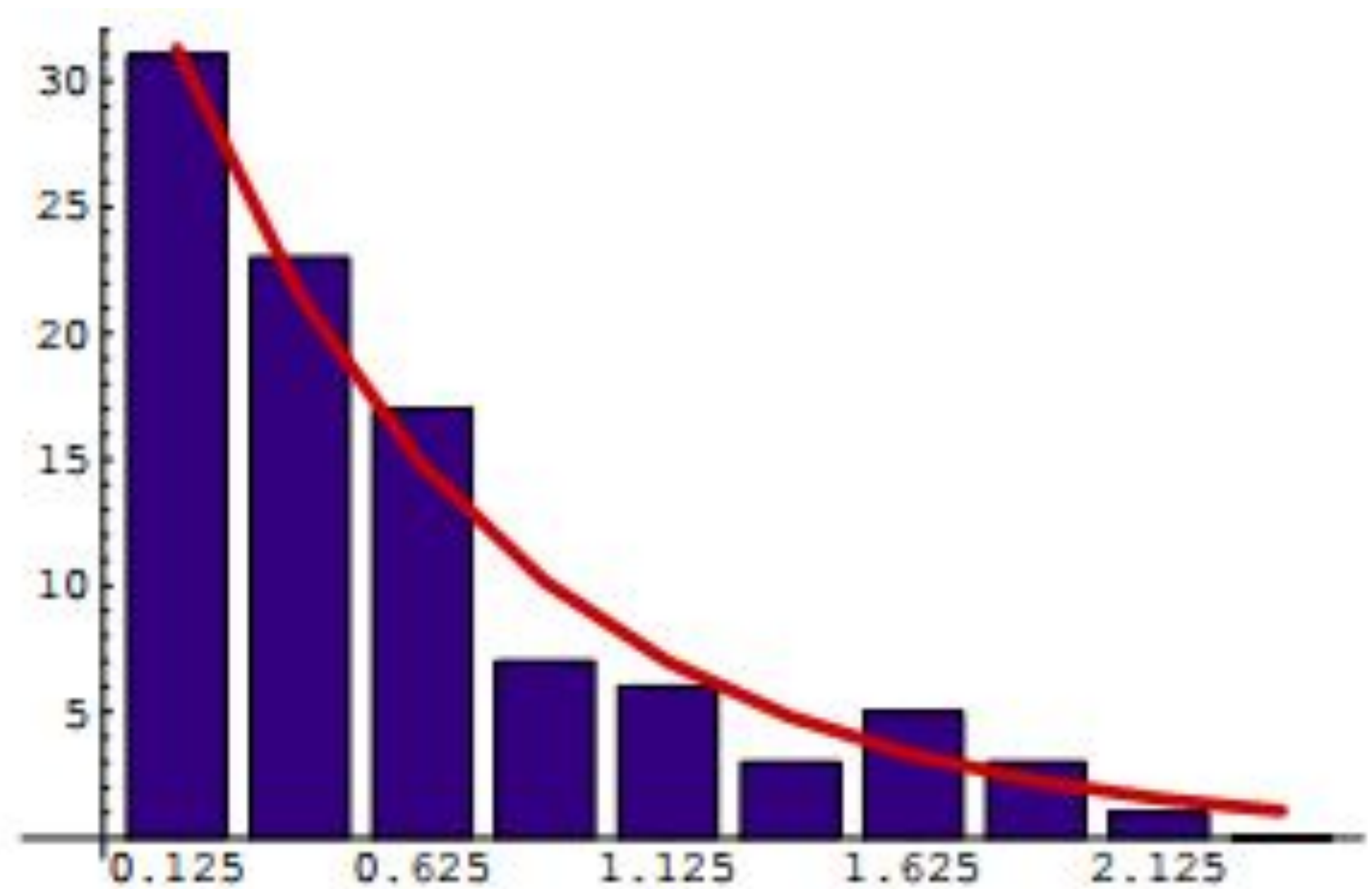
λ=1.5

$\to \frac{1}{\lambda} = \text{media}$

Para r = 0.5520 se obtiene: $a = F^{-1}(0{,}5520) = -\frac{1}{1{,}5}\ln\,(0{,}5520) = 0{,}3961$

r		a
0.5520		0.3961
0.4881		0.4782
0.7512	$\Rightarrow$	0.1907
0.3124		0.7756
0.5696		0.3752
0.7238		0.2155
0.9438		0.0386

Diagrama de frecuencia de 100 números pseudoaleatorios con distribución exponencial



Distribución uniforme U(A,B)

Fix

La función de distribución acumulada de la distribución uniforme U(A,B) está dada por:

$$F(x) = \frac{x-A}{B-A}$$

[0, 1)
A' B

[2, 5)

$\frac{4-2}{3} =$

[A, B)

$$F(a) = r => \frac{a-A}{B-A} = r => a = F^{-1}(r) = r(B-A) + A$$

B A

Es decir, simplemente se multiplica el aleatorio en [0,1] con la longitud B-A y se suma el punto inicial A de la uniforme U(A,B). Para r aleatorio en [0,1].

[0, 1)

[A, B)

$$a = (B-A)r + A$$

*Si r es uniforme en U(0,1]}

Distribución uniforme U(10,20)

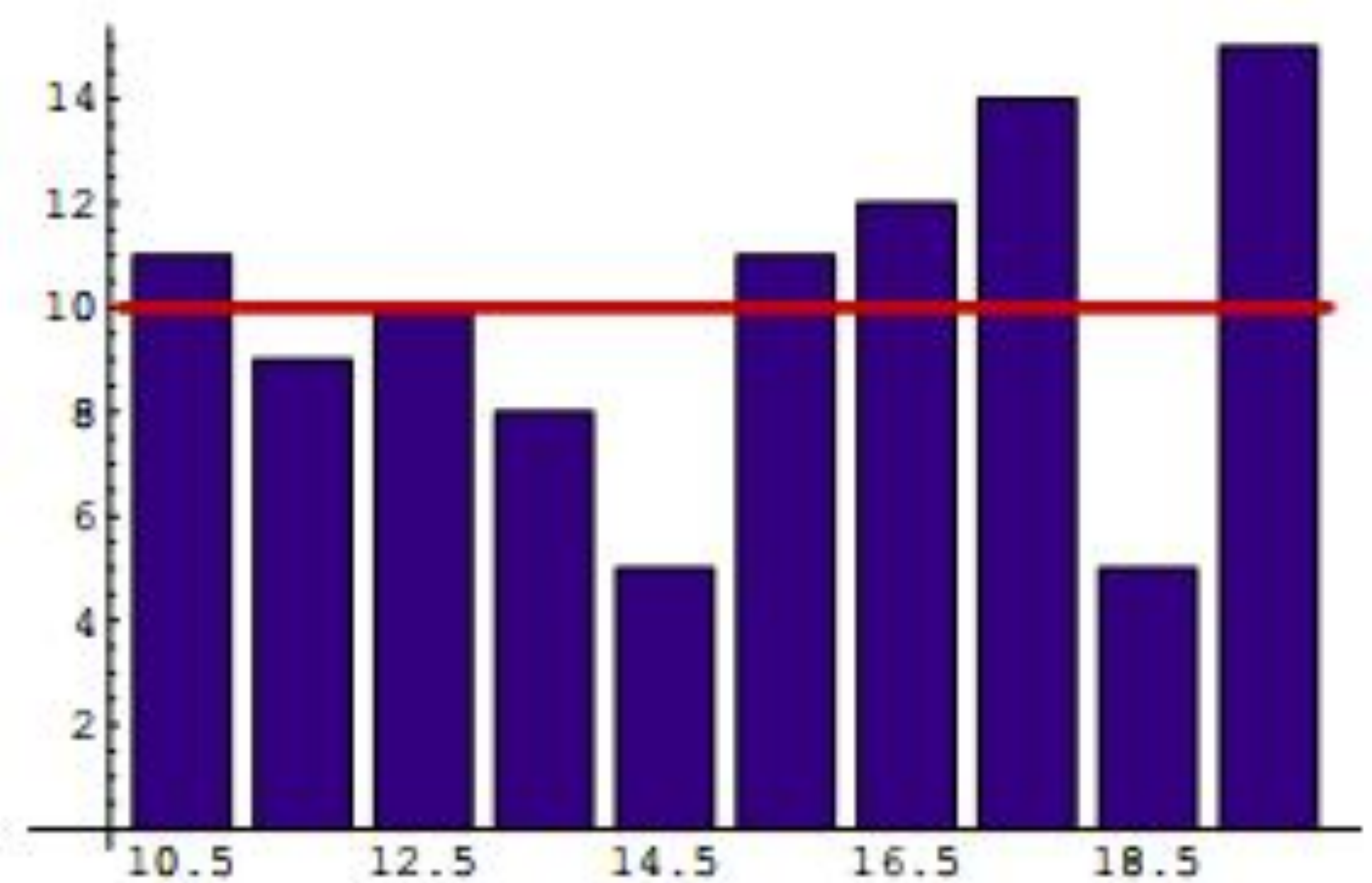
Para r=0.5520, se obtiene a=F^{-1}(0.5520)=0.5520*10 + 10

$A=10$ $B=20$

$\underbrace{(B-A)}_{10} r + \underbrace{A}_{10}$

r		a
0.5520		15.520
0.4881		14.881
0.7512	$\Rightarrow$	17.512
0.3124		13.124
0.5696		15.696
0.7238		17.238
0.9438		19.438

Diagrama de frecuencias de 100 números pseudoaleatorios con distribución U(10,20)



Distribución Normal Estandarizada N(0,1)

e^{-x^2}

La función de distribución acumulada de la normal:

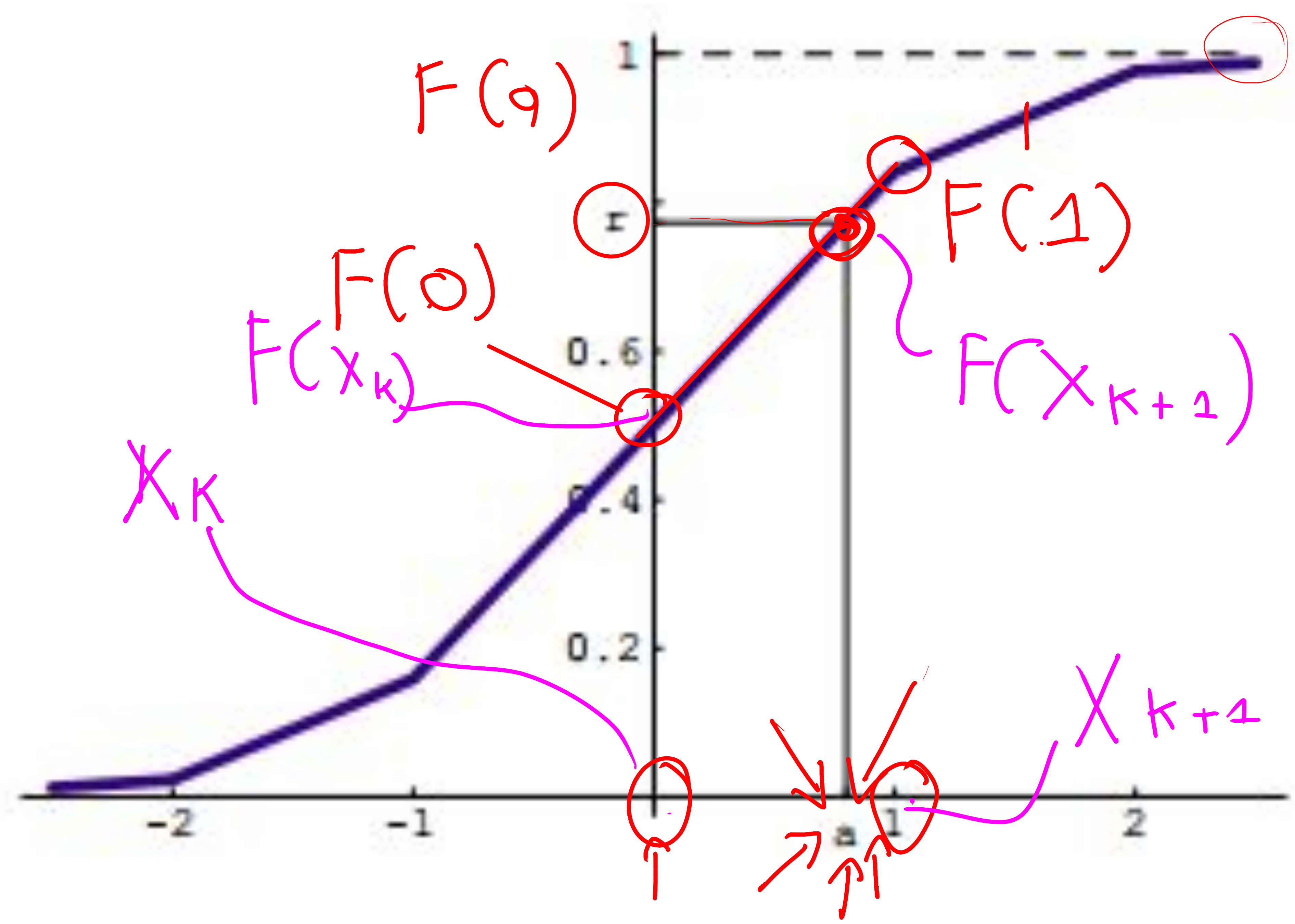
$$F_{N(0,1)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt\,,$$

Se tiene que determinar por integración numérica. Como resultado se obtiene una tabla de puntos, las valores intermedios se obtiene por interpolación lineal. A esta aproximación de la función de distribución acumulada se puede aplicar el método de la transformada inversa.

Distribución Normal Estandarizada N(0,1)

Ejemplo: Método de la inversa transformación, aplicada para la aproximación de F

x	$F(x)$
-3	0.0013
-2	0.0228
-1	0.1586
0	0.5
1	0.8413
2	0.9772
3	0.9987



En realidad se debe tabular F en muchos puntos
La tabla aqui presentada presenta un nivel de granulidad muy alto (saltos de 1 en 1)

$r = 0{,}257$

$h = \dfrac{0.5 - 0.257}{0.5 - 0{,}1586} = \underline{0{,}7118}$

$q = 0.7118 * -1 + (1 - 0{,}7118)\, 0$

$\boxed{q = -0{,}7118}$

-0.7	0.24196
-0.6	0.27425

$r = 0.257$

$h = \dfrac{0.27425 - 0.257}{0{,}27425 - 0.24196} = 0.53422$

$q = -0.7 \times 0.53422 + (1 - 0{,}53422) \times (-0.6)$

$\boxed{q = -0.6534}$

$$\boxed{r = 0.828}$$

x	$F(x)$
0,9	0,81594
1	0,84134

$$\boxed{\lambda = 0{,}5251} \qquad \lambda = \frac{F(x_{k+1}) - r}{F(x_{k+1}) - F(x_k)}$$

$$\boxed{Q = 0{,}9474} \qquad q = \underline{\lambda X_k} + (1-\lambda)\underline{X_{k+1}}$$

Distribución Normal Estandarizada N(0,1)

1. Para $r \in U(0,1)$ se busca en la tabla el intervalo que contiene r

$$[F(x_k), F(x_{k+1})]$$

2. r es combinación de los extremos del intervalo:

$$r = \lambda F(x_k) + (1 - \lambda) F(x_{k+1});$$

despejando λ se obtiene:

$$\lambda = \frac{F(x_{k+1}) - r}{F(x_{k+1}) - F(x_k)}$$

3. a es la combinación correspondiente de x_k y x_{k+1}:

$$a = \lambda x_k + (1 - \lambda) x_{k+1}.$$

Distribución Normal N(0,1)

La tabla de la distribución acumulada normal estandarizada para x <=0:

x	F(x)	x	F(x)	x	F(x)
-3	0.00135	-1.9	0.02872	-0.9	0.18406
-2.9	0.00187	-1.8	0.03593	-0.8	0.21186
-2.8	0.00256	-1.7	0.04457	-0.7	0.24196
-2.7	0.00347	-1.6	0.05480	-0.6	0.27425
-2.6	0.00466	-1.5	0.06681	-0.5	0.30854
-2.5	0.00621	-1.4	0.08076	-0.4	0.34458
-2.4	0.00820	-1.3	0.09680	-0.3	0.38209
-2.3	0.01072	-1.2	0.11507	-0.2	0.42074
-2.2	0.01390	-1.1	0.13567	-0.1	0.46017
-2.1	0.01786	-1	0.15866	0	0.50000
-2	0.02275				

Distribución Normal N(0,1)

La tabla de la distribución acumulada normal estandarizada para x >=0:

x	F(x)	x	F(x)	x	F(x)
0	0.50000	1	0.84134	2	0.97725
$F(x_k)$ 0.1	0.53983	1.1	0.86433	2.1	0.98214
$F(x_{k+1})$ 0.2	0.57926	1.2	0.88493	2.2	0.98610
0.3	0.61791	1.3	0.90320	2.3	0.98928
0.4	0.65542	1.4	0.91924	2.4	0.99180
0.5	0.69146	1.5	0.93319	2.5	0.99379
0.6	0.72575	1.6	0.94520	2.6	0.99534
0.7	0.75804	1.7	0.95543	2.7	0.99653
0.8	0.78814	1.8	0.96407	2.8	0.99744
0.9	0.81594	1.9	0.97128	2.9	0.99813
				3	0.99865

Distribución Normal N(0,1)

Ahora con una secuencia de números pseudoaleatorios con una distribución U(0,1), hacemos el proceso utilizando las tablas de la distribución normal acumulada estandarizada

Supongamos una secuencia r = {0.5520, 0.4884, ...}

$$\lambda = \frac{F(x_{k+1}) - r}{F(x_{k+1}) - F(x_k)}$$

$$r = 0.552$$

$$r \in [0.53983, 0.57926]$$

$$\lambda = \frac{0.57926 - 0.5520}{0.57926 - 0.53983} = 0.695$$

$$a = 0.695 * 0.1 + (1 - 0.695) * 0.2 = 0.130$$

$$\lambda \times F_k + (1 - \lambda) F_{x_{k+1}}$$

$$r = 0.4881$$

$$r \in [0.46017, 0.5]$$

$$\lambda = \frac{0.5 - 0.4881}{0.5 - 0.46017} = -0.030$$

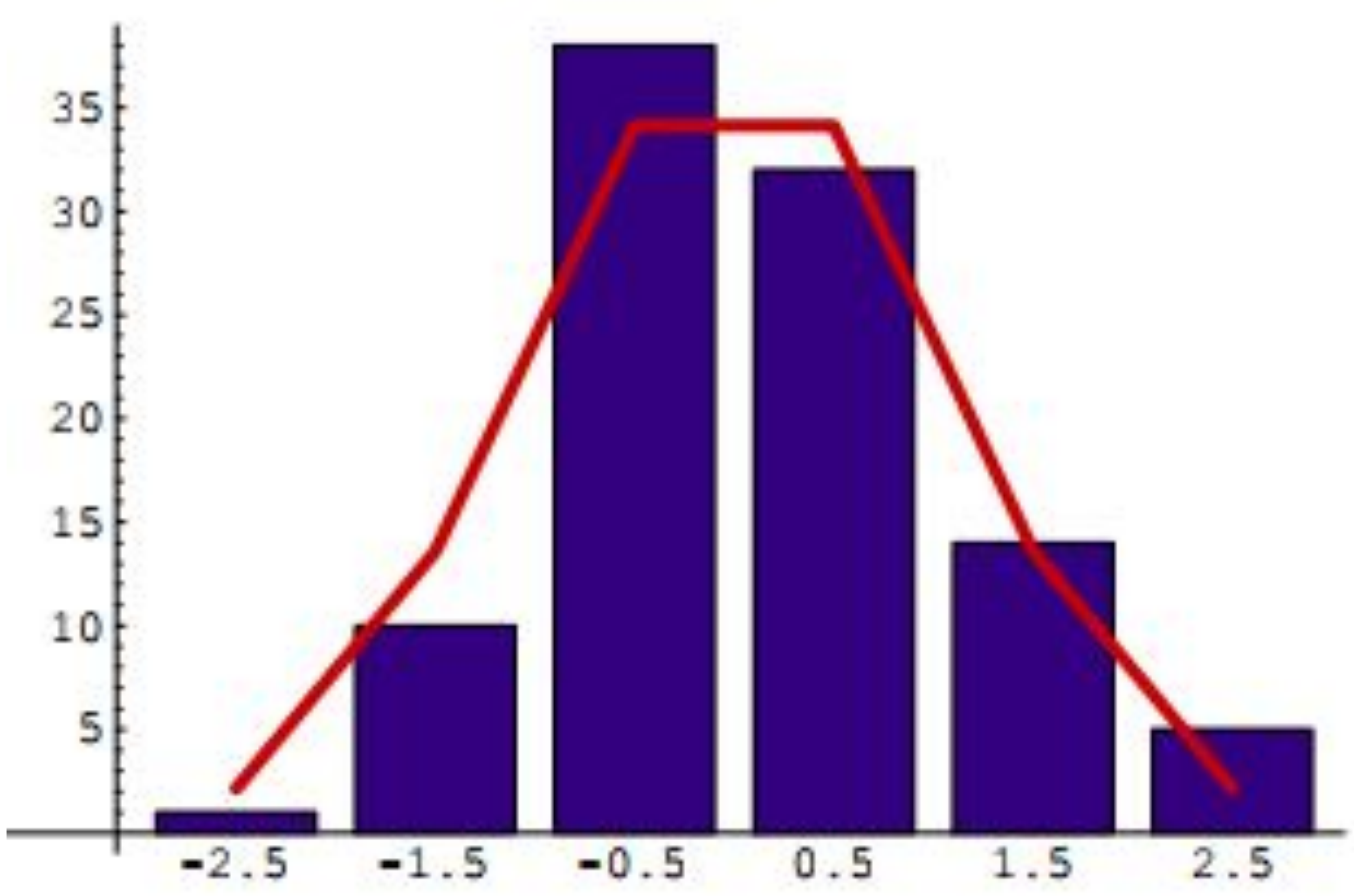
$$a = 0.299 * (-0.1) + (1 - 0.299) * 0 = 0.130$$

Distribución Normal N(0,1)

r	$F(x_k)$	$F(x_{k+1})$	λ	intervalo de a		a
0.552	0.53983	0.57976	0.695	0.1	0.2	0.130
0.4881	0.46017	0.5	0.299	-0.1	0	-0.030
0.7512	0.72575	0.75804	0.212	0.6	0.7	0.679
0.3124	0.30854	0.34458	0.893	-0.5	-0.4	-0.489
0.5696	0.53983	0.57926	0.245	0.1	0.2	0.176
0.7238	0.69146	0.72575	0.057	0.5	0.6	0.594
0.9438	0.93319	0.9452	0.117	1.5	1.6	1.588

Distribución Normal N(0,1)

Diagrama de frecuencias de 100 números pseudoaleatorios con la distribución que aproximaN(0, 1):



1 Obsérvese que esta aproximación no es buena dado que se usa muy pocos puntos. Para aplicaciones prácticas se deberían usar mucho más puntos tabulados.

2 Existen otros métodos para generar datos con distribución N(0,1) que no requieren la tabulación de la normal (ver más adelante).

Otras distribuciones Normales

Si se tiene datos que siguen la distribución normal estandarizada, se pueden convertir a datos con distribución N(μ, σ), sumando la media μ(traslada la distribución) y multiplicando por σ (anchura de la distribución)

Datos con distribución Normal N(μ, σ):

$$a_{N(\mu,\sigma)} = a_{N(0,1)} \cdot \sigma + \mu$$

Ejemplo: distribución N(10,2)

Convertir los datos del ejemplo anterior a datos con N(10, 2)

$0.152359 \times 2 + 10$

r	$a_{N(0,1)}$	$a_{N(10,2)}$
0.5520	0.152359	10.3047
0.4881	-0.0348667	9.93027
0.7512	0.736009	11.472
0.3124	-0.549663	8.90067
0.5696	0.203926	10.4079
0.7238	0.655728	11.3115
0.9438	1.75423	13.5085

Distribuciones discretas

Aunque la función de distribución acumulada de una distribución discreta no tiene inversa, se puede aplicar un procedimiento parecido al método de la transformada inversa para generar distribuciones discretas.

Para la distribución discreta dada por sus valores y su probabilidad (v_i, p_i) se usa la probabilidad acumulada pa_i en lugar de la función de distribución F:

v_i	p_i	pa_i
v_1	p_1	p_1
v_2	p_2	$\sum_{i=1}^{2} p_i$
$\vdots$	$\vdots$	$\vdots$
v_n	p_n	$\sum_{i=1}^{n} p_i = 1$

Para r ~ U[0, 1] se define:

$$a = v_i \text{ si } r \in [pa_{i-1}, pa_i)$$

Distribución discreta empírica

v_i	p_i	pa_i
1	0.2	0.2
5	0.1	0.3
8	0.5	0.8
10	0.2	1

Por ejemplo:

r	a
0.5520	8
0.4881	8
0.7512	8
0.2124	5
0.5696	8
0.0723	1
0.9438	10

Para $r \sim U[0,1]$ se define:

$$a = \begin{cases} 1 & \text{si } r < 0{,}2 \\ 5 & \text{si } 0{,}2 \leq r < 0{,}3 \\ 8 & \text{si } 0{,}3 \leq r < 0{,}8 \\ 10 & \text{si } 0{,}8 \leq r < 1 \end{cases}$$

Distribución Bernoulli

$p = 0.8$

La generación de datos con distribución Bernoulli con parámetro p es muy sencillo y rápido

Datos con distribución Bernoulli

$$a = \begin{cases} \text{éxito} & \text{si } r < p \\ \text{fracaso} & \text{en el otro caso} \end{cases}$$

Ejemplo para $p = 0{,}7$

r	a
0.5520	éxito
0.4881	éxito
0.7512	fracaso
0.3124	éxito
0.5696	éxito
0.7238	fracaso
0.9438	fracaso

Distribución Binomial

Ejemplo: $n = 6$, $p = 0{,}7$

Definición de las probabilidades, usando

$$p(i) = \binom{n}{i} 0{,}7^i \, 0{,}3^{n-i}, \; 0 \le i \le n$$

$C(n, i)$ $\quad \underset{E}{p^i} \, \underset{F}{(1-p)^{n-i}}$

valor	probabilidad	probabilidad acumulada
0	0.000729	0.000729
1	0.010206	0.010935
2	0.059535	0.07047
3	0.18522	0.25569
4	0.324135	0.579825
5	0.302526	0.882351
6	0.117649	1

Datos a con distribución binomial $(0{,}3, 6)$:

r	a
0.5520	4
0.4881	4
0.7512	5
0.3124	4
0.5696	3
0.7238	5
0.9438	6

$n = 7$ $\quad p = 0.3$ $\quad (nCi) \times p^{i} \times (1-p)^{n-i}$

n	P_n	P_ϕ
0	0,0823543	0,0823543
1	0,2470629	0,3294172
2	0,317653	0,6470702
3	0,2268945	0,8739647
4	0,0972405	0,9712052
5	0,0250047	0,9962099
6	0,0035721	0,999782
7	0,0002187	1

$r = 0,48$

$\boxed{q = 2}$

Distribución Binomial

El método de la inversa transformada para generar una binomial (n, p) es eficiente, si n es pequeño.

Si n es grande (por ejemplo n = 200) se tiene que calcular muchas probabilidades p(i) y probabilidades acumuladas pa(i).

En este caso se recomienda otro método de generación: el método de convolución para generar la distribución binomial.

Método de convolución

Principio:

1. Se expresa la variable aleatoria con distribución deseada como función de otras con distribuciones que se generan fácilmente.
2. Se generan estas variables aleatorias y se aplica la función.
3. El resultado es un valor de la variable aleatoria deseada

Distribución binomial

con el método de convolución

Dado un experimento Bernoulli con probabilidad de éxito p, la variable aleatoria X: número de éxitos en *n* repeticiones del experimento Bernoulli sigue una distribución binomial con parámetros *p* y *n*.
Es decir, generando *n* números con distribución Bernoulli, y contando el número de éxitos se obtiene un número con distribución binomial.

Ejemplo

p=0.7,n=6

E	F	F $r_1, r_2, \ldots, r_6$	E	F	E	a
0.0530374	0.808378	0.903386	0.538051	0.875958	0.513395	3
0.932824	0.261447	0.879235	0.456081	0.851364	0.566246	3
0.926326	0.648505	0.0858267	0.603867	0.297316	0.126606	5
0.990079	0.367285	0.693317	0.125736	0.434217	0.300056	5

Distribución normal

por el método de convolución

El promedio de *n* datos con igual distribución cualquiera tiende a una distribución normal (teorema central de límite). El método de convolución usa este teorema para generar datos normalmente distribuidos con promedio de datos ~ U[0, 1].
La distribución uniforme U[0, 1] tiene media 1/2 y varianza 1/12, por eso: la suma de 12 datos ~ U[0, 1] tiene media 6 y varianza 1, por eso

$$a = \left(\sum_{i=1}^{12} r_i \right) - 6 \sim N[0, 1]$$

Ejemplo

Distribución normal

N(30, 2)

$r_1, r_2, \ldots, r_{12}$				a
0.0530374	0.808378	0.903386	0.538051	
0.875958	0.513395	0.851364	0.566246	
0.932824	0.261447	0.879235	0.456081	1.6394
0.926326	0.648505	0.0858267	0.603867	
0.297316	0.126606	0.434217	0.300056	
0.990079	0.367285	0.693317	0.125736	-0.400863